THE UNIVERSITY
of EDINBURGH

Exploring natural and engineered resistance to potyviruses

Douglas Pyott

Doctor of Philosophy

The University of Edinburgh

2017

Table of contents

Table of contents ..... **ii**

Declaration ..... **x**

Acknowledgements ..... **xi**

Abstract ..... **xii**

Abbreviations ..... **xiii**

Chapter 1: General introduction

1.1 The importance of potyvirus research **1**

1.1.1 The impact of potyvirus diseases to global agriculture 1

1.1.2 Potato Virus Y 2

1.1.3 Plum Pox Virus 3

1.1.4 Turnip Mosaic Virus 4

1.2 The potyvirus genome and molecular structure **6**

1.3 Potyvirus transmission **8**

1.4 The potyvirus translation-replication process **10**

1.4.1 Co-ordination of viral processes 10

1.4.2 Virus Replication Complexes 11

1.5 Movement of potyviruses through the plant **12**

1.5.1 Cell-cell (local) movement 12

1.5.2 Long distance (systemic) movement 13

1.6 Genetic resistance to potyvirus infection **14**

1.6.1 General principles of genetic resistance to pathogens 14

1.6.2 Dominant potyvirus resistance genes 15

1.6.3 Recessive potyvirus resistance genes 17

1.7 Immune responses to potyvirus infection **18**

1.8 RNA silencing **19**

1.8.1 General features of RNA silencing 19

1.8.2 Exogenous RNA silencing 22
1.8.3 Endogenous RNA silencing 23
1.8.4 Viral suppressors of RNA silencing 25
1.8.5 Cross-talk in viral RNA silencing pathways 26

1.9 Infection of *Arabidopsis thaliana* with Turnip Mosaic Virus as a model system for potyvirus infection. **28**

Chapter 2: Materials and methods

2.1 Genetic material **30**
2.1.1 Details and sources of plasmid vectors *30*
2.1.2 Details and sources of transgenic/mutant plant material *31*
2.2 Plant growth conditions **31**
2.3 Bacterial plasmid purification **32**
2.4 Vector cloning **32**
2.4.1 pDe-Ubi:SpCas9-U6:sgRNA_AteIF(iso)4E 32
2.4.2 pDe-35S:SpCas9-U6:sgRNA_GFP 32
2.4.3 pK7FWG2-35S:AsCpf1 and pK7FWG2-35S:LbCpf1 32
2.4.4 pK7FWG2-U6:AscrRNA_GFP and pK7FWG2-U6:LbcrRNA_GFP 33
2.4.5 pK7FWG2-Ubi:LbCpf1-Oleosin:RFP 33
2.4.6 pK7FWG2-Ubi:SpCas9-Oleosin:RFP 33
2.4.7 pK7FWG2-Ubi:LbCpf1-U6:crRNA_arf19_1 and pK7FWG2-Ubi:LbCpf1-U6:crRNA_arf19_2 34
2.4.8 pK7FWG2-Ubi:SpCas9-Oleosin:RFP-U6:sgRNA_arf19_1 and pK7FWG2-Ubi:SpCas9-Oleosin:RFP-U6:sgRNA_arf19_2 34
2.4.9 pK7FWG2-Ubi:LbCpf1-U6:crRNA_GFP_1-U6:crRNA_GFP_2-U6:crRNA_Donor_1-U6:crRNA_Donor_2-Donor_DNA 35
2.5 Bacterial transformations **38**
2.5.1 Transformation of Escherichia coli cells with plasmid vectors 38
2.5.2 Transformation of Agrobacterium tumefaciens cells with plasmid vectors 38

2.6 Plant transformation **39**
2.6.1 Floral dipping 39
2.6.2 Transient infiltration 39
2.7 Selection of T_1 seed **39**
2.7.1 Screening T_1 seedlings by BASTA selection 39
2.7.2 Screening T_1 seeds by FAST-red fluorescence sorting 40
2.8 TuMV-GFP rub-inoculations **40**
2.9 UV imaging of TuMV-GFP **40**
2.10 Plant DNA/RNA purification **41**
2.11 PCR and gel imaging **41**
2.11.1 PCR cycle conditions 41
2.11.2 Gel imaging 42
2.12 qRT-PCR **42**
2.12.1 DNase treatment 42
2.12.2 cDNA synthesis 43
2.12.3 SYBR qRT-PCR reaction 43
2.12.4 Primer design and validation 43
2.12.5 SYBR qRT-PCR analysis 44
2.13 Northern blotting **46**
2.14 Western blotting **47**
2.15 Sanger sequencing **48**
2.15.1 PCR and product purification 48
2.15.2 BigDye reaction and capillary sequencing 48
2.15.3 Sequence analysis 48
2.16 T7 endonuclease assay **49**
2.17 Physiological analysis of *eIF(iso)4E* mutants **50**
2.17.1 Flowering time measurements 50
2.17.2 Dry weight measurements 50
2.18 *arf19* hypocotyl assay **50**

2.18.1 Seed sterilisation 50
2.18.2 Growth conditions 50
2.18.3 Hypocotyl measurements 51
2.19 Statistical analyses **51**
2.19.1 Standard error calculations 51
2.19.2 Analysis of Variance (ANOVA) and post-hoc comparison of means 52
2.19.3 Boxplot/Scatterplots 52
2.20 Figure mounting **52**
2.21 Buffer/solution recipes **52**
2.21.1 Luria Bertani media (LB media) 52
2.21.2 Yeast Extract Peptone media (YEP media) 53
2.21.3 Infiltration Media 53
2.21.4 1M Phosphate buffer (pH 7.0) 53
2.21.5 Tris Borate EDTA (TBE) buffer (10X) 53
2.21.6 Tris EDTA (TE) buffer 53
2.21.7 Total Nucleic Acid Extraction Buffer (TNA EB) 54
2.21.8 RNA loading dye (2X) 54
2.21.9 15% polyacrylamide denaturing gel 54
2.21.10 Saline Sodium Citrate (SSC) buffer (20X) 54
2.21.11 Hybridisation buffer 54
2.21.12 Protein Extraction Buffer (Protein EB) 55
2.21.13 Tris Buffered Saline Tween Triton (TBSTT) 5% Milk 55
2.21.14 2,4-diclorophenoxyacetic acid (2,4-D) selection media (pH 5.7) 55

Chapter 3: Environmental influences on viral RNA silencing

3.1 Chapter introduction **56**
3.1.1 The importance of environmental conditions on the outcome of plant diseases. 56
3.1.2 Thermotherapy as a method for cleaning virus infected plant material 56
3.1.3 Virus-induced RNA silencing is promoted by increasing temperatures 59
3.1.4 Genetic components of temperature-activated viral RNA silencing 60
3.1.5 Effects of temperature on potyvirus infection 61

3.1.6 Effects of light on potyvirus infection 62

3.2 Chapter aims **64**

3.3 *In vivo* monitoring of early TuMV-GFP infection at different temperatures in local leaves of *Nicotiana benthamiana* **65**

3.4 Analysis of early TuMV-GFP infection at different temperatures in local leaves of *Arabidopsis thaliana* **70**

3.5 Analysis of TuMV-GFP infection and recovery at different temperatures in systemic tissue of *Arabidopsis thaliana* **75**

3.6 Genetic factors responsible for activation of viral PTGS at high temperatures **77**

3.7 *In vivo* monitoring of early TuMV-GFP infection at different light intensities in local leaves of *Nicotiana benthamiana* **81**

3.8 Analysis of early TuMV-GFP infection at different temperatures in local leaves of *Arabidopsis thaliana* **84**

3.9 Analysis of TuMV-GFP infection and recovery at different temperatures in systemic tissue of *Arabidopsis thaliana* **86**

3.10 Genetic factors responsible for activation of viral PTGS at low light intensity **88**

3.11 Combined effects of temperature and light intensity shifts on the expression profiles of *RDR1* and *RDR6* **91**

3.12 Chapter discussion **93**

3.12.1 Temperature increases are likely to affect multiple aspects of the potyvirus lifecycle and host defence responses 93

3.12.2 Low light intensity appears to reduce TuMV synthesis and enhance anti-TuMV PTGS by induction of RDR1 and RDR6 95

3.12.3 Potential for antiviral 'light therapy' as an alternative or complementary approach to thermotherapy 96

3.12.4 Different antiviral defence responses may be alternately activated/repressed different environmental conditions 96

3.12.5 Biological relevance of the light and temperature shift experiments 97

Chapter 4: The impact of inoculation time on potyvirus susceptibility

4.1 Chapter introduction **99**

4.1.1 General features of circadian rhythms 99

4.1.2 Molecular mechanisms of the plant circadian clock 100

4.1.3 Functional relevance of the plant circadian clock 102

4.2 Chapter aims **105**

4.3 Rhythmic transcription of antiviral RNA silencing genes **105**

4.4 DCL4 protein abundance is rhythmic under circadian conditions **112**

4.5 Susceptibility to TuMV-GFP infection exhibits diurnal and circadian rhythmicity in *Arabidopsis* **115**

4.6 The role of antiviral *DCL* genes in night-time suppression of TuMV-GFP infection **118**

4.7 eIF(iso)4E is a candidate host factor which may generate the time-of-day rhythms in TuMV-GFP infection **122**

4.8 Chapter discussion **124**

4.8.1 Virus susceptibility in plants varies over day/night cycles in a circadian-clock dependent manner 124

4.8.2 Antiviral RNA silencing operates primarily as a 'night-time' viral defence pathway. 124

4.8.3 Rhythmic expression of host factors may be the underlying mechanism for time-of-day differences in virus susceptibility 126

Chapter 5: Engineering potyvirus resistance in *Arabidopsis* using CRISPR/Cas9

5.1 Chapter introduction **128**

5.1.1 Eukaryotic translation 128

5.1.2 Viral usurpation of host translation factors 128

5.1.3 CRISPR/Cas9 as a tool for genome editing 130

5.2 Chapter aims **131**

5.3 Site specific mutation of *eIF(iso)4E* by transgenic expression of a sgRNA guided Cas9 ..... **131**

5.4 Segregation of the induced mutation from the transgene in the T_2 generation **134**

5.5 The induced mutation in *eIF(iso)4E* confers complete resistance to TuMV **141**

5.6 The *eIF(iso)4E* mutants show no growth defects compared to wild type plants, when grown under standard growth conditions **145**

5.7 Chapter discussion **146**

5.7.1 Introducing R genes into crops by plant breeding 146

5.7.2 Introducing R genes into crops using biotechnological approaches 147

5.7.3 The potential for CRISPR/Cas9 technology for generating virus resistance in crops. 148

5.7.4 Technological outlook: the potential to copy natural mutations by genomic editing.. 150

Chapter 6: Exploring new CRISPR methods

6.1 Chapter introduction **152**

6.1.1 Recombinant forms of Cas9 expand its biotechnological repertoire 152

6.1.2 Natural diversity of Cas9 orthologues 153

6.1.3 Cpf1: a novel Cas orthologue with distinct features to Cas9 154

6.1.4 Using CRISPR systems for targeted gene replacement 157

6.2 Chapter aims **161**

6.3 Testing Cpf1 for *in planta* genome editing **161**

6.3.1 Selecting a suitable Cpf1 orthologue for in planta genome editing 161

6.3.2 Design of a phenotypic screen to compare the efficiencies of SpCas9 and LbCpf1 for inducing heritable mutations 162

6.3.3 Design of ARF19-targeting sgRNAs and crRNAs for CRISPR/Cas9/Cpf1-induced mutations 164

6.3.4 Testing the arf19 haploinsufficient phenotype 167

6.3.5 Scoring CRISPR/Cas9/Cpf1-induced mutations in the T_1 generation 171

6.3.6 Assessing the determinants for efficient Cpf1 crRNAs 177

6.4 Microhomology-mediated targeted DNA replacement using Cpf1 ..... **184**

6.5 Chapter discussion **191**

6.5.1 The LbCpf1 orthologue can induce heritable mutations in plants 191

6.5.2 The function determinants of LbCpf1 crRNAs are yet to be adequately deduced 192

6.5.3 Exploring the potential for microhomology directed sequence knock-ins 193

Chapter 7: Conclusions and future outlook

7.1 The importance of potyvirus research in the context of an expanding global population **195**

7.2 The effects of temperature on the potyvirus synthesis/degradation equilibrium **195**

7.3 'Skototherapy' and biological relevance of the influence of light signals on viral RNA silencing defence pathways **197**

7.4 The potential for biotechnological advances to create novel potyvirus resistant germplasm **198**

7.5 Concluding Remarks **201**

References **202**

Declaration

I declare that the work presented in this thesis is my own. Any contributions from other parties is explicitly acknowledged. It has not been submitted in any previous application for a degree or qualification. Chapter 5 has been excerpted from a publication in Molecular Plant Pathology (Pyott *et al.*, 2016), and sections of Chapter 1 have been excerpted from a publication in Plant Biotechnology Journal (Pyott & Molnar, 2015). Both of these publications were written by myself.

Acknowledgements

It was a pleasure to have spent the last four years at the Institute of Molecular Plant Sciences (IMPS) at the University of Edinburgh. I would like to thank all members of IMPS for their companionship, support, and friendship. I am grateful for the support of my PhD supervisors; Dr Attila Molnar, Prof. Karen Halliday, and Dr Lesley Torrance. Specifically, I would like to thank Dr Attila Molnar for providing me with extensive lab training and frequent stimulating discussions. I would also like to thank my thesis committee (Prof. Andrew Hudson and Dr Steven Spoel) for their continued support and input, and Dr Graham Cowan for technical assistance and advice. My experience in the last four years would not have been nearly as enjoyable without the close friendships formed with fellow lab members, namely: Dr Kirsten Knox, Dr Danaé Paultre, Karen Bell, Fei Yue, Aron Ferenczi, Lila Grandgeorge, and Muriel Monteiro. Finally, I am extremely grateful to my family for their constant support and guidance.

Abstract

Viruses are ubiquitous in natural growth environments and cause severe losses to crop yields, globally. Approximately 30% of plant viruses described to date are grouped within the family *Potyviridae*, making it one of the largest plant virus families. Furthermore, certain potyvirus species can cause devastating diseases in several agriculturally and economically important crops. Hence, gaining insight into potyvirus resistance and recovery mechanisms in plants is an important research focus. This thesis firstly explores how environmental cues can modulate the activity of a central form of viral defence, namely RNA silencing. Specifically, high temperatures and low light intensities were found to increase the efficacy of viral RNA silencing in *Arabidopsis*, resulting in recovery from infection by Turnip Mosaic Virus. The biological context and potential for agricultural exploitation of these phenomena are discussed. Secondly, this thesis explores the ability to engineer resistance alleles using the latest genome editing techniques. Specifically, resistance to Turnip Mosaic Virus was successfully engineered in *Arabidopsis* by CRISPR/Cas9-induced deletion of a known susceptibility factor eIF(iso)4E. Biotechnological methods to implement this proof of concept research in crop species were also investigated.

Abbreviations

AGO – ARGONAUTE

ANOVA – analysis of variance

avr – avirulence

BCMV – Bean Common Mosaic Virus

bp – base pair

Cas – CRISPR-associated

CBS – CCA1 binding element

CCA1 – CIRCADIAN CLOCK ASSOCIATED 1

CCA1-OX – CCA1 overexpressor

CI – cylindrical inclusion

CK2 – CASEIN KINASE II

CMV – Cucumber Mosaic Virus

CO - CONSTANS

CP – coat protein

Cpf1 – CRISPR from *Prevotella* and *Francisella* 1

CPIP – coat protein interacting protein

CRIPSR – clustered regularly spaced short palindromic repeats

crRNA – CRISPR RNA

CymRSV – Cymbidium Ringspot Virus

DI-RNA – defective interfering RNA

DCL – DICER-LIKE

dpi – days pots inoculation

dps – days post shift

DSB – double stranded breaks

DSRB – double strand break repair

dsRNA – double-stranded RNA

EC – evening complex

EE – *EVENING ELEMENT*

eIF – eukaryotic translation initiation factor

ELISA – enzyme-linked immunosorbent assay

EM – electron microscope

EMS – ethyl methanesulfonate

ER – endoplasmic reticulum

ERES – ER-exit sites

ETI – effector-triggered immunity

GFP – GREEN FLUORESCENT PROTEIN

HC-Pro – helper component-protease

HDR – homology directed repair

hcsiRNA – heterochromatic siRNA

HR – hypersensitive response

HSD – honest significant difference

HSE – *HEAT SHOCK ELEMENT*

HSP70 – HEAT SHOCK PROTEIN 70

indel – insertion/deletion

IR – inverted repeat

KI – knock-in

KO – knock-out

LMV – Lettuce Mosaic Virus

LHY – LATE ELONGATED HYPOCOTYL

LRE – *LIGHT RESPONSIVE ELEMENT*

MDMV – maize dwarf mosaic virus

MATH – meprin and TRAF homology

miRNA – microRNA

MMEJ – microhomology-mediated end joining

MP – movement protein

Myb - myoblastosis

NB-LRR – nucleotide-binding leucine rich repeat

NHEJ – non-homologous end joining

NLS – nuclear localisation signal

NMD – nonsense mediated decay

natsiRNA – natural antisense siRNA

nt – nucleotide

ORF – open reading frame

PABP – polyA binding protein

PAM – proto-spacer adjacent motif

PAMP – pathogen associated molecular pattern

PCD – programmed cell death

PD – plasmodesmata

PIPO – pretty interesting potyviral protein

PPV – Plum Pox Virus

PRR – PSEUDO RESPONSE REGULATOR

PRSV – Papaya Ringspot Virus

PTGS – post-transcriptional gene silencing

PTI – PAMP-triggered immunity

PTNRD – potato tuber necrotic ringspot disease

PVX – Potato Virus X

PVY – Potato Virus Y

RISC – RNA induced silencing complex

RDR – RNA dependent RNA polymerase

RFP – RED FLUORESCENT PROTEIN

R-gene – resistance gene

RNAi – RNA interference

RTM – restricted TEV movement

ROS – reactive oxygen species

SA – Salicylic acid

SAR – systemic acquired resistance

SDSA – synthesis dependent strand annealing

SEL – size exclusion limit

SEM – standard error in the mean

sgRNA – single guide RNA

siRNA – short interfering RNA

SPFMV – sweet potato feathery mottle virus

sRNA – small RNA

SSB – single stranded break

+ssRNA – positive sense single-stranded RNA

-ssRNA – negative sense single-stranded RNA

TALEN – transcription activator-like effector nuclease

tasi-RNA – trans-acting short interfering RNA

TBSV – Tomato Bushy Stunt Virus

TCV – Turnip Crinkle Virus

TEV – tobacco etch virus

TGS – transcriptional gene silencing

TILLING – targeting induced local lesions in genomes

TMV – Tobacco Mosaic Virus

TNV – Tobacco Necrosis Virus

TPD – tomato plant decline

tracrRNA – trans-acting CRISPR RNA

TuMV – Turnip Mosaic Virus

USDA – US Department of Agriculture

UTR – un-translated region

UV – ultraviolet

vsiRNA – viral short interfering RNA

VPg – viral protein genome-linked

VRC – viral replication complex

VSR – viral suppressor of RNA silencing

WT – wild type

ZFN – zinc finger nuclease

ZT – zeitgeber time

2,4-D – 2,4-diclorophenoxyacetic acid

Chapter 1: General introduction

1.1 The importance of potyvirus research

1.1.1 The impact of potyvirus diseases to global agriculture

The global population is predicted to reach 9 billion by the middle of this century (Godfray *et al.*, 2010) which is anticipated to be accompanied by an approximately 2.5% annual increase in food demand, per capita (Tilman *et al.*, 2011). Consequently, there will be an appreciable and annually compounding demand for increased food production in the coming years. Estimates state that food production between the years 2005 and 2050 will need to at least double to meet this rising demand (Tilman *et al*., 2011). While several socio-economic changes (including reducing food waste, altering agricultural practices, and changing diets) will be required to attain global food security, a significant impetus remains with plant science research to improve germplasm, technologies, and our general understanding of these fascinating organisms that, directly or indirectly, supply the human population with most of its calories (Vitousek *et al.*, 1986, Haberl *et al.*, 2007).

Mitigating the yield losses imposed by plant pathogens will be integral to improving food supply in coming years. Broad estimates of total crop losses caused by plant pathogens are in the region of 30-40% (Flood, 2010), while global crop yield losses due to plant viruses have been estimated at 10-15% (Regenmortel & Mahy, 2009). Considered by a different metric, plant viruses have been estimated to cost the global economy more than $30 billion per year (Sastry & Zitter, 2014).

Potyviruses (Genus *Potyvirus*, Family *Potyviridae*) are one of the largest groupings of plant virus species (Gibbs & Ohshima, 2010), containing approximately 30% of all recorded species (Riechmann *et al.*, 1992). In addition to representing a large proportion of all plant viruses, potyviruses are an important focus for research because certain potyvirus species can cause devastating diseases in important crops. Interestingly phylogenetic analysis has indicated that there was a major radial speciation event for potyviruses approximately 6000 years ago, in the Holocene epoch, which is coincident with the emergence of agriculture (Gibbs *et al.*, 2008). While

speculative, these findings could suggest considerable anthropogenic influences on the evolution of potyviruses. Furthermore, the concurrent emergence of many extant potyvirus species with the development of agriculture substantiates their status as significant crop pathogens as they are likely to have co-adapted alongside the evolutionary domestication of wild plants into crops.

Some examples of agriculturally relevant potyviruses are presented in Table 1.1 and three particularly relevant species are discussed under the following sub-headings:

1.1.2 Potato Virus Y

Potato virus Y (PVY) is the type member (the species after which the family is named) of *Potyviridae* (Gibbs & Ohshima, 2010). As the name suggests, it was first characterised as the causal agent of severe mosaic diseases on potato plants (Smith, 1931). To this day, PVY remains one of the biggest pathological threats to potato production and is considered to be the most damaging virus of cultivated potatoes (Singh *et al.*, 2008). Virus diseases can be particularly difficult to manage for vegetatively propagated crops such as potatoes, because viruses can easily pass to subsequent generations. This is in contrast with 'true seed' crops (those which are sexually propagated by seeds produced from a floral meristem) as most viruses are excluded from floral meristems, or early embryos so are not transmitted to subsequent generations through seeds (Wang & Maule, 1994, Wang *et al.*, 1997, Schwach *et al.*, 2005). Potatoes have been ranked as the fourth most important staple crop, based on global yields (Scholthof *et al.*, 2011). This, along with the worldwide distribution of PVY (Cuevas *et al.*, 2012) testifies the importance of research into this virus. Furthermore, in recent decades a recombinant strain of PVY, called PVY^{NTN}, has placed further emphasis on PVY research (Beczner *et al.*, 1984). The PVY^{NTN} strain can overcome the genetic resistance which is effective against its parent strains PVY^{N} and PVY^{O} (van den Heuvel *et al.*, 1994). Moreover, infections with PVY^{NTN} are associated with severe disease symptoms in tubers, described as potato tuber necrotic ringspot disease (PTNRD) (Romancer *et al.*, 1994). The presence of PTNRD renders tubers un-marketable and hence the emerging prevalence of PVY^{NTN} can result in effective yield losses of up to 100% (Karasev & Gray, 2013). In addition to potato,

PVY is a threat to several other crops within the *Solanaceae* family, including: tomato (*Solanum lycopersicum*), pepper (*Capsicum annuum*), and tobacco (*Nicotiana tabacum*) (Hasiow-Jaroszewska *et al.*, 2015, Sharma *et al.*, 1989, Thomson & Wright, 1966). Taken together, the symptom severity, global distribution and agricultural relevance of its hosts make PVY an important focus for plant virus research.

1.1.3 Plum Pox Virus

Plum Pox Virus (PPV) was identified and characterised in Bulgaria shortly after the first description of PVY (Atanasoff, 1935). PPV primarily infects stone fruits within the *Prunus* family, including: plums (*P. domestica*), apricots (*P. armeniaca*), peaches (*P. persica*), almonds (*P. dulcis*), and cherries (*P. avium*) (Sochor *et al.*, 2012). Infection of these species with PPV causes a disease commonly known as Sharka, with diverse symptoms such as: leaf and petal mosaics, reduced fruit size, fruit deformation, altered fruit sugar/acid content, and premature fruit drop (Ilardi & Tavazza, 2015, Usenik *et al.*, 2015, Usenik & Marn, 2017). Arguably, the latter can cause the most devastating losses as fruit which falls from the tree before ripening has no economic value to either the fresh fruit or fruit-based-alcohol industries. In extreme cases PPV-induced fruit drop can occur at frequencies of up to 100%, resulting in total yield loss (Kamenova & Milusheva, 2005). Globally, PPV-incurred losses to commercial stone fruit was valued at approximately 10 billion euros over a 30-year period (1976-2006) (Cambra *et al.*, 2006). While early reports of Sharka disease were confined to eastern Europe, its incidence is now globally widespread (Rimbaud *et al.*, 2015). It is thought that commercial trading of infected *Prunus* plants is the main cause for the global spreading of the Sharka disease (Rimbaud *et al*., 2015). Interestingly, a few regions with substantial stone fruit industries (including Australia, New Zealand, California (USA) and South Africa) remain free from Sharka disease, and hence several governing bodies across the globe have classified PPV as a quarantine pathogen (Rimbaud *et al*., 2015). Phytosanitary measures such as this remain the main method for prevention and protection against Sharka disease, and there is an apparent paucity of resistance genes effective against PPV infection in commercial cultivars (Kegler *et al.*, 1998). Interestingly, transgenic RNAi technology (see RNA silencing below) has

been utilised to engineer a plum variety, 'HoneySweet', with stable resistance to PPV (Scorza *et al.*, 1994, Scorza *et al.*, 2001). After extensive field trials, 'HoneySweet' has recently been approved for commercial cultivation in the USA (Scorza *et al.*, 2012). However, the use of transgenes in agriculture remains controversial in most of the world so application of this technology is limited. The disease severity, global distribution, and limited knowledge of natural resistance mechanisms related to PPV infection means that this virus, in addition to PVY, supports the significance potyvirus research.

1.1.4 Turnip Mosaic Virus

The first recorded symptoms of Turnip Mosaic Virus (TuMV) infections date back to 1862, in relation to the 'flower-breaking' petal colour variegation seen on certain ornamental plants such as *Matthiola incana* (Tompkins, 1939). The viral cause of these symptoms, however, was unknown. For several decades, many viral diseases of Brassica crops were described resulting in diverse naming of the causal viruses. Subsequent discovery that these manifold diseases were all caused by the same potyvirus species resulted in their reclassification under the unifying name, TuMV (Shattuck, 2010). The historical complexity of the nomenclature for this virus is a reflection of its extensive host range. While *Brassicaceae* species are the hosts with greatest agricultural relevance, TuMV is known to infect at least 318 different plant species representing 156 genera and 46 families (Chung *et al.*, 2015). Due to this wide host range, along with the disease severity and associated economic losses in important crops (Spence *et al.*, 2007), TuMV was classified as the second most damaging viral pathogen to field-grown vegetables (Tomlinson, 1987). In addition to its agricultural relevance, TuMV is an important virus for potyvirus research in general. Early development of TuMV infectious clones (Sanchez *et al.*, 1998) represented a paradigm shift from classical virological methods by providing a tool for studying the molecular genetics of potyviruses. Furthermore, the ability of TuMV to infect the model plant *Arabidopsis thaliana* and the genetic synteny between this plant and agriculturally important Brassica hosts makes the TuMV/*Arabidopsis* patho-system a useful model for potyvirus research (Walsh & Jenner, 2002).

Potyvirus	Main Host Family	Important Crop Species	References
Potato Virus Y (**PVY**)	*Solanaceae*	Potato, Tomato, Pepper, Tobacco	(Karasev & Gray, 2013)
Plum Pox Virus (**PPV**)	*Prunus*	Plum, Apricot, Peach, Almonds, Cherry	(Sochor *et al*., 2012)
Turnip Mosaic Virus (**TuMV**)	*Brassicaceae*	Oilseed Rape, Turnip, Chinese Cabbage, Cabbage	(Walsh & Jenner, 2002)
Papaya Ringspot Virus (**PRSV**)	*Cucurbitaceae*	Papaya	(Gonsalves, 1998)
Maize Dwarf Mosaic Virus (**MDMV**)	*Poaceae*	Maize, Sorghum	(Louie & Knoke, 1975)
Bean Common Mosaic Virus (**BCMV**)	*Fabaceae*	Bean	(Sharma *et al.*, 2015)
Sweet Potato Feathery Mottle Virus (**SPFMV**)	*Convolvulaceae*	Sweet Potato	(Loebenstein, 2015)

Table 1.1: List of agriculturally relevant potyvirus species, and their respective hosts.

1.2 The potyvirus genome and molecular structure

Potyviruses exist as flexuous filamentous nanoparticles which are approximately 700-800nm in length and clearly visible under an electron microscope (EM) (Figure 1.1). These viral nanoparticles, or virions, are ribonucleoproteins composed of virus-encoded coat proteins (CP) encapsulating a single copy of the viral genome. The potyvirus genome is a single-stranded positive sense RNA molecule (+ssRNA) that is about 10kb in length (Allison *et al.*, 1985). The viral RNA is covalently linked at its 5' terminus to a virally encoded VPg protein (Murphy *et al.*, 1991) and contains a polyadenylated tail at its 3' terminus (Hari *et al.*, 1979, Hari, 1981). The +ssRNA genome is used directly as a template for replication, as well as translation. Both proceed by the coordinated action of viral proteins and host factors which are usurped by the virus. The genome is translated into a single polyprotein which is autocatalytically cleaved into ten multifunctional viral proteins (Allison *et al.*, 1985). An eleventh protein (PIPO – pretty interesting potyviral protein) is produced by a translational frameshift in the P3 coding region of the genome (Chung *et al.*, 2008). This frameshift protein was recently renamed P3N-PIPO after the discovery that the full protein sequence contains a portion of N-terminus of P3 (Vijayapalani *et al.*, 2012). The genomic position and biological functions of each of the potyviral proteins is summarised in Table 1.2 and Figure 1.2.

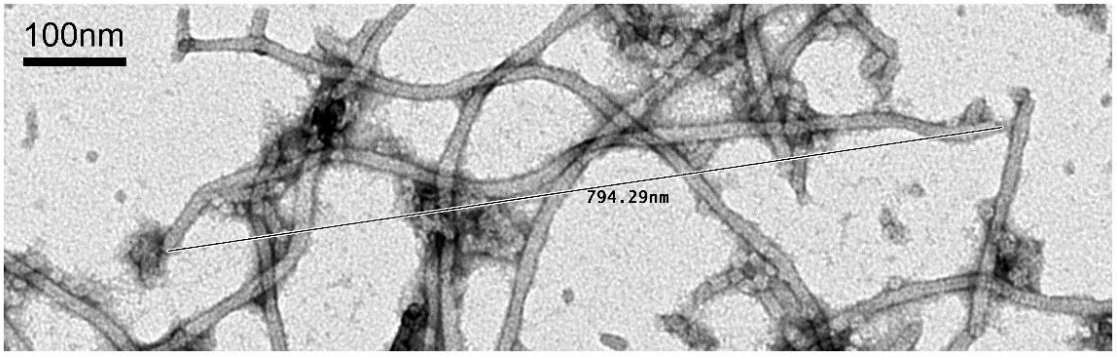

Figure 1.1 Electron micrograph of purified PVY virions. Scalebars indicate the approximate virion length. Purification and EM imaging by Douglas Pyott.

Protein	Genomic position	Function	References
P1	131…1216	Cleavage of the viral polyprotein ssRNA binding	(Verchot *et al.*, 1992) (Soumounou & Laliberte, 1994)
Hc-Pro	1217…2590	Cleavage of the viral polyprotein Suppression of RNA silencing Virion transmission Intercellular movement	(Carrington *et al.*, 1989) (Kasschau & Carrington, 1998) (Thornbury *et al.*, 1990) (Rojas *et al.*, 1997)
P3	2591…3655	Intercellular movement Symptom development	(Cui *et al.*, 2010) (Riechmann *et al.*, 1995)
P3N-PIPO	2591…3258	Intercellular movement	(Wei *et al.*, 2010b)
6K1	3656…3811	Viral replication	(Cui & Wang, 2016)
CI	3812…5743	Intercellular movement RNA helicase activity	(Carrington *et al.*, 1998) (Lain *et al.*, 1990)
6K2	5744…5902	Insertion into host membranes to form replication complexes	(Schaad *et al.*, 1997)
VPg	5903…6478	Viral genome stability and translation	(Murphy *et al.*, 1991) (Wittmann *et al.*, 1997)
NIa	6479…7207	Cleavage of the viral polyprotein	(Carrington *et al.*, 1993)
NIb	7208…8758	Genome replication	(Hong & Hunt, 1996)
CP	8759…9622	Encapsidation Intercellular movement Virion transmission Regulating viral replication Regulating viral translation	(Allison *et al.*, 1985) (Rojas *et al.*, 1997) (Baulcombe *et al.*, 1993) (Mahajan *et al.*, 1996) (Hofius *et al.*, 2007) (Hafren *et al.*, 2010)

Table 1.2 Genomic position and biological function(s) of the potyviral proteins. Genomic positions are given relative to the start of translation (+1) and are based on the annotated sequence of the TuMV UK1 strain (genbank ID AF169561). The position of P3N-PIPO is as defined by (Wei *et al.*, 2010b).

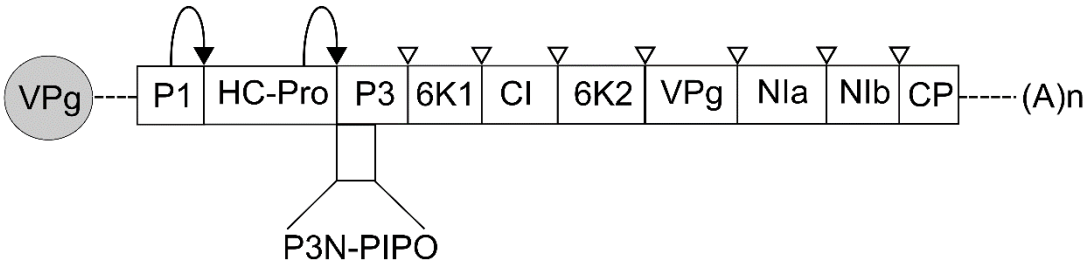

Figure 1.2 Schematic of the potyvirus polyprotein. Dashed lines represent the 5'/3' UTRs. The potyviral VPg (grey) is covalently attached to the 5'UTR, and the 3'UTR contains a polyA tail (A)n. The boxed areas indicate the coding region of the viral genome. Black arrows indicate self-processing protease activities of P1 and HC-Pro. White arrow-heads indicate trans-processing by the NIa protease. P3N-PIPO is produced by translational slippage in the P3 coding region.

1.3 Potyvirus transmission

In contrast to viruses infecting bacterial or animal cells, plant viruses have a strong reliance on vectors for their ultimate survival. The first reason for this is that, unlike bacterial or animal hosts, plants are sessile so viral movement between plants requires motile vectors. Secondly, plant virus vectors provide a means of entry into the host cell because plant viruses do not possess mechanisms to penetrate host cells, as bacterial and animal viruses do. While a diverse range of organisms can serve as vectors for plant viruses, including parasitic plants (Hosford, 1967) and fungi (Hollings & Stone, 1969), the most common mode of transmission is by insect vectors. Potyviruses are predominantly transmitted by aphids in a non-persistent manner, meaning that the virus only associates transiently with the mouthpart of the aphid but does not enter the aphid's digestive or circulatory systems. The interaction between potyvirus virions and aphid mouthparts (stylets) was first clearly shown by EM imaging (Taylor & Robertson, 1974). However, when purified potyvirus virions were fed to aphids they were found to be non-transmissible (Pirone & Megahed, 1966) leading to speculation that a helper component (HC) must be required for viral transmission. Evidence for a viral origin of this HC came from experiments showing

that purified potyvirus virions could become transmissible if aphids were previously fed on potyvirus-infected plants which had been UV treated to inactivate the virions in the sap (Kassanis & Govier, 1971). Similarly, fractions of infectious sap containing no potyvirus virions were shown to facilitate the aphid transmission of purified virions (Govier & Kassanis, 1974a, Govier & Kassanis, 1974b). By comparing sequences of different potyvirus strains with varying transmissibility, HC was mapped to what is now defined as the Helper Component-Protease (HC-Pro) coding region of the potyvirus genome (Thornbury *et al*., 1990). Further sequence comparisons and site directed mutagenesis revealed that a conserved PTK (proline, threonine, lysine) domain within a cysteine rich region of the N-terminal domain of HC-Pro is necessary for aphid transmission (Granier *et al.*, 1993, Peng *et al.*, 1998, Robaglia *et al.*, 1989). The favoured model by which HC-Pro enables virion transmission is known as the 'bridge hypothesis', in which HC-Pro acts as a molecular bridge to adhere virions to the aphid stylet. Evidence supporting the bridge hypothesis came from experiments where radioactively labelled potyvirus virions were fed to aphids in the presence or absence of HC-Pro. HC-Pro was found to not influence the uptake of virions by aphids but to increase their affinity to specific parts of the aphid stylet (Berger & Pirone, 1986). The specific molecular interactions underlying the bridge hypothesis are now relatively well understood. A conserved DAG (aspartate, alanine, glycine) motif in the viral CP is necessary for interacting with the PTK motif of HC-Pro, and hence is required for transmission (Baulcombe *et al*., 1993, Lee *et al.*, 1993). While the details of the HC-Pro/CP interaction have been well characterised, the molecular details for the adhesion of HC-Pro to the aphid stylet have been more elusive, particularly concerning the transmission determinants on the aphid stylet. A conserved KITC (lysine, isoleusine, threonine, cysteine) motif on HC-Pro has been identified as a necessary component for the interaction with aphid stylets (Atreya & Pirone, 1993, Blanc *et al.*, 1998). Several aphid proteins have been shown to interact with HC-Pro, including cuticular proteins (Dombrovsky *et al.*, 2007) and ribosomal protein homologues (Fernandez-Calvino *et al.*, 2010), but the minimal domains necessary and sufficient for transmission have yet to be resolved. As natural variations in the PTK, DAG, and KITC motifs have been shown to affect the transmissibility of virions (Antignus *et al.*, 1989, Granier *et al*., 1993) and determine which insect species can act as vectors (Stenger *et al.*, 2005, Dombrovsky *et al.*, 2005), transmission by vectors not only plays an important role in the viral lifecycle but can also provide selective pressures that drive potyvirus evolution.

1.4 The potyvirus translation-replication process

1.4.1 Co-ordination of viral processes

To successfully infect host cells, viral RNAs must orchestrate a number of processes which can be broadly categorised as translation, replication, encapsidation, and movement to un-infected cells (Mäkinen & Hafren, 2014). Each of these processes is ultimately controlled by viral proteins which commandeer a plethora of host factors to carry out the viral processes. Recent efforts to model virus-host protein interaction networks clearly illustrate the complexity and sophisticated organisation of this molecular parasitism (Elena & Rodrigo, 2012, Martinez *et al.*, 2016). As viral proteins are required for each stage in the infection cycle, logically the first stage involves translation of the viral genome. This occurs by usurpation of the host cell's translation apparatus, whereby interaction between core translational components and the VPg, which is bound to the 5' untranslated region (UTR) of the viral genome, results in translation of the potyviral polyprotein (reviewed in (Robaglia & Caranta, 2006)). A more detailed discussion of this process is provided in chapter 5.

The role of VPg in translation of the potyvirus genome is supported by several lines of evidence. Firstly, potyviral VPgs have been shown to interact with eIF4E and eIF(iso)4E proteins which play important roles in eukaryotic translation (Leonard *et al.*, 2000, Wittmann *et al.*, 1997). Furthermore, VPg was shown to stimulate translation of recombinant reporter genes containing potyvirus 5'UTRs, in a dose-dependent manner (Eskelin *et al.*, 2011). Likewise, dose-dependent VPg-stimulation of potyvirus gene expression was observed for non-replicating potyviruses, suggesting a direct role for VPg in potyvirus translation (Eskelin *et al.*, 2011). Nevertheless, the role of VPg in the potyvirus lifecycle is likely to not be limited to translation functions alone. For instance, uridylated VPg may serve as a primer for the potyviral replicase (NIb) for synthesis of the -sense viral RNA (Puustinen & Makinen, 2004). Additionally, VPg has been implicated in potyvirus cell-cell movement (possibly forming a movement complex with cytoskeletal elements and eIF4E) and systemic movement (Rajamaki & Valkonen, 1999, Rajamaki & Valkonen, 2002). The potyviral VPg also operates to ensure stability of the viral genome, by binding to the 5'UTR (Murphy *et al.*, 1991).

In some cases, the VPg protein may also protect the viral RNA by disrupting host RNA silencing (Rajamäki & Valkonen, 2009),which is an important antiviral defence response (see below).

The potyviral CP has emerged as an important regulator for the transition between different stages of the infection cycle. While the CP itself is not strictly required for replication, it was discovered that a translation into part of the CP coding region of the genome is required before viral replication can occur, due to a downstream cis-active replication element (Mahajan *et al*., 1996). The phenomenon is thought to represent a 'check-point' style of regulation that ensures aberrant, truncated viruses do not proceed to the replication stage. A control mechanism involving the CP also functions to temporally divide the processes of translation/replication and encapsidation which are antagonistic, as CP accumulation has been shown to limit viral translation in a dose-dependent manner (Besong-Ndika *et al.*, 2015). The current model for this temporal control involves a plant encoded Coat Protein Interacting Protein (CPIP) which, together with its co-chaperone Heat Shock Protein 70 (HSP70), target CP for degradation during the early phase of viral translation (Hofius *et al*., 2007, Hafren *et al*., 2010). Phosphorylation of CP by another host protein, Casein Kinase II (CK2) has been proposed as an additional mechanism by which premature virion assembly is prevented (Ivanov *et al.*, 2003). These mechanisms ensure that viral encapsidation only occurs at a late stage in the infection cycle, when the rate of CP production overcomes the sequestration mechanisms.

Host factors involved in translation of the viral genome may also be involved in the initiation of replication. For instance, the potyviral RNA dependent RNA polymerase, NIb, has been shown to form a complex with several host translation factors, by associating with the potyviral proteins NIa and VPg (Li *et al.*, 1997, Beauchemin *et al.*, 2007, Thivierge *et al.*, 2008). These protein-protein interactions along with a protein-RNA interaction involving a stem-loop structure at the 3'UTR of the potyvirus genome (Haldeman-Cahill *et al.*, 1998) deliver NIb to the 3' terminus of viral RNA. This results in the synthesis of a -sense RNA strand, from which many more +sense RNA copies can be synthesised (Martinez *et al.*, 2011). This replication process takes

place within virus-induced membrane structures known as Virus Replication Complexes (VRCs).

1.4.2 Virus Replication Complexes

The formation VRCs which hi-jack cellular membranes and cytoskeletal components is a conserved feature of +sense RNA viruses (Miller & Krijnse-Locker, 2008, den Boon & Ahlquist, 2010, Laliberte & Sanfacon, 2010, Verchot, 2011, Grangeon *et al.*, 2012b). VRCs are proposed to serve several functions including; concentrating replication factors around the viral RNA, providing a physical scaffold for the replication complex, and physically impeding host immunity effectors (Laliberte & Sanfacon, 2010).

Potyvirus VRCs are induced by a virally encoded integral membrane protein, 6K2 (Schaad *et al*., 1997), which inserts into the membrane of ER-Exit Sites (ERES) (Wei & Wang, 2008). Interestingly, ectopic expression of the 6K2 protein alone can induce the formation of VRC-like structures (Schaad *et al*., 1997). Potyvirus VRCs are highly dynamic and motile structures which utilise the host's cytoskeleton (Cotton *et al.*, 2009) and Golgi-mediated secretory pathways for both intra- and inter-cellular movement (Grangeon *et al.*, 2012a, Wei *et al.*, 2013). While the potyviral VRCs originate at the ER, they have been shown to accumulate in perinuclear bodies (Cotton *et al*., 2009) and chloroplast aggregates (Wei *et al.*, 2010a). Recently, an interaction between a chloroplast protein, PsbO1, and 6K2 has been proposed as a mechanism by which potyviral VRCs are anchored to chloroplast membranes (Geng *et al.*, 2017). Moreover, time-course imaging of fluorescently labelled 6K2 vesicles suggests a one-way flow of VRCs from the ER to chloroplasts, via the Golgi apparatus (Grangeon *et al*., 2012a). This, along with the finding that the majority of double stranded RNA (dsRNA) co-localised to chloroplast-bound VRCs, led to the conclusion that chloroplast membranes are the main site of potyvirus replication (Wei *et al*., 2010a). The implications of this close association between chloroplasts and potyvirus VRCs have yet to be fully investigated. Recently, it has been shown that disrupting photosynthesis (Manfre *et al.*, 2011) or chloroplast retrograde signalling (Abdelkefi *et al.*, 2017) can influence the outcome of potyvirus infections.

1.5 Movement of potyviruses through the plant

1.5.1 Cell-cell (local) movement

Plasmodesmata (PD) are pore-like conduits which enable cytoplasmic continuity of adjacent plant cells (Maule, 2008). They are dynamic structures which can adjust their size exclusion limit (Sambrook & Russell), to regulate the intercellular trafficking of macromolecules (Roberts & Oparka, 2003). Plant viruses must pass through PD to move between adjacent cells during local infection, and to enter the phloem and/or xylem for successful systemic infection. However, the native SEL of PD is too small to allow the passage of virions or viral RNA. To overcome this, plant viruses encode movement proteins (MPs) which facilitate viral movement by increasing the SEL of PD and by interacting with viral CPs and RNA (Schoelz *et al.*, 2011). Potyviruses are an unusual exception among plant viruses, with respect to MPs, because they do not encode proteins with dedicated movement functions. Instead, movement of potyviruses proceeds by the concerted action of several viral proteins which also possess non-movement related functions (Carrington *et al*., 1998). EM imaging revealed that the potyviral cylindrical inclusion (CI) protein forms a transient cone-like structure around PD, during the early stages of infection (Roberts *et al.*, 1998). Similar EM studies revealed the presence of viral RNA and CP oligomers at the centre of the cylindrical inclusions (CIs), which were interpreted to represent virions passing through a CI/PD movement complex (Rodriguez-Cerezo *et al.*, 1997). Interestingly, the role of CP in potyvirus movement is not limited to virion assembly, as site directed mutagenesis of the CP coding region revealed that certain mutations could affect cell-cell movement with no effect on virus encapsidation or replication (Dolja *et al.*, 1994). Furthermore, by microinjecting purified CP and HC-Pro into plant cells it was suggested that both of these proteins could independently increase the SEL of PD (Rojas *et al*., 1997). These results should be interpreted cautiously however, as artificial delivery of these over-expressed viral proteins may not accurately reflect their native functions in natural potyvirus infections. Site directed mutagenesis of the CI (Carrington *et al*., 1998), HC-Pro (Klein *et al.*, 1994, Cronin *et al.*, 1995, Kasschau *et al.*, 1997), and P3N-PIPO (Wen & Hajimorad, 2010) coding regions provided genetic evidence that these proteins are also necessary for potyvirus movement. Recently it has been shown that the P3N-PIPO protein interacts with both CI (Wei *et al*., 2010b)

and a PD localised membrane protein, PCaP1 (Vijayapalani *et al*., 2012), leading to the conclusion that P3N-PIPO may serve an anchoring function to allow CI movement structure to form around PD. Potyvirus movement was severely restricted in the *pcap1* mutant confirming its essential role for anchoring potyvirus movement complexes to PD (Vijayapalani *et al*., 2012). Disruption of the actin/myosin cytoskeletal network (Agbeci *et al.*, 2013) and vesicular secretory pathways (Wei *et al*., 2010b, Grangeon *et al*., 2012a) also prevented intercellular movement, highlighting the need for functional intracellular trafficking mechanisms to deliver movement complexes and VRCs to PD.

1.5.2 Long distance (systemic) movement

By comparison to cell-cell movement, long distance (systemic) movement of potyviruses has been less well characterised. This may partly be because cell-cell movement is a pre-requisite for systemic movement and hence it is hard to distinguish the two using genetic approaches. However, a mutation in the central region of HC-Pro was found to debilitate systemic movement with no apparent effects on replication or intercellular movement (Kasschau *et al*., 1997). Later, this mutation was shown to abrogate HC-Pro-mediated inhibition post-transcriptional gene silencing (Kasschau & Carrington, 1998) which forms an important antiviral defence response (see below). Therefore, the role of HC-Pro in long distance movement is likely to be an indirect effect of curtailing a systemic defence response, rather than promoting systemic movement *per se*. It has long been assumed that systemic movement of potyviruses involves virions moving exclusively through the phloem. However, recently it was shown that active potyvirus VRCs could be detected in both the phloem and xylem tissue, dramatically altering our understanding of the routes and mechanisms of systemic potyvirus infection (Wan *et al.*, 2015). Evidently, our understanding of how potyviruses move systemically is still vague.

1.6 Genetic resistance to potyvirus infection

1.6.1 General principles of genetic resistance to pathogens

When considering resistance to plant pathogens it is worth recalling the mantra that any given plant will be resistant to most pathogens and hence it is susceptibility, not resistance, which is exceptional (Kang *et al.*, 2005, Maule *et al.*, 2007). Non-host resistance is an extreme form of resistance whereby no strains of a pathogenic species are able to infect any genotypes of a given plant species (Fraser, 1990). While non-host resistance is clearly prevalent and important in natural contexts, the underlying mechanisms remain poorly understood. This is partly due to the fact that it is difficult to genetically map the determinants because, by definition, there is an absence of within-species genetic variation associated with the response. In contrast, host resistance has been better studied. Upon invasion by pathogens, plant cells can mount a general immune response known as PAMP(pathogen associated molecular pattern)-triggered immunity (PTI). Common PAMPs that trigger PTI include chitin and flagellin for fungal and bacterial pathogens, respectively, while viral dsRNA could be considered as a viral PAMP. To evade PTI, many pathogens produce effector molecules (or avirulence (avr) factors) that interfere with PAMP perception or downstream signalling. Perception of these effector molecules initiates an additional immune response known as effector triggered immunity (ETI), which relies on the products of plant resistance genes (R genes) to enable detection of pathogen effectors. Therefore, allelic variation at resistance gene (*R gene*) loci determines whether a plant will be resistant or susceptible to a pathogen. Similarly, avirulence (avr) factors are the pathogen-derived determinants of infection. Hence it is the specific combination of *R genes* and *avr* loci which determine the onset of for a given pathogen-host interaction. This has been described as the 'gene for gene' model of resistance (Flor, 1971). The co-evolution of host and pathogen results in a perpetual evolutionary arms race, whereby novel *R genes* and *avr* loci are selected to detect pathogens and evade detection, respectively (Dangl & Jones, 2001). A summary of some dominant *R genes* and their corresponding avr factors, which have been determined for potyvirus infections in various host species is presented in Table 1.2.

R gene	Gene product	Avr product	Virus	Host	References
BcTuR3	NB-LRR	unknown	TuMV	Chinese Cabbage	(Ma *et al.*, 2010)
TuN1	NB-LRR	P3	TuMV	*Arabidopsis*	(Kim *et al.*, 2008)
rsv1	NB-LRR	P3 HC-Pro	SMV	Soybean	(Hajimorad *et al.*, 2003)
Pvr1	NB-LRR	unknown	PRSV	Melon	(Anagnostou *et al.*, 2000)
Pvr4	NB-LRR	NIb	PVY	Pepper	(Kim *et al.*, 2015)
I	NB-LRR	unknown	BCMV	Bean	(Collmer *et al.*, 2000)
Y-1	NB-LRR	unknown	PVY	Potato	(Vidal *et al.*, 2002)
Ny	NB-LRR	HC-Pro	PVY	Potato	(Tian & Valkonen, 2013)
TuRBO1 /2/3/4/5	Quantitative trait loci	P3 CI	TuMV	Oilseed rape	(Jenner *et al.*, 2000, Jenner *et al.*, 2002)
RTM1 *RTM2* *RTM3*	Jacalin protein HSP homologue Meprin/MATH	CP	TEV PPV LMV	*Arabidopsis*	(Mahajan *et al.*, 1998, Whitham *et al.*, 2000, Cosson *et al.*, 2010)

Table 1.2: List of dominant potyvirus *R genes* and corresponding Avr products.

1.6.2 Dominant potyvirus resistance genes

Many of the genetically dominant potyvirus *R genes* encode nucleotide binding leucine rich repeat (NB-LRR) proteins. These well conserved proteins play an important role in a canonical disease resistance response. NB-LRR proteins can perceive avr products either directly or indirectly, and initiate a signalling cascade involving the production of reaction oxygen species (ROS) and the phytohormone

salicylic acid (SA). The production of ROS and SA results in programmed cell death (PCD) and *PATHOGENESIS RELATED* (*PR*) gene expression, respectively. The outcome of these events is defined as the hypersensitive response (HR). It has been proposed that PCD mainly operates to prevent viruses from spreading out from infected cells (Ross, 1961), while the SA pathway leads to systemic acquired resistance (SAR) (Conrath, 2006). As the name suggests, SAR involves mobile signals moving systemically to prime healthy tissue in advance of the spreading infection (Dean & Kuć, 1986). Over the years various chemical signals have been proposed to operate as a mobile signal for SAR and the topic is still a matter of scientific debate (reviewed in (Fu & Dong, 2013)).

Interestingly, a class of genetically dominant potyvirus *R genes* have been discovered which act independently to the canonical HR (Mahajan *et al*., 1998). Certain *Arabidopsis* ecotypes are fully susceptible to systemic infection by the potyvirus Tobacco Etch Virus (TEV), whereas other ecotypes are susceptible to the same TEV strain in local tissue only. Map-based cloning of these ecotypes resulted in the identification of *RESTRICTED TEV MOVEMENT* (*RTM*) loci (Mahajan *et al*., 1998, Whitham *et al.*, 1999). RTM1 belongs to the jacalin family of proteins (Chisholm *et al.*, 2000), while RTM2 is a transmembrane protein with homology to heat shock proteins (Whitham *et al*., 2000). RTM1 and RTM2 were both shown to be localised to the sieve elements of phloem tissue (Chisholm *et al.*, 2001), corroborating their role as suppressors of long distance TEV movement. However, their mode of action remained obscure for many years. More recently, RTM3 has been characterised as protein with meprin and TRAF homology (MATH) domains, and has been shown to form homodimers and also to heterodimerise with RTM1 (Cosson *et al*., 2010). This led to a hypothesis that RTM complexes may function in the phloem to sequester virions or factors required for systemic viral movement. While RTM-based resistance was originally described as a specific response to TEV (Mahajan *et al*., 1996), it was later shown to be effective against strains of two other potyviruses – PPV and Lettuce Mosaic Virus (LMV)(Decroocq *et al.*, 2006). Interestingly, mutations adjacent to the DAG motif in the CP coding region of PPV have been associated with a subversion of RTM-based resistance, giving further support that RTM complexes act by sequestering virus particles or viral movement complexes (Decroocq *et al.*, 2009). However, there

is no evidence to support a direct interaction between the potyvirus CP and any of the RTM proteins. Recently, a protein-protein interaction study identified several host proteins capable of interacting with both the LMV CP and at least one of the three RTM proteins, supporting the notion of a multimeric sequestration complex (Sofer *et al.*, 2017). A further two *RTM* genes (*RTM4* and *RTM5*) have been identified genetically but have not yet been characterised (Cosson *et al.*, 2012). Hence, significant research is still required to elucidate the mechanistic properties for RTM-based potyvirus resistance.

1.6.3 Recessive potyvirus resistance genes (susceptibility factors)

An interesting observation has been made that most potyvirus resistance genes are genetically recessive (Provvidenti & Hampton, 1992). In almost all cases, these recessive *R genes* are the result of loss-of-function mutations in genes which encode factors (known as susceptibility factors) that are required for part of the potyvirus lifecycle. The susceptibility factors that have most frequently been associated with recessive potyvirus resistance are members of the eukaryotic translation initiation factors (eIFs) family of proteins. Specifically, eIF4E and its paralogue, eIF(iso)4E have been identified as susceptibility factors for many potyvirus-host interactions. An interaction between the potyvirus VPg and the eIF4E/eIF(iso)4E protein(s) is essential for the potyvirus to usurp the host cell's translational apparatus to translate its genome (Wittmann *et al*., 1997, Schaad *et al.*, 2000, Leonard *et al.*, 2000). Hence, mutations at the *eIF4E/eIF(iso)4E* loci which prevent VPg binding restrict the translation/replication of the viral genome, resulting in resistance (Leonard *et al*., 2000, Duprat *et al.*, 2002, Lellis *et al.*, 2002, Sato *et al.*, 2005). In addition to eIF4E/eIF(iso)4E, other components of the eukaryotic translation complex, and related signalling pathways have been identified as susceptibility factors for certain potyviruses (Nicaise *et al.*, 2007, Thivierge *et al*., 2008, Gao *et al.*, 2015, Ouibrahim *et al.*, 2015, Shopan *et al.*, 2017). A summary of susceptibility factors underlying recessive resistance to various potyviruses is given in Table 1.3.

R gene	Gene product	Virus	Host	References
lsp1	eIF(iso)4E	TuMV	*Arabidopsis*	(Lellis *et al*., 2002)
pvr2	eIF4E	PVY	Pepper	(Ruffel *et al.*, 2002)
pot1	eIF4E	PVY, TEV	Tomato	(Ruffel *et al.*, 2005)
mo1	eIF4E	LMV	Lettuce	(Nicaise *et al.*, 2003)
bc3	eIF4E	BCMV	Bean	(Naderpour *et al.*, 2010)
sbm1	eIF4E	Pea seed-born mosaic virus	Pea	(Gao *et al.*, 2004)
rym4	eIF4E	Barley Yellow Mosaic Virus	Barley	(Stein *et al.*, 2005)
nsv	eIF4E	Melon Necrotic Spot Virus	Melon	(Nieto *et al.*, 2006)
retrO3	eIF2Bβ	TuMV	Mustard	(Shopan *et al*., 2017)
Rpv1	Chloroplast phospho-glycerate kinase	PPV	*Arabidopsis*	(Poque *et al.*, 2015)

Table 1.3 List of recessive potyvirus R genes

1.7 Immune responses to potyvirus infection

The induction of HR (mentioned above) is an example of an immune response to virus infection. Immune responses can be defined as cellular processes, triggered by a foreign substance, which act to remove or contain the invading entity. In addition to HR, several immune responses occur in response to virus infections. For example, a lyposomal-based degradation process called autophagy has been described as an immune response to PVY infection in *Nicotiana benthamiana* (Choi *et al.*, 2016). Nonsense mediated decay (NMD), a regulatory response which degrades aberrant mRNAs, was recently shown to operate as an immune response to a +ssRNA plant virus (Garcia *et al.*, 2014). Another form of RNA-based viral immunity, which is much

more widely conserved and better characterised, is the process referred to as RNA silencing. The process of RNA silencing and its role in virus immunity is discussed in the subsequent section.

1.8 RNA silencing

1.8.1 General features of RNA silencing

RNA silencing is an evolutionarily conserved mechanism of genetic regulation in eukaryotes, which involves sequence-specific silencing of gene expression by short (21-25 nucleotide (nt)) RNA fragments (Baulcombe, 2004). The presence dsRNA within a cell initiates a common pathway whereby RNAse III-like ribonulceases of the Dicer or Dicer-like (DCL) families cleave the dsRNA precursor into small RNA (sRNA) fragments (Bernstein *et al.*, 2001). These sRNA molecules can form an RNA-induced silencing complex (RISC) by association with ARGONAUTE (AGO) proteins, along with other factors, to silence the expression of homologous sequences guided by Watson-Crick base paring. A subset of RNA silencing pathways additionally utilise RNA-dependent RNA polymerases (RDRs) to amplify the sRNA silencing signals (Voinnet, 2008). RNA silencing operates at several levels including transcriptional gene silencing (TGS) and post-transcriptional gene silencing (PTGS). The former involves methylation of DNA cytosine bases by RNA directed DNA methylation (RdDM) (Jones *et al.*, 1998, Mette *et al.*, 2000, Wassenegger *et al.*, 1994), while the latter occurs either by cleavage (Baumberger & Baulcombe, 2005) or translational inhibition (Brodersen *et al.*, 2008) of homologous RNA sequences. Broadly, sRNAs can be divided into microRNAs (miRNAs) and short-interfering RNAs (siRNAs), of which there are several types (Table 1.4). The term 'sRNA' will be used herein to refer to all short RNA species (miRNAs and siRNAs), while the term 'siRNA' refers specifically to this class of sRNA to the exclusion of miRNAs.

In plants, the genes encoding for the core components of RNA silencing have multiplied and diversified (Baulcombe, 2004). They play distinct roles in diverse RNA silencing pathways to control development (Tang *et al.*, 2003), abiotic stress signalling (Borsani *et al.*, 2005), and defence against invading nucleic acids such as viruses or

transposons (Aravin *et al.*, 2007, Ratcliff *et al.*, 1999, Waterhouse *et al.*, 2001). *Arabidopsis* possesses four *DCL*, ten *AGO*, and six *RDR* homologues. Functional divergence has resulted in different combinations of these RNA silencing components mediating the various RNA silencing pathways, although it is now emerging that there is a high degree of redundancy and promiscuity between the different pathways. At the broadest level, RNA silencing can be classified based on whether the dsRNA initiator derives from an endogenous (within the host genome) or exogenous (outside the host genome) source (Figure 1.3).

Small RNA (sRNA)	**Role**	**Origin**	**Main mode of action**
miRNA	Post-transcriptional gene regulation	Endogenous	PTGS
tasiRNA	Post-transcriptional gene regulation	Endogenous	PTGS
natsiRNA	Post-transcriptional gene regulation	Endogenous	PTGS
hcsiRNA	Epigenetic regulation	Endogenous	TGS
vsiRNA	Viral defence	Exogenous	PTGS/TGS
Transgene-induced siRNA	Defence against 'foreign' DNA	Exogenous	PTGS/TGS

Table 1.4 Types of sRNAs, their origin, and main mode of action.

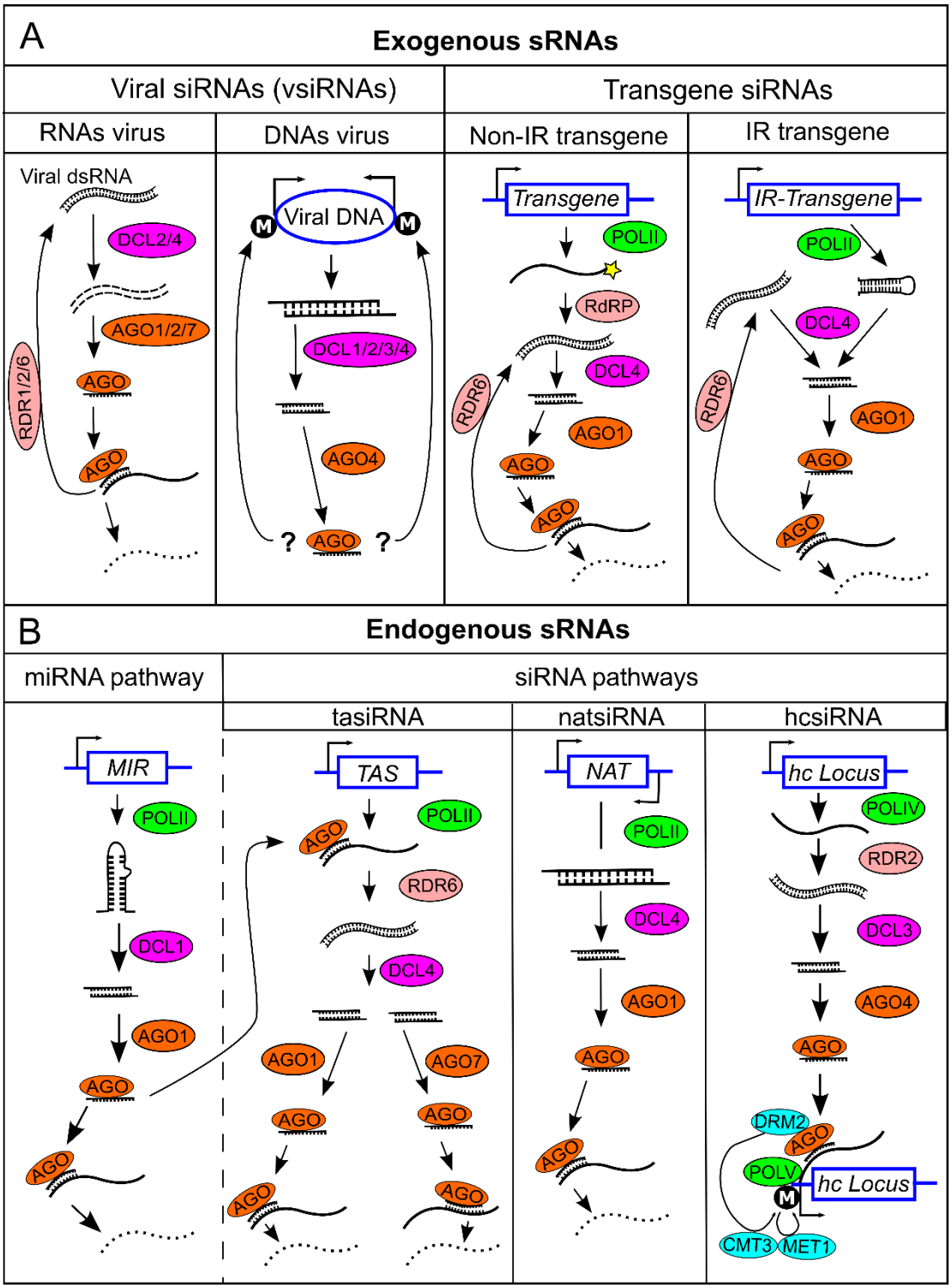

Figure 1.3 Summary of RNA silencing pathways in plants. DNA/RNA are represented by blue/black lines respectively, yellow star represents structural aberrancy, white M indicates cytosine methylation. Dashed lines indicate sRNA generation by dsRNA cleavage, dotted lines represent RNA degradation. **A:** RNA silencing pathways initiated by exogenous DNA or RNA. (IR = inverted repeat) **B:** RNA silencing pathways initiated from endogenous loci.

1.8.2 Exogenous RNA silencing

Exogenous RNA silencing can be initiated by dsRNA derived from viruses or by expression of transgenes within the host (Figure 1.3A). For RNA viruses, the dsRNA trigger can result from dsRNA replication intermediates and, most likely, also from secondary structures in the viral genome (Molnar *et al.*, 2005). Secondary structures in the transcripts of DNA viruses may also direct the processing of viral short interfering RNAs (vsiRNAs), though the best understood mechanism of vsiRNA generation from DNA viruses involves bi-directional transcription of the viral genome giving rise to overlapping fragments forming dsRNAs (Pooggin, 2013). All four DCLs have been associated with viral RNA silencing, directly or indirectly, though for several viruses DCL4 and DCL2 are most important (Blevins *et al.*, 2006, Deleris *et al.*, 2006, Garcia-Ruiz *et al.*, 2010, Qu *et al.*, 2008). The vsiRNAs produced can be 21nt, 22nt, or 24nt in length and associate predominantly with AGO1, AGO2 and AGO7 (Carbonell *et al.*, 2012, Qu *et al*., 2008, Wang *et al.*, 2012) which silence the viral genome either by endonucleolytic cleavage or, for DNA viruses, by DNA methylation (Wassenegger *et al*., 1994). RDR1, RDR2 and RDR6 have been most frequently associated with the amplification of vsiRNAs from RNA viruses (Wang *et al*., 2012). None of the *RDR* homologues, however, have been directly implicated in silencing pathways for DNA viruses. Furthermore, mutation of key components in the RdDM pathway (such as RDR2, POLIV and POLV – see below) was found to have no effect on the infection of DNA viruses (Aregger *et al.*, 2012, Blevins *et al*., 2006). This suggests that the methylation of viral DNA does not involve the exact same pathway as endogenous RdDM. The extent to which vsiRNA-directed methylation of viral DNA plays a direct role in viral defence is still an open question (Pooggin, 2013).

An important feature of vsiRNAs is that they are able to move between cells via PD (Dunoyer *et al.*, 2005) and over long distances through the phloem (Yoo *et al.*, 2004). This allows vsiRNA to reach systemic tissue with, and even before, the invading virus which serves a defence priming function analogous to SAR. This phenomenon was first described based on observations that transgenic plants carrying viral sequences showed enhanced systemic recovery when challenged by cognate viruses (Lindbo *et al.*, 1993). The role of non-coding RNA in this recovery response was deduced from

observations that viral recovery was enhanced in transgenic plants carrying non-coding viral sequences (Dougherty *et al.*, 1994), or plants which had been previously infected with related viruses (Ratcliff *et al*., 1999). More recently, vsiRNA pathways have been directly implicated in systemic viral recovery as RNA silencing mutants (at *DCL2*, *DCL4*, *RDR1*, *RDR6*, *AGO1*, *AGO2* and *AGO7* loci) involved in vsiRNA processing and activity are impaired in systemic recovery to virus infections (Carbonell *et al*., 2012, Garcia-Ruiz *et al*., 2010).

In transgene-induced RNA silencing the dsRNA trigger is thought to arise when 'aberrant' RNA transcribed by the transgene is recognised by a host RNA-dependent RNA polymerase (RdRP) and is subsequently copied into dsRNA (Baulcombe, 2004)(Figure 1.3A). In addition to the spontaneous production dsRNA from transgenes by aberrant RNA formation, transgenes can be engineered to trigger silencing with a high frequency by introducing inverted repeats (IR) into the transgenic sequence. Transcription of IR sequences leads to the formation of perfect fold-back dsRNA structures, which can lead to TGS and PTGS by the same mechanisms described above. IR containing transgenes can be tailored to silence potentially any sequence, and have frequently been used for research purposes to 'knock-down' the expression of genes in a process termed RNA interference (RNAi). This technology has also been used to engineer virus resistance in crops (as in the 'HoneySweet' example above) by inserting non-coding viral sequences in transgenes to prime the plant with vsiRNAs.

1.8.3 Endogenous RNA silencing

The first example of endogenous RNA silencing to be discovered in plants was the microRNA (miRNA) pathway (Reinhart *et al.*, 2002)(Figure 1.3B). Transcription of *MIR* genes by RNA polymerase II (PolII) produces long non-coding transcripts, which serve as miRNA precursors. Internal sequence complementarity within these precursor transcripts leads to the formation of imperfect fold-back dsRNA, which are recognised and cleaved by DCL1 to produce a 20-22nt dsRNA heteroduplex. One of the heteroduplex strands (the passenger strand) is degraded upon association with AGO1

to produce a mature miRNA, which can then silence homologous mRNA sequences by cleavage or translational inhibition (Brodersen *et al*., 2008).

In general, only one mature miRNA is produced from a partially double-stranded region of a fold-back single-stranded transcript of *MIR* genes, which directs the cleavage and subsequently, the degradation of mRNAs with complementary target sequences. However, some miRNAs can trigger the production of secondary sRNAs, referred to as trans-acting siRNAs (tasiRNAs) (Figures 1.3B). The biogenesis of tasiRNAs involves PolII transcription of long non-coding tasiRNA precursors from *TAS* loci. miRNA directed cleavage of the *TAS* transcript by AGO1 allows RDR6 to produce a dsRNA, primed by the bound miRNA. Phased cleavage of the dsRNA produces a population of 21nt tasiRNAs, which silence diverse endogenous loci in *trans* bound to AGO1 or AGO7. As the secondary (or transitive) tasiRNAs produced can have different sequences to the generative miRNA, tasiRNA pathways can operate as important signalling hubs for genetic regulation (MacLean *et al.*, 2010). The RDR6-dependent amplification of tasiRNAs critically distinguishes them from miRNAs, which lack a similar amplification process.

Heterochromatic siRNAs (hcsiRNAs) are a third class of endogenous sRNAs (Figure 1.3B). A large proportion of hcsiRNAs are generated from loci containing active (or potentially active) transposable elements (TEs). Hence an important function of hcsiRNAs is to silence TE expression to minimise their damaging effects on the genome. In addition, hcsiRNAs can be generated from domesticated TE sequences where they serve as regulators of endogenous gene expression. A notable example of this involves the imprinting of maternal or paternal genes by hcsiRNAs (Calarco *et al.*, 2012). The hcsiRNA/RdDM pathway is initiated by DCL3-dependent production of 24nt siRNAs, which can operate either in *cis* or in *trans* (targeting the same/different loci respectively from the generating locus) in association with AGO4. The RdDM pathway additionally involves a maintenance phase whereby PolIV transcription at the target locus produces a long non-coding RNA. AGO4 directed cleavage of this transcript primes the RDR2-mediated synthesis of a dsRNA, which is then processed by DCL3 to generate more hcsiRNAs (Figure 1.3B). A second non-coding transcript

is produced by PolV which serves as a scaffold for hcsiRNA/AGO4 binding whereupon DNA methylation factors such as Domains Rearranged DNA methyltransferase 2 (DRM2) and Chomomethylase 3 (CMT3) along with other RdDM components are recruited to methylate cytosine bases in all sequence contexts. This complex process has been recently reviewed (Matzke & Mosher, 2014).

A final class of endogenous sRNA to mention are the natural antisense siRNAs (natsiRNAs). Convergent, overlapping PolII transcription at natsiRNA generating loci produces transcripts with regions of perfect complementarity, resulting in a dsRNA, which is cleaved by DCL2 to produce 21/22nt natsiRNAs that can mediate PTGS of homologous sequences (Figure 1.3B). Though less characterised than the other sRNA silencing pathways, natsiRNAs have often been associated with responses biotic and abiotic stress (Borsani *et al*., 2005, Katiyar-Agarwal *et al.*, 2006, Moldovan *et al.*, 2010).

1.8.4 Viral suppressors of RNA silencing

The potency and ubiquity of vsiRNAs as antiviral defence molecules is reflected by the fact that plant viruses have evolved diverse strategies to overcome this central form of defence. Most of the characterised viral suppressors of RNA silencing (VSR) are proteins which reduce vsiRNA efficacy by disrupting various stages of the viral RNA silencing pathway. In addition to the well characterised protein VSRs, some viruses encode decoy RNAs that sequester vsiRNAs, thereby protecting the replicating viral genome (Blevins *et al.*, 2011).

The main potyviral VSR is the multi-functional HC-Pro, which binds 21nt, 22nt, and 24nt to sequester them away from antiviral RISCs (Lakatos *et al.*, 2006). A conserved FRNK (phenylalanine, arginine, asparagine, lysine) motif in the central region of HC-Pro is required for efficient siRNA binding (Sahana *et al.*, 2014, Shiboleth *et al.*, 2007). Mutations around this FRNK motif result in severely debilitated viruses which are unable to spread throughout the plant due to an active RNA silencing response (Gal-On, 2000). Interestingly, the infectivity of such mutant viruses is restored in *dcl2*, *dcl4 rdr1*, *rdr2*, *rdr6*, *ago1*, *ago2*, and *ago10* mutant plants demonstrating the

coevolutionary defence/counter-defence tug-of-war between vsiRNAs and VSRs (Garcia-Ruiz *et al*., 2010, Garcia-Ruiz *et al.*, 2015).

1.8.5 Cross-talk in viral RNA silencing pathways

Several examples of signalling cross-talk exist for viral RNA silencing pathways. The potyviral VSR, HC-Pro, does not discriminate between vsiRNAs and endogenous sRNAs such as miRNAs and tasiRNAs (Lakatos *et al*., 2006). Hence potyvirus infections can severely disrupt plant developmental pathways due to HC-Pro sequestering developmentally important miRNAs and tasiRNAs (Chapman *et al.*, 2004, Endres *et al.*, 2010, Jay *et al.*, 2011, Kasschau *et al.*, 2003). This cross-talk between HC-Pro and endogenous sRNAs has been demonstrated to be the cause of at least some potyvirus symptoms (such as stunting, leaf curling and abnormal flower development) as transgenic expression of Hc-Pro in *Arabidopsis* was found to re-capitulate these viral symptoms and sequester miRNAs involved in developmental regulation (Chapman *et al*., 2004, Jay *et al*., 2011, Kasschau *et al*., 2003).

Several studies have demonstrated that vsiRNAs can exhibit trans-acting effects by targeting host mRNAs, hence disrupting gene expression in the host plant. Using bioinformatic approaches combined with small RNA deep sequencing, it was shown that there are abundant Tobacco Mosaic Virus (TMV) vsiRNAs with the potential to disrupt gene expression in *Arabidopsis*, by vsiRNA-mRNA cross-talk. Evidence that these vsiRNA can directly repress host targets was provided for two of the candidate genes, encoding a polyadenylation specificity factor (At1g30460) and a translocon-associated protein α (At2g16595) (Qi *et al.*, 2009). Similarly, vsiRNAs produced by Cucumber Mosaic Virus (CMV) satellite RNAs have been shown to degrade transcripts of the *N. benthamiana CHL1* gene, which encodes an enzyme involved in chlorophyll synthesis (Shimura *et al.*, 2011, Smith *et al.*, 2011). While these examples illustrate that viral and host genomes can interact at multiple levels, the evolutionary pressures driving the vsiRNA/host gene cross-talk remain speculative.

Many genes which are central to RNA silencing pathways are themselves regulated by the action of miRNAs. For instance, miR162 and miR168 have been shown to direct

the cleavage of *DCL1* and *AGO1* transcripts, respectively (Xie *et al.*, 2003, Vazquez *et al.*, 2004). As DCL1 and AGO1 are required for miRNA processing and activity, their miRNA-mediated regulation provides a homeostatic feedback loop which maintains the production and activity of plant miRNAs. The importance of this feedback mechanism was demonstrated by creating transgenic plants expressing a miR168-resistant *AGO1*, hence abolishing the homeostatic control of *AGO1*. These plants showed severe developmental defects which were linked to aberrant activity of diverse miRNAs (Vaucheret *et al.*, 2004). Some VSRs can disrupt these homeostatic responses to indirectly supress host immunity. For example, the p38 VSR of Turnip Crinkle Virus (TCV) binds to AGO1 to suppress its activity. The reduced AGO1 activity results in an overexpression of *DCL1*, due to a de-repression of the miR162-mediated homeostatic control. *DCL1* overexpression subsequently reduces expression of *DCL2* and *DCL4* (through an unknown mechanism) thereby hindering the production of vsiRNAs, which requires DCL2 and DCL4 (Azevedo *et al.*, 2010). Similarly, Cymbidium ringspot virus (CymRSV) infections were shown to result in a reduction of AGO1 protein due to enhanced production of an endogenous miRNA (miR168) which is involved in a negative feedback mechanism regulating *AGO1* expression (Várallyay *et al.*, 2010, Vaucheret *et al*., 2004). Another example of viral de-repression of a miRNA-regulated defence response involves a three-way cross-talk between miRNAs, *R genes* and virus infections. In tomato, a miRNA (miR482) was discovered which represses several NB-LRR type *R genes*. Infection by several viruses led to an increase in the expression of these *R genes* due to de-repression of miR482 and its associated tasiRNAs (Shivaprasad *et al.*, 2012). This provides an interesting example of a virus-inducible *R gene* defence response.

The overlapping nature of RNA silencing pathways provides great scope for signalling cross-talk. While there are several examples of viruses subverting or exploiting these cross-talk mechanisms, bioinformatic analysis suggests that there are many more to be discovered. Furthermore, the functional relevance of these intriguing plant-pathogen interactions is yet to be properly understood.

1.9 Infection of *Arabidopsis thaliana* with Turnip Mosaic Virus as a model system for potyvirus infection.

The *Arabidopsis*/TuMV patho-system has emerged as a prevalent model for investigating the molecular aspects of potyvirus infections. Using *Arabidopsis* as a potyvirus host offers several advantages of working with a well-established model organism. These include extensive genetic resources, a small and fully sequenced genome, a wealth of microarray data, and a relatively fast growth rate and life cycle. While there are several potyvirus species which can infect *Arabidopsis,* TuMV is a logical choice as it naturally infects a wide range of plants within the *Brassicaceae* family, to which *Arabidopsis* also belongs. Hence, good genetic resources have been developed for TuMV. An infectious clone of TuMV has been developed in which the coding sequence for GREEN FLUORESCENT PROTEIN (GFP) is inserted in between the P1 and HC-Pro coding regions of the TuMV genome (hereafter TuMV-GFP). In this clone, an artificial NIa protease recognition sequence was inserted at the 3'terminus of the GFP sequence (at the GFP/HC-Pro junction), and the native P1 protease recognition sequence was maintained at the P1/GFP junction (Lellis *et al*., 2002). Hence, the GFP polypeptide is excised from the translated TuMV-GFP polyprotein by the action of the P1 and NIa proteases at its 5' and 3' termini, respectively (Figure 1.4). This releases a soluble GFP protein which can be easily detected using an ultraviolet (UV)-emitting lamp (Figure 1.5). One of the advantageous features of TuMV-GFP is that it allows the synthesis and spread of the virus to be indirectly monitored by non-destructive imaging under UV illumination. Moreover, recovery from virus infections can be inferred by the loss of GFP signal in systemic tissue. For these reasons, TuMV-GFP and *Arabidopsis* were chosen as model organisms in this thesis, to study natural and engineered defence mechanisms to potyvirus infection.

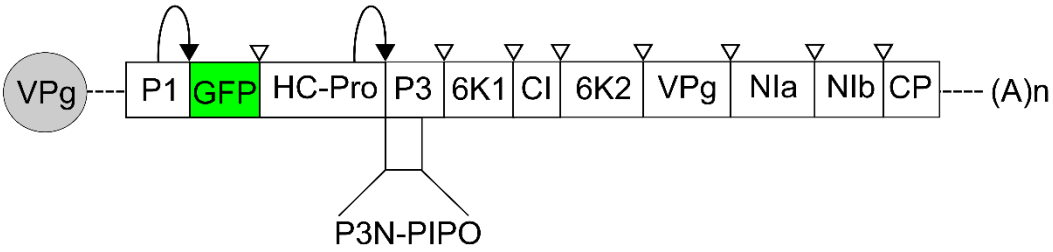

Figure 1.4 Genome structure of TuMV-GFP. The coding sequence for GFP (green box) is inserted between the P1 and HC-Pro regions of TuMV. Addition of an artificial NIa protease site immediately downstream of the GFP sequence allows the GFP polypeptide to be released from the viral polyprotein. Black arrows indicate self-processing protease activities of P1 and HC-Pro. White arrow-heads indicate trans-processing of the NIa protease.

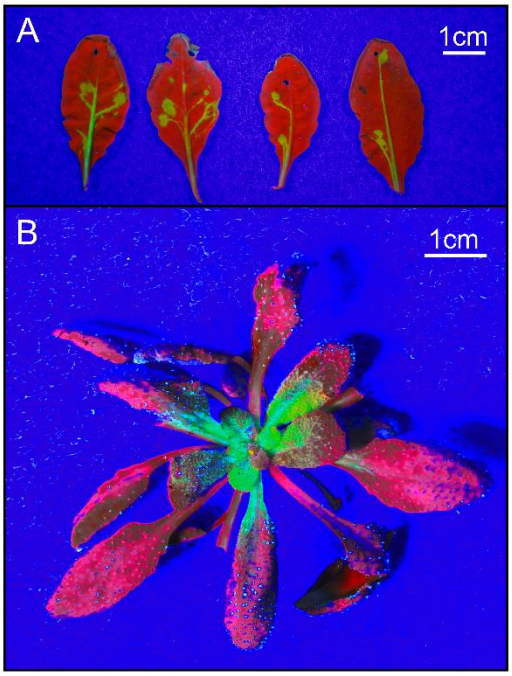

Figure 1.5 Detection of GFP (green) produced by infection with TuMV-GFP in *Arabidopsis* **A** inoculated tissue and **B** systemic tissue. Image was taken under ultraviolet (UV) light. Un-infected plant tissue is red due to chlorophyll auto-fluorescence.

Chapter 2: Materials and methods

2.1 Genetic material

2.1.1 Details and sources of plasmid vectors

Vector name	Vector type	Source	Reference
pCB-TuMV-GFP	Binary vector	James Carrington	(Garcia-Ruiz *et al*., 2010)
pDE-Cas9	Binary vector	Holger Puchta	(Fauser *et al.*, 2014)
pEN-Chimera	Gateway entry vector	Holger Puchta	(Fauser *et al*., 2014)
pY010	Mammalian expression vector	Feng Zhang	(Zetsche *et al.*, 2015)
pY016	Mammalian expression vector	Feng Zhang	(Zetsche *et al*., 2015)
pK7FGW2	Binary vector	Ann Depicker	(Karimi *et al.*, 2002)
pICSL70008	Binary vector	Nicola Patron	(Engler *et al.*, 2014)

Table 2.1 Plasmid vector names, types and source.

2.1.2 Details and sources of transgenic/mutant plant material

Mutant allele	Gene locus	Mutation type	Source	T-DNA ID/ Reference
dcl2-1	AT3G03300	null	David Baulcombe	SALK_064627C
dcl3-1	AT3G43920	null	David Baulcombe	SALK_005512
dcl4-2	AT5G20320	null	David Baulcombe	GABI_160G05
rdr1-1	AT1G14790	null	David Baulcombe	SAIL_672_F11
rdr6-15	AT3G49500	null	David Baulcombe	SAIL_617_H07
CCA1-OX	AT2G46830	Over-expresser	Andrew Millar	(Wang & Tobin, 1998)
eif(iso)4E	AT5G35620	null	Christophe Robaglia	(Duprat *et al*., 2002)
arf19-1	AT1G19220	null	Yunde Zhao	SALK_009879
16c	-	Integrated GFP transgene	David Baulcombe	(Ruiz *et al.*, 1998)

Table 2.2 Mutated and genetically transformed plant material. All table entries apply to *Arabidopsis* except for 16c which is a transgenic *N. benthamiana* stably expressing GFP.

2.2 Plant growth conditions

Plants were grown in 6cmx6cmx8cm (length x width x depth) pots on Levington F2+S professional growth compost. Seeds were imbibed and stratified in 1.5ml Eppendorf tubes in darkness at 4^{o}C and then directly planted onto compost. For the majority of experiments plants were grown in SANYO/Panasonic growth chambers at 21^{o}C with 16:8 hour light:dark cycles with side illumination from cool, white fluorescent bulbs at approximately 200µMolesm^{-2}s^{-1} light intensity. Deviations from these standard conditions with respect to light intensity,

temperature and photoperiod are clearly stated in the relevant figure legends. Light intensity was varied using neutral density light filters (Blacklights Ltd., Neutral density filters #210 and #299).

2.3 Bacterial plasmid purification

Bacterial plasmids were purified from *Escherichia coli* (DH5α strain) cells using a Qiagen miniprep kit (Qiagen Catalogue #27106), according to the manufacturer's protocol.

2.4 Vector cloning

2.4.1 pDe-Ubi:SpCas9-U6:sgRNA_AteIF(iso)4E

An *eIF(iso)4E*-targeting sgRNA was designed using online software: CRISPR Design (http://crispr.mit.edu/), and DNA2.0 CRISPR gRNA design tool (https://www.dna20.com/eCommerce/cas9/input). The corresponding oligonucleotides Iso_Fw and Iso_Rv (Table 2.3) were annealed and cloned into BbsI digested entry vector pEn-Chimera (Fauser *et al*., 2014). After sequence verification, the sgRNA expression cassette was recombined into the destination vector pDe-Cas9 (Fauser *et al*., 2014) by Gateway cloning (LR reaction) according to the manufacturer's instructions (Life Technologies) resulting in pDe-Cas9-sgRNA_AteIF(iso)4E.

2.4.2 pDe-35S:SpCas9-U6:sgRNA_GFP

A *GFP*-targeting sgRNA was designed manually by inspecting the transgenic GFP sequence present in 16c plants. The corresponding oligonucleotides Cas9-sgRNA_GFP_Fw and Cas9-sgRNA_GFP_Rv (Table 2.3) were annealed and cloned into BbsI digested entry vector pEn-Chimera. After sequence verification, the sgRNA expression cassette was recombined into the destination vector pDe-Cas9 by Gateway cloning (LR reaction) according to the manufacturer's instructions (Life Technologies) resulting in pDe-Cas9-sgRNA_GFP.

2.4.3 pK7FWG2-35S:AsCpf1 and pK7FWG2-35S:LbCpf1

Plasmid vectors containing human-codon optimised Cpf1 from *Acidamonococcus* sp. BV3L6 (AsCpf1) and *Lachnospiraceae* bacterium ND2006 (LbCpf1) were a gift from Feng Zhang

(Addgene plasmids #69982 and #69988, respectively). The Cpf1 coding sequence and C-terminal nuclear localization signal were PCR amplified to include 5' SpeI and 3' BsrGI sites and a C-terminal HIS tag, using AsCpf1_Fw/LbCpf1_Fw and Cpf1-6XHIS_Rv (Table 2.3). PCR fragments were cloned into plant transformation binary vector pK7FWG2 as SpeI-BsrGI fragments.

2.4.4 pK7FWG2-U6:AscrRNA_GFP and pK7FWG2-U6:LbcrRNA_GFP

A *GFP*-targeting crRNA was designed manually by inspecting the transgenic GFP sequence present in 16c plants. The crRNA sequences were added downstream of the U6 promoter by PCR amplification with U6_Fw and As_crRNA_GFP_Rv or Lb_crRNA_GFP_Rv (Table 2.3), using pEN-Chimera as template DNA. Resultant PCR products were cloned as SmaI-EcoRV fragments into separate pK7FWG2 vectors pK7FWG2 to create pK7FWG2-U6:AscrRNA_GFP and pK7FWG2-U6:LbcrRNA_GFP

2.4.5 pK7FWG2-Ubi:LbCpf1-Oleosin:RFP

The 35S terminator sequence from pK7FWG2-35S:LbCpf1 was excised by BsrGI/ApaI restriction digestion. The ubiquitin terminator sequence was PCR amplified from pDE-Cas9, adding 5' BsrGI and 3' ApaI sites, using pea3A_Fw and pea3A_Rv (Table 2.3). This amplicon was ligated into the BsrGI/ApaI opened pK7FWG2-35S:LbCpf1. The 35S promoter sequence was then excised from this intermediate vector by SacI/EcoRI restriction digestion. The ubiquitin promoter sequence was PCR amplified from vector pDE-Cas9, adding 5' SacI and 3' EcoRI sites, using Ubi_Fw and Ubi_Rv (Table 2.3). This insert was ligated into the SacI/EcoRI opened intermediate vector to create pK7FWG2-Ubi:LbCpf1. An oleosin:RFP cassette was PCR amplified from vector pICSL70008 (Engler *et al*., 2014) using Oleosin:RFP_Fw and Oleosin:RFP_Rv (Table 2.3). The PCR fragment was ligated into the SmaI site of pK7FWG2-Ubi:LbCpf1 to create pK7FWG2-Ubi:LbCpf1-Oleosin:RFP. The SmaI site was restored, for future cloning, by addition of a SmaI sequence to Oleosin:RFP_Rv.

2.4.6 pK7FWG2-Ubi:SpCas9-Oleosin:RFP

The Cpf1 coding sequence was excised from pK7FWG2-Ubi:LbCpf1by EcoRI/PacI restriction digestion. The Cas9 coding sequence was PCR amplified from pDE-Cas9, adding

5’ EcoRI and 3’ PacI sites, using Cas9_Fw and Cas9_Rv (Table 2.3). The PCR fragment was ligated into the EcoRI/PacI opened pK7FWG2-Ubi:LbCpf1 vector to create pK7FWG2-Ubi:SpCas9. The oleosin:RFP was inserted into the SmaI site of pK7FWG2-Ubi:SpCas9, as described above, to create pK7FWG2-Ubi:SpCas9-Oleosin:RFP.

2.4.7 pK7FWG2-Ubi:LbCpf1-U6:crRNA_arf19_1 and pK7FWG2-Ubi:LbCpf1-U6:crRNA_arf19_2

ARF19-targeting crRNAs were designed manually by inspecting the full genomic sequence of the *ARF19* locus (AT1G19220). The crRNA sequences were added downstream of the U6 promoter by PCR amplification with U6_Fw and crRNA_arf19_1_Rv or crRNA_arf19_2_Rv (Table 2.3), using pEN-Chimera as template DNA. The PCR amplicons were cloned into the SmaI site of pK7FWG2-Ubi:LbCpf1-Oleosin:RFP to create pK7FWG2-Ubi:LbCpf1-Oleosin:RFP-U6:crRNA_*arf19*_1 and pK7FWG2-Ubi:LbCpf1-Oleosin:RFP-U6:crRNA_*arf19*_2.

2.4.8 pK7FWG2-Ubi:SpCas9-Oleosin:RFP-U6:sgRNA_arf19_1 and pK7FWG2-Ubi:SpCas9-Oleosin:RFP-U6:sgRNA_arf19_2

ARF19-targeting sgRNAs were designed manually by inspecting the full genomic sequence of the *ARF19* locus (AT1G19220). The corresponding oligonucleotides, sgRNA_arf19_1_Fw/ sgRNA_arf19_1_Rv and sgRNA_arf19_2_Fw/ sgRNA_arf19_2_Rv were annealed and cloned into BbsI digested entry vector pEn-Chimera to create vectors pEn-Chimera:U6_sgRNA_arf19_1 and pEn-Chimera:U6_sgRNA_arf19_2 . The U6:sgRNA_arf19 cassettes were then PCR amplified from the pEn-Chimera:U6_sgRNA_arf19 vectors as templates, using U6_Fw and sgRNA_Rv. The resultant PCR amplicons were cloned into the SmaI site of pK7FWG2-Ubi:SpCas9-Oleosin:RFP to create pK7FWG2-Ubi:SpCas9-Oleosin:RFP-U6:sgRNA_*arf19*_1 and pK7FWG2-Ubi:SpCas9-Oleosin:RFP-U6:sgRNA_*arf19*_2

2.4.9 pK7FWG2-Ubi:LbCpf1-U6:crRNA_GFP_1-U6:crRNA_GFP_2-U6:crRNA_Donor_1-U6:crRNA_Donor_2-Donor_DNA.

GFP-targeting crRNAs were designed manually to direct CRISPR/Cpf1 cleavage of two sites, approximately 250bp apart, at the transgenic *GFP* locus. crRNAs were also manually designed to excise an artificial Donor DNA sequence. The crRNA sequences were added downstream of the U6 promoter by PCR amplification with HindIII-U6_Fw/NcoI-U6_Fw/EcoRI-U6_Fw/SacII-U6_Fw and crRNA_GFP_1_Rv/ crRNA_GFP_2_Rv / crRNA_Donor_1_Rv/ crRNA_Donor_2_Rv (Table 2.3), using pEN-Chimera as template DNA. Restriction sites were added to the amplification primers resulting in the following amplicons: HindIII-U6:crRNA_GFP_1-NcoI ; NcoI-U6:crRNA_GFP_2-EcoRI ; EcoRI-U6:crRNA_Donor_1-SacII ; SacII-U6:crRNA_Donor_2-SpeI. These fragments were cloned into the HindIII/SpeI sites of pK7FWG2-Ubi:LbCpf1 by 5-way ligation, to create pK7FWG2-Ubi:LbCpf1-U6:crRNA_GFP_1-U6:crRNA_GFP_2-U6:crRNA_Donor_1-U6:crRNA_Donor_2. An artificial Donor DNA sequence was PCR amplified from pCB-TuMV-GFP using Donor_DNA_Fw and Donor_DNA_Rv. Sequences complementary to the crRNA_Donor_1/2 guides, along with PAM sequences, were added to Donor_DNA_Fw and Donor_DNA_Rv oligonucleotide, to flank the Donor DNA with crRNA_Donor_1/2 recognition sites. The resulting PCR amplicon was cloned into the SmaI site of pK7FWG2-Ubi:LbCpf1-U6:crRNA_GFP_1-U6:crRNA_GFP_2-U6:crRNA_Donor_1-U6:crRNA_Donor_2 to create pK7FWG2-Ubi:LbCpf1-U6:crRNA_GFP_1-U6:crRNA_GFP_2-U6:crRNA_Donor_1-U6:crRNA_Donor_2-Donor_DNA.

Oligonucleotide name	Sequence (5'-3')	T_A (°C)	Extension time (s)
Iso_Fw	ATTGTGTGAACGAGCCTCTCCCGG	-	-
Iso_Rv	AAACCCGGGAGAGGCTCGTTCACA	-	-
Cas9-sgRNA_GFP_Fw	ATTGTAAGGGAATCGATTTCAAGG	-	-
Cas9-sgRNA_GFP_Rv	AAACCCTTGAAATCGATTCCCTTA	-	-
AsCpf1_Fw	TAACTAGTTGAATTCATGACACAGTTCGAGGGCTTT	64	120
LbCpf1_Fw	TAACTAGTTGAATTCATGAGCAAGCTGGAGAAGTTTACAA	64	120
Cpf1-6XHIS_Rv	AATGTACATTAGTGATGGTGATGGTGATGGGATCCCTTTTTCTTTTTTGCCTGG	64	120
U6_Fw	TTTTTTCTTCTTCTTCGTTCATACAG	58	15
As_crRNA_GFP_Rv	AAAAAACCGAGGATGTTTCCGTCCTCCTTATCTACAAGAGTAGAAATTACAATCACTACTTCGACTCTAGC	58	15
Lb_crRNA_GFP_Rv	AAAAAACCGAGGATGTTTCCGTCCTCCTTATCTACACTTAGTAGAAATTACAATCACTACTTCGACTCTAGC	58	15
pea3A_Fw	ATTGTACATTAATTAACAGGCCTCCCAGCTTTCG	63	20
pea3A_Rv	ATGGGCCCAAGCCTATACTGTACTTAACTTGATTGCATAATTAC	63	20
Ubi_Fw	GCTTGAGCTCACTAGTAAAAATTACGGATATGAATATAGGCATA	60	30
Ubi_Rv	AATTGAATTCGCTGCACATACATAACATATCAAGATC	60	30
Oleosin:RFP_Fw	ATGTCGCGGAACAAATTTTAAAAC	63	120
Oleosin:RFP_Fw	GGGACTAAATGGAGCAACCTACTGTTTTTG	63	120
Cas9_Fw	AATTGAATTCATGGATAAGAAGTACTCTATCGGACTC	60	180
Cas9_Rv	TTAATTAATCAAACCTTCCTCTTCTTCTTAGG	60	180
crRNA_arf19_1_Rv	AAAAAAACAAGACTTCCCACAGGAGGTAATATCTACACTTAGTAGAAATTACAATCACTACTTCGACTCTAGC	62	10
crRNA_arf19_2_Rv	AAAAAATTACAAACAAAACTGAATCACCCGATCTACACTTAGTAGAAATTACAATCACTACTTCGACTCTAGC	62	10

(Continued overleaf)

sgRNA_arf19_1_Fw	ATTGACCTCCTGTGGGAAGTCTTG	-	-
sgRNA_arf19_1_Rv	AAACCAAGACTTCCCACAGGAGGT	-	-
sgRNA_arf19_2_Fw	ATTGGTGATTCAGTTTTGTTTGTA	-	-
sgRNA_arf19_2_Rv	AAACTACAAACAAAACTGAATCAC	-	-
sgRNA_Rv	GAAAAAAAGCACCGACTCG	58	15
HindIII-U6_Fw	AATAAGCTTTTTTTTCTTCTTCTTCGTTCATACAG	58	15
NcoI-U6_Fw	AATCCATGGTTTTTTCTTCTTCTTCGTTCATACAG	58	15
EcoRI-U6_Fw	AATGAATTCTTTTTTCTTCTTCTTCGTTCATACAG	58	15
SacII-U6_Fw	AATCCGCGGTTTTTTCTTCTTCTTCGTTCATACAG	58	15
crRNA_GFP_1_Rv	TAACCATGGAAAAAACCTTAAATTTATTTGCACTACTGATCTACACTTAGTAGAAATTACAATCACTACTTCGACTCTAGC	58	15
crRNA_GFP_2_Rv	TAAGAATTCAAAAAACCGAGGATGTTTCCGTCCTCCTTATCTACACTTAGTAGAAATTACAATCACTACTTCGACTCTAGC	58	15
crRNA_Donor_1_Rv	TAACCGCGGAAAAAATAAGGAGTATCGTGCACTAGCTTATCTACACTTAGTAGAAATTACAATCACTACTTCGACTCTAGC	58	15
crRNA_Donor_2_Rv	TAAACTAGTAAAAAACTCGGATCCATTATGCCGTACGGATCTACACTTAGTAGAAATTACAATCACTACTTCGACTCTAGC	58	15
Donor_DNA_Fw	AATTTCAGCTAGTGCACGATACTCCTTAACCCGGGGCAAACTAATAGTCCCAACGAACTTCGAAG	68	15
Donor_DNA_Rv	AATTTCCGTACGGCATAATGGATCCGAGGCCCGGGAACTTGCTCACACACTCGACTGTTAC	68	15
EcoRI-LbCpf1_Fw	TTAAGAATTCATGAGCAAGCTGGAGAAGTTTACAAAC	67	120
6XHIS-SacI_Rv	ATTGAGCTCTTAGTGATGGTGATGGTGATGGG	67	120
6XHIS-DR-2xBbsI-DR-SacI_Rv	ATTCCATGGATCTACACTTAGTAGAAATTATTTAATCTTTGAAACGTCTTCTCGAAGACATCTACACTTAGTAGAAATTATTTAATCTTTGAAACTTAGTGATGGTGATGGTGATGGG	67	120
SacI-DR_Fw	ATTGAGCTCCCATCACCATCACTAAGTTTCAAAG	63	15
DR-MluI_Rv	TTAACGCGTCTCGAGGGATCCCCATGGATCTACACTTAG	63	15

(Continued overleaf)

Cpf1_guide_GFP_Fw	GTAGATAAGGAGGACGGAAACATCCTCGGGT	-	-
Cpf1_guide_GFP_Rv	TGAAACCCGAGGATGTTTCCGTCCTCCTTAT	-	-

Table 2.3 List of oligonucleotides used for vector cloning. PCR amplifications were performed using Q5 Hi-fidelity DNA polymerase (New England Biolabs, Catalogue# M0491). PCR conditions are outlined in *section 2.11.1*. The specific annealing temperature (T_A) and extension time for each PCR reaction is listed in the table.

2.5 Bacterial transformations

2.5.1 Transformation of Escherichia coli cells with plasmid vectors

Chemically competent *E.coli* cells were prepared using the Inoue method (Sambrook & Russell, 2006). Approximately 1µg of plasmid DNA was added to a 50µl aliquot of competent cells in a 1.5ml Eppendorf tube. The transformation mixture was thoroughly mixed by pipetting and incubated on ice for 30 minutes and then transferred to a water bath set to 42^{o}C for 45s, and returned to ice for 20 minutes. 500µl of LB medium (*section 2.21.1*) was added to the mixture and the tube was incubated for 50 minutes in a shaking incubator at 37^{o}C/250rpm. Transformed cells were selected on LB-agar plates containing the relevant selection antibiotics.

2.5.2 Transformation of Agrobacterium tumefaciens cells with plasmid vectors

Electrically competent *Agrobacterium tumefaciens* cells (AGL1 strain) were prepared by growing a cell culture in YEP medium (*section 2.21.2*) overnight at 28^{o}C with 250rpm shaking. Once the cell culture reached an OD_{600} of 1.0, cells were pelleted by centrifugation at 4000rpm and resuspended in and equal volume of YEP to wash the cells. Cells were re-pelleted and then resuspended in an equal volume of 10% glycerol. Approximately 1µg of plasmid DNA was added to a 50µl aliquot of competent cells in a 1.5ml Eppendorf tube. The transformation mixture was thoroughly mixed by pipetting and immediately transferred to an electroporation cuvette (Scientific Laboratory Supplies Ltd, Catalogue# FBR-101). Cells were electroporated at 1800V in a cell electroporator (Eppendorf, Catalogue# 2510). Cells were recovered from the cuvette by flushing them with 500µl of YEP medium, and transferred to a 1.5ml Eppendorf

tube. Cells were left to recover for 50 minutes in a shaking incubator at 28ºC/250rpm. Transformed cells were selected on YEP-agar plates containing the relevant antibiotics.

2.6 Plant transformation

2.6.1 Floral dipping

Agrobacterium tumefaciens (AGL1 strain) cells were transformed with the relevant plasmid by electroporation and transformed cultures were grown overnight in LB liquid medium at 28ºC with 250rpm shaking to a final OD_{600} of 1.0. Cells were pelleted by centrifugation at 4000rpm for 20 minutes. Pelleted cells were washed then resuspended in liquid LB and the OD_{600} was adjusted to 0.8. Silwet was added to the culture at a final concentration of 0.1% (v/v) and Acetosyringone was added to a final concentration of 150µM. 6 week old flowering *Arabidopsis* (Col-0 accession) plants were dipped in the *Agrobacterium* suspension for approximately 10s. This process was repeated 7 days later.

2.6.2 Transient infiltration

Agrobacterium cells were grown as described above and resuspended in infiltration medium (section *2.21.3*). The cells were incubated at room temperature without shaking for 1 hour before syringe infiltration into the abaxial surface of 3 fully expanded leaves of 4 week old *N. benthamiana* plants.

2.7 Selection of T_1 seed

2.7.1 Screening T_1 seedlings by BASTA selection

T_1 seeds were collected and scattered evenly on compost filled 70cmx30cmx6cm (lengh x width x depth) trays. 7 days after germination, seedlings were sprayed with a 120mg/l solution of BASTA by applying a fine mist of the herbicide across the entire area of the tray, using a hand-held spray bottle. The BASTA spray treatment was repeated at 14 and 21 days after germination. Plants were grown in a temperature controlled glasshouse at 21ºC with 16:8 light:dark cycles and approximately 300µMolesm^{-2}s^{-1} light intensity provided by halogen light bulbs and daylight. BASTA resistant plants were transplanted into 6cmx6cmx8cm (length x width x depth) pots to produce T_2 seed.

2.7.2 Screening T_1 seeds by FAST-red fluorescence sorting

T_1 seeds were spread evenly on a sheet of paper in a dark room. The seeds were illuminated with a green LED torch (Joyland, Green Light 530nm) and viewed through a red filter (iTekLife global, Transparent colour correction light gel filter red) to visualise the red seed fluorescence conferred by the oleosin:RFP transgene. Red fluorescing seeds were sorted onto a separate sheet of paper using a wetted wooden toothpick and then transferred to a 1.5ml Eppendorf tube.

2.8 TuMV-GFP rub-inoculations

Infectious TuMV-GFP sap was prepared by syringe infiltrating 4 week old *N.benthamiana* plants with *Agrobacterium tumefaciens* cells (as outlined in *section 2.6.2* above) containing the pCB-TuMV-GFP plasmid (Garcia-Ruiz *et al*., 2010). 14 days after infiltration, systemic leaves showing viral symptoms were harvested and homogenised with a sterile mortar and pestle. The homogenate was diluted 1:5 (w/v) in 1mM sodium phosphate buffer (pH 7.0) (section *2.21.4*) and frozen in 0.5ml aliquots at -80ºC. *Arabidopsis/N. benthamiana* leaves were rub-inoculated on the adaxial surface with 10µl of the 1:5 diluted viral sap aliquots, using aluminium oxide powder as an abrasive. The third, fourth, fifth, and sixth oldest rosette leaves of 4 week old *Arabidopsis* plants were used for rub-inoculations. For *N. benthamiana* rub-inoculations, the second, third and fourth oldest leaves were used.

2.9 UV imaging of TuMV-GFP

Viral expression of GFP was monitored using hand-held UV lamps (UVP, B-100AP Lamp 100W 365nm) and imaged using a Canon powershot G16 digital camera with automatic shooting mode. JPEG images were imported to 'Digital Photo Professional' (version 4.0) Canon software (http://cpn.canon-europe.com) for colour adjustment. The hue, saturation, and monochrome settings were adjusted manually using the sliders in the 'tool palette' function. Settings were adjusted to optimise the red/green/blue contrast for a representative image and the settings were saved as a script file to batch apply to all images.

2.10 Plant DNA/RNA purification

Total nucleic acids (TNA) were purified from plant tissue by phenol/chloroform extraction using a method adapted from (White & Kaper, 1989). Plant tissue was collected in 2ml Eppendorf tubes containing 2 metal ball-bearings (Spheric Trafalgar Ltd. 1/8") and frozen in liquid nitrogen. Plant tissue was ground in pre-chilled blocks in a Qaigen TissueLyzer (Qaigen, Catalogue# 85300). 650µl of tris-saturated phenol (pH8.0) (Severn Biotech Ltd. Catalogue# 40-1000-20) and 650µl of TNA extraction buffer (*section 2.21.7*) was added to the powdered plant tissue. The samples were thoroughly mixed by vortexing and then centrifuged as 13,000rpm for 5 minutes in an AccuSpin Micro 17R centrifuge (Fisher Scientific, Catalogue# 75002462). The supernatant was transferred to a 1.5ml Eppendorf tube containing 650µl of 25:24:1 phenol:chloroform:isoamyl-alcohol (Sigma, Catalogue# 1002480564). Samples were vortexed and centrifuged as before. The supernatant was transferred to a fresh 1.5ml Eppendorf tube containing 25:24:1 phenol:chloroform:isoamyl-alcohol, and mixed/centrifuged as before. The supernatant was then transferred to a fresh 1.5ml Eppendorf tube containing 24:1 chloroform:isoamyl-alcohol and mixed/centrifuged as before. The supernatant was then transferred to a siliconised 1.5ml Eppendorf tube (alpha laboratories Catalogue# LW2410) containing 550µl isopropanol and 25µl 4M sodium acetate. Samples were mixed by inverting the tubes and then incubated at -20ºC for 30 minutes. TNA was pelleted by centrifugation at 13,000rpm for 30 minutes at 4ºC. The pellet was washed with 500µl 80% ethanol. The ethanol was poured off and aspirated with a pipette tip and tubes were left open on ice for 30 minutes to allow residual ethanol to evaporate. The dried TNA pellets were resuspended in 50µl of autoclaved, twice distilled water. The concentration of the TNA was determined using a spectrophotometer (Nanodrop, ND-1000). The RNA quality was checked by mixing 2µl of sample with 2µl of gel loading dye (*section 2.21.8*), heating the samples at 65ºC for 5 minutes, and separating the TNA on a 2% agarose gel. The presence of sharp rRNA bands confirmed the integrity of the RNA.

2.11 PCR and gel imaging

2.11.1 PCR cycle conditions

PCR amplifications were performed using Q5 Hi-Fidelity DNA polymerase (New England Biolabs, Catalogue# M0491S). PCR reagents were added according to the manufacturer's guidelines. PCR cycles were as follows: 98ºC 120s > (98oC 10s > 'T_A' 20s > 72oC 't's)x35 > 72oC 120s. The annealing temperature ('T_A') and extension time ('t') are specific to each PCR

reaction and can be found in the along with the oligonucleotide information in Tables 2.3/2.4/2.5.

2.11.2 Gel imaging

PCR products were separated on a 1% agarose gel (unless otherwise stated) containing SYBR safe gel stain (Invitrogen, Catalogue# S33102) at the recommended dilution. Gels were run in 1X TBE buffer (*section 2.21.5*) and imaged in a digital gel imager (UVP, Biodoc-It^{TM}).

Oligonucleotide name	**Sequence (5'-3')**	**T_A (ºC)**	**Extension time (s)**	**PCR product (bp)**
Cas9_Fw	ATGGATAAGAAGTACTCTATCGGAC	55	30	664
Cas9_Rv	GCCTTCTTGACTTAGAGAGCCTAG			
EF1a_Fw	CTTCTTGAGGCTCTTGACCAG	55	30	418
EF1a _Rv	TGAGAGGTGTGGCAATCGAG			
GFP_Fw	CACTGGAGTTGTCCCAATTC	60	30	710
GFP_Rv	ATATGAATTCTTAAAGCTCATCATGTTTGTATAGTTC			
Donor-DNA_Fw	CTGTTGTTTTTACTTTAATGGGTTG	60	20	486
Donor-DNA_Rv	AAACGTAAAACGGCTTGTCC			

Table 2.4 List of oligonucleotides used for PCR diagnostic assays PCR amplifications were performed using Q5 Hi-fidelity DNA polymerase (New England Biolabs, Catalogue# M0491). PCR conditions are outlined in *section 2.11.1*, above. The specific annealing temperature (T_A) and extension time for each PCR reaction is listed in the table.

2.12 qRT-PCR

2.12.1 DNase treatment

DNA was removed from purified TNA samples by treating 3µg of TNA with TurboDNAse (Life Technologies, Catalogue# AM1907) according to the manufacturer's instructions. The RNA was precipitated (to wash out degraded nucleotides) by mixing the sample with 150µl isopropanol and 2µl 4M sodium acetate. RNA was pelleted, washed and resuspended as

outlined in section **2.10**. The concentration of the purified RNA was determined using a spectophotometer (Nanodrop, ND-1000).

2.12.2 cDNA synthesis

cDNA was synthesised from purified RNA using Superscript II reverse transcriptase, according to the manufacturer's instructions. 1µg of RNA was used for each reaction and random hexameric primers (Sigma-Aldrich) were used for template priming. cDNA was diluted 1:5 with twice distilled water for subsequent applications.

2.12.3 SYBR qRT-PCR reaction

Quantitative PCR of the cDNA was performed using SYBR green master mix (Roche, Catalogue# 04887352001) and gene-specific oligonucleotides. 10µl reactions were used, containing 5µl SYBR green master mix, 1µl of 10µM forward primer, 1µl of 10µM reverse primer, and 3µl of 1:5 diluted cDNA. 3 technical replicate reactions were performed for each biological replicate. Reactions were performed in 384-well PCR plates (Axygen, Catalogue# PCR-384-LC480-W) and a LightCycler480 thermocycler (Roche, Catalogue# 05015243001). The following cycling conditions were used for all reactions: 95ºC 5 min > (95ºC 10s > 58ºc 10s > 72ºc 10s) x 45.

2.12.4 Primer design and validation

Primers for quantitative real-time cDNA amplification were designed using PerlPrimer software (version 1.1.21) (perlprimer.sourceforge.net/) with the following parameters: Tm=60-61ºC, difference=1ºC; primer length=20-24nt; amplicon size=80-100nt; %GC/overlap/GC-clamp=off; span intron/exon boundary=on. A cDNA standard curve was used to test the amplification efficiency for each primer pair. For this, non-diluted cDNA was used for the highest concentration and 5 10-fold serial dilutions were made. Amplification efficiency was determined using the LightCycler480 software (https://molecular.roche.com/systems/lightcycler-480-system/) and only primer pairs with an efficiency of 2 (±0.01) were used for experiments. Oligonucleotides used for priming quantitative real-time cDNA PCR reactions are listed in Table 2.5.

2.12.5 SYBR qRT-PCR analysis

Analysis of the quantitative PCR reaction was performed using LightCycler480 software. Technical replicates were checked manually by visualising the amplification curves, to confirm accurate pipetting. Technical replicates were then grouped to use an average Cp value (an interpolated cycle number corresponding to the point of inflection of the exponential amplification curve) for the 3 technical replicates for subsequent analysis. Relative expression of target genes was determined using the 'advanced relative quantification' function, which applies a ΔΔCp calculation (Schmittgen & Livak, 2008). Briefly, this calculation involves a ΔCp value which is the difference in Cp values for the target gene and an internal control (or 'housekeeping') gene, for the same cDNA sample. ΔΔCp is defined as the difference in ΔCp between a positive control sample and each test sample. The relative pre-amplification concentration (c_0) of each sample can then be calculated by using the ΔΔCp as the negative exponent of base 2 ($c_0 = 2^{-\Delta\Delta Cp}$). For each experiment, an extra sample was collected to serve as a positive control to use in the ΔΔCp calculation. The housekeeping gene used in each experiment is indicated in the figures and figure legends.

Oligonucleotide name	Sequence (5'-3')	Target Locus	Gene ID
TuMV_Fw	GTGGCTCTAAACCTCGATCAT	TuMV Coat Protein	-
TuMV_Rv	AACCATGTGTCAAACTGCTTTC		
EF1a_Fw	CACCACTGGAGGTTTTGAGG	*EF1a*	AT5G60390
EF1a_Rv	TGGAGTATTTGGGGGTGGT		
DCL2_Fw	TGAAGCCAAACTCTCCAAGAAA	*DCL2*	AT3G03300
DCL2_Rv	TGTGTTCCCAATGTCATCTTCAA		
DCL4_Fw	CCATCCATCTTACAACAGGCA	*DCL4*	AT5G20320
DCL4_Rv	CATCAAGTAGTCCAGAACAGCA		
RDR1_Fw	ATCTATCTCCACGCTCTTCTACT	*RDR1*	AT1G14790
RDR1_Rv	TCAAACTTCTTATCGCCTATGACA		
RDR6_Fw	TGAAGTAGGAATTGTTTGGCGT	*RDR6*	AT3G49500
RDR6_Rv	GGATATTTGAAGTGTCTGCGATTT		
AGO1_Fw	CTCTCAATGTTGCCTGACAAAG	*AGO1*	AT1G48410
AGO1_Rv	CTGTGTTTCTTCCTCCAACCT		
AGO2_Fw	AGTTTAGGAGGCGTGATGTG	*AGO2*	AT1G31280
AGO2_Rv	CAGCCCTACAATGGTGAGTT		
AGO7_Fw	ATCGTAATAGCTTCTCTGGTGTTG	*AGO7*	AT1G69440
AGO7_Rv	ACTCAAGCCTATCTCCCTGAA		
TOC1_Fw	ATCTTCGCAGAATCCCTGTGATA	*TOC1*	AT5G61380
TOC1_Rv	GCACCTAGCTTCAAGCACTTTACA		
ACT7_Fw	TGAACAATCGATGGACCTGA	*ACT7*	AT5G09810
ACT7_Rv	CAGTGTCTGGATCGGAGGAT		
eif(iso)4E_Fw	ATTTGTGGTGTGGTTGCTAGT	*eIF(iso)4E*	AT5G35620
eif(iso)4E_Rv	ATACCCATCAGAACAGCTTCATT		

Table 2.5 List of oligonucleotides used as primers for qRT-PCR. All primers were designed to use a T_A of 58ºC.

2.13 Northern blotting

For northern blotting of viral siRNAs, 1-5µg of purified RNA was separated on a 0.75mm 15% denaturing polyacrylamide gel (section *2.21.9*). RNA sample concentrations were normalised with autoclaved, twice distilled water to a final volume of 10µl. Normalised samples were mixed with 10µl of RNA loading dye (section *2.21.8*) and heated at 65ºC for 5 minutes before loading onto the gel. The gel was pre-run at 50V for 30 minutes in a biorad tank (Biorad, Catalogue# 1658040) filled with 0.5X TBE and wells were thoroughly flushed with 0.5X TBE before sample loading. The RNA samples were separated for 30 minutes at 50V, then a further 60 minutes at 100V. The gel was removed from the glass casting plates and equilibrated in 20X SSC (section *2.21.10*) for 10 minutes. The separated RNA was blotted from the gel onto a positively charged nylon membrane (GE Healthcare Catalogue# RPN303B) by capillary blotting. The membrane was equilibrated in distilled water, then 20X SSC prior to setting up the transfer, and a transfer stack containing 20X SSC soaked filter paper and tissue paper was left overnight to allow sufficient blotting of RNA onto the membrane. After overnight blotting, the RNA was crosslinked to the membrane with UV at 120000µJoules (Statagene, UV Stratalinker 2400). The membrane was cut slightly above the upper xylene xyanol dye to allow the gel to be hybridised with a siRNA specific probe (lower section) and a U6 probe (upper section) for qualitative sample normalisation. Membranes were washed in hybridisation buffer (section *2.21.11*) at 40ºC for 30 minutes, and then hybridised with a ^{32}P-ATP-labelled probe in hybridisation buffer at 40ºC overnight. Probes were radioactively labelled using the following reaction: 2µl 10µM heat-denatured DNA oligonucleotide, 2µl 10X PNK buffer (Thermo Fisher, Catalaogue# EK0031), 1µl T4 polynucleotide kinase (Thermo Fisher, Catalaogue# EK0031), 5µl ^{32}P-ATP, 10µl water. The labelling reaction was incubated at 37ºC for 30 minutes. Un-incorporated nucleotides were removed from the labelled probe using a Microspin G-25 column (GE Healthcare, Catalogue# 27-5325-01) according to the manufacturer's instructions. 2µl of 0.5M EDTA was added to the labelled probe and the probe was denatured at 90ºC for 5 minutes before adding it to the hybridisation buffer. After overnight hybridisation, the membranes were washed twice in 2X SSC, 0.1% SDS at 40ºC for 10 minutes. The two sections of the cut membrane were aligned and wrapped together in Saran wrap and exposed to a phosphor imaging plate (Kodak Catalogue# SO230) for 24 hours. The radioactive signal was measured by scanning the phosphor plate in a Typhoon scanner (GE healthcare, Typhoon FLA 7000). Oligonucleotides used for Northern blotting are listed in Table 2.6, below.

Oligonucleotide name	Sequence (5'-3')	Target Locus	Gene ID
TuMV	GATTTGTGTTTGCTTTA ATCGTTTTGTGTATGTT ATGTTG	TuMV 5' UTR	-
U6-26	GCTAATCTTCTCTGTAT CGTTCC	U6	AT3G13855

Table 2.6 List of oligonucleotides used as probes for Northern blotting.

2.14 Western blotting

Samples were collected and powdered as previously described in section **2.10**. Total protein was extracted by adding 100µl of extraction buffer to 100mg of the lysed plant tissue. Samples were vortexed for 20s immediately after addition of protein extraction buffer (section *2.21.12*). Crude protein extract was centrifuged at 13,000 rpm at 4ºC for 10 minutes. The supernatant was transferred to a clean Eppendorf tube and the original tube containing the pelleted plant material was discarded. 5µl of the protein extract was used to determine the protein concentration using a Bradford assay (Sigma, Catalogue# B6916). Samples were normalized to 25µg for loading onto a polyacrylamide electrophoretic gel. 25µg total plant protein (in 13µl) was mixed with 2µl of 10X Bolt™ LDS sample loading buffer (Novex, Catalogue# B0007) and 2µl of 10X Bolt™ sample reducing agent (Novex, Catalogue# B0009). Samples were heated in a dry heat block at 70ºC for 10 minutes, and immediately loaded onto a pre-cast 7% tris-acetate NuPAGE™ gel (Novex, Catalogue# EA03555BOX). The gel was run at a constant 165V for 5 hours at 25ºC, then the gel was washed for 5 minutes in 20% ethanol and soaked for one minute in transfer buffer (Novex, Catalogue# BT0006 & BT0005) An overnight transfer to a 0.45µm nitrocellulose membrane (Bio-Rad, Catalogue# 1620115) was performed at a constant 15V at 4ºC, using the Novex™ miniblot module (Novex, Catalogue# B1000). After overnight transfer, the membrane was blocked for 1 hour at 25ºC in 5% milk TBSTT (section *2.21.13*). The blocking solution was removed and replaced with fresh 5% milk TBSTT containing the primary antibody at a dilution of 1:2000. The membrane was incubated with this primary antibody solution for 1 hour at 25ºC Primary antibody solution was discarded and the

membrane was washed 6x for 5 minutes with 5% milk TBSTT. The membrane was then incubated with Pierce® Goat Anti-Rabbit IgG (H+L) peroxidase conjugated secondary antibody (Thermo Fisher, Catalogue# 31466) diluted 1:10000 in 5% milk TBSTT for 1 hour at 25ºC. The secondary antibody solution was discarded and the membrane was washed 6x for 10 minutes with 5% milk TBSTT. The membrane was incubated for 2 minutes at 25ºC in Pierce® western blotting substrate solution containing equal volumes of ECL reagents A/B (Thermo Scientific, Catalogue# 32106) and SuperSignal® ECL reagents A/B (Thermo Scientific, Catalogue# 37071). After incubation, excess western blotting substrate solution was removed. The membrane was exposed to an X-Ray film (Thermo Scientific, Catalogue# 34089) to determine protein abundance.

2.15 Sanger sequencing

2.15.1 PCR and product purification

PCR products for sequencing were purified using the following SAP/EXO reaction: 5µl PCR product, 0.66µl shrimp alkaline phosphatase (New England Biolabs, Catalogue# M0371S), 0.06µl ExonucleaseI (New England Biolabs, Catalogue# M0293S), 4.28µl water. The reaction was incubated at 37ºC for 30 minutes and terminated by incubating at 80ºC for 10 minutes.

2.15.2 BigDye reaction and capillary sequencing

The BigDye v3.1 (Applied Biosystems, Catalogue# 4337455) terminator reaction was performed with the following reagents: 5µl SAP/EXO purified PCR product, 1.6µl of 2µM sequencing primer, 1.5µl Big Dye buffer, 1µl Big Dye, 0.9µl water. The Big Dye reaction was performed with the following cycling conditions: 96ºC 45s > (96ºC 10s > 50ºc 5s > 60ºc 4min) x 25. The forward direction primer used for PCR amplification was used as a sequencing primer. Sanger capillary sequencing was performed by Edinburgh Genepool (http://genepool.bio.ed.ac.uk/).

2.15.3 Sequence analysis

The CRISP-ID online software (Dehairs *et al.*, 2016) was used to align sequencing traces and identify indels, with 'background cut-off' set to 50% and 'max insert size' set to 10bp.

2.16 T7 endonuclease assay

RNA was removed from extracted TNA, to yield pure DNA, by adding 5µl of RNAse A/T1 (Thermo Fisher, Catalogue# EN0551) to TNA resuspended in 50µl TE buffer (section *2.21.6*), and incubated at 37ºC for 30 minutes. RNAse was removed from the sample by phenol/chloroform extraction, as outlined in section **2.10**. Primers were designed to produce a ~500bp amplicon by PCR with the predicted mutation site at approximately 200bp from one end of the amplicon (Table 2.7). For the denaturing/annealing reaction the total Q5 PCR product (10µl) was mixed with 1.5µl 10X NEB buffer 2 (NEB #B7002S) and 1.5µl water. This mixture was incubated at 95ºC for 10 minutes, then ramped from 95ºC-85ºC at a rate of -2ºC s^{-1}, then from 85ºC-25ºC at a rate of -0.3ºC s^{-1}. T7 enodnuclease I (NEB #B7002S) was diluted to a concentration of 2U μl^{-1} (in NEB buffer 2 #B7002S) and 2µl of this was added to the denaturing/annealing reaction product, then incubated at 37ºC for 1 hour. The total T7 digestion product (15µl) was loaded onto a 2% agarose gel and separated at 100V.

Oligonucleotide name	Sequence (5'-3')	T_A (ºC)	Extension time (s)	PCR product (bp)	Digest products (bp)
eif(iso)4E_T7_Fw	AAGAGTTAAATGCTC TGATGGAC	57	10	574	196 + 378
eif(iso)4E_T7_Rv	ACAAGTGAATTTCAG CATTCG				
GFP_T7_Fw	AAGGACGACGGGAA CTACAAG	62	10	431	106 +325
GFP_T7_Rv	GAACTATACAAACAT GATGAGCTTTAAGA				

Table 2.7 List of oligonucleotides used as PCR primers for the T7 endonuclease assay. The T_A and extension time are given for each primer pair, which were used for a Q5 PCR reaction as outlined in section *2.11.1*. The expected PCR product sizes and expected cleavage products after incubation with T7 endonucleaseI are also listed.

2.17 Physiological analysis of *eIF(iso)4E* mutants

2.17.1 Flowering time measurements

Seeds were carefully monitored to record the date of germination, defined by the first sign of a radical emerging from the seed coat. Mature plants were carefully monitored to record the onset of flowering, defined by the presence of a 1cm floral 'bolt' emerging from the foliar rosette. The difference between germination date and flowering date was used to determine the flowering time (in days) for each plant.

2.17.2 Dry weight measurements

4 week old *Arabidopsis* plants were harvested for dry weight measurements. Whole plants were prepared by carefully removing the plants and soil from the pots and soaking the plant roots in water to wash the soil from the root system. 2 plants per genotype were pooled and placed in pre-weighed waxed paper cups, giving 15 replicates of 2 pooled plants per genotype. The plants were dried in a Binder drying and heating chamber (Model E 28) set to 100^{o}C for 48 hours. The difference in weight between the empty paper cups and the paper cups with dried plants was used to estimate the dry mass of the 2 pooled plants.

2.18 *arf19* hypocotyl assay

2.18.1 Seed sterilisation

Seeds were sterilised by incubation with chlorine gas for 12 hours. Seeds were placed in 1.5ml Eppendorf tubes with lids left open which were placed inside a sealed plastic container, together with a beaker containing 97ml sodium hypochlorite (thin bleach) and 3ml of 10N HCl.

2.18.2 Growth conditions

Sterilised seeds were sowed, using a sterile toothpick, onto 10x10cm petri dishes containing 2,4-D (2,4-dichlorophenoxyacetic acid) selection media (section *2.21.14*). Seeds were stratified by wrapping the plates in tin foil and placing them in a cold room at 4^{o}C for 3 days. A 5 hour light pulse of 200µE white light at 21^{o}C was used to synchronise germination. After the light pulse, plates were wrapped in tin foil to exclude light and maintained at 21^{o}C for 8 days.

2.18.3 Hypocotyl measurements

After 8 days of growth in complete darkness, the seedlings were imaged under white light using a Canon powershot G16 digital camera with automatic shooting mode. JPEG images were imported into ImageJ software to measure hypocotyl lengths, defined as the distance (in mm) between the apical hook and the hypocotyl/root junction.

2.19 Statistical analyses

2.19.1 Standard error calculations

Error bars in bar charts represent the standard error of the arithmetic mean ($SE_{\bar{x}}$), calculated in excel by the following formulae:

$$SE_{\bar{x}} = \frac{s}{\sqrt{N}}$$

$$s = \sqrt{\frac{1}{N-1} \times \sum_{i=1}^{N} (x - \bar{x})^2}$$

Where:

s = standard deviation

N = number of samples

$(x_1 \ldots x_i)$ = individual sample values

$\bar{x}$ = mean of sample values

2.19.2 Analysis of Variance (ANOVA) and post-hoc comparison of means.

ANOVA calculations were performed using R software, along with the R Commander graphic user interface (GUI) (https://www.r-project.org/). When the global ANOVA indicated a significant difference in the means, Tukey's honest significant difference (HSD) test was used for a pairwise comparison of means.

2.19.3 Boxplot/Scatterplots

Boxplots/Scatterplots were plotted using R software, along with the R Commander graphic user interface (GUI). The solid lines in boxplots represent the median of the sample values, boxes represent the sample value upper and lower quartiles, error bars represent the sample value range, and circles represent outlier values. Outliers are automatically identified as values lying out with the most parsimonious normal distribution curve.

2.20 Figure mounting

Figures were compiled using inkscape software (https://inkscape.org/). Gel images were black/white inverted using GIMP software (version 2.8) (https://www.gimp.org/downloads/). The brightness and contrast of inverted gel images and phosphor plate scans were adjusted manually using GIMP software, applying changes uniformly across the entire image.

2.21 Buffer/solution recipes

2.21.1 Luria Bertani media (LB media) (pH 7.5)

- 10g Bacto-tryptone
- 5g Yeast extract
- 10g NaCl
- (15g Bacto-agar)
- H_20 to make 1L

2.21.2 Yeast Extract Peptone media (YEP media) (pH 7.5)

- 10g Yeast extract
- 10g Bacto-peptone
- 5g NaCl
- (15g Bacto-agar)
- H_20 to make 1L

2.21.3 Infiltration Media

- 1ml 1M MES
- 1ml 1M $MgCl_2$
- 150µl 0.1M Acetosyringone
- H_20 to make 100ml

2.21.4 1M Phosphate buffer (pH 7.0)

- 57.7ml 1M Na_2HPO_4
- 42.3ml 1M NaH_2PO_4

2.21.5 Tris Borate EDTA (TBE) buffer (10X) (pH 8.0)

- 121.1g Tris base
- 61.8g Boric acid
- 7.4g EDTA (disodium salt)
- H_20 to make 1L

2.21.6 Tris EDTA (TE) buffer (pH 8.0)

- 1ml 1M Tris-HCl (pH 8.0)
- 0.2ml 0.5M EDTA
- H_20 to make 100ml

2.21.7 Total Nucleic Acid Extraction Buffer (TNA EB) (pH 8.0)

- 0.75g glycine
- 2ml 0.5M EDTA (pH 8.0)
- 0.58g NaCl
- 2g SDS
- H_20 to make 100ml

2.21.8 RNA loading dye (2X)

- 10ml deionised formamide
- 10mM EDTA
- 1mg xylene cyanol FF
- 1mg bromophenol blue

2.21.9 15% polyacrylamide denaturing gel

- 4.2g urea
- 0.5ml 10X TBE
- 3.75ml 40% (w/v) 19:1 acrylamide:bis-acrylamide
- 2.5ml H_2O
- 70µl 10% (w/v) ammonium persulphate
- 3.5µl TEMED

2.21.10 Saline Sodium Citrate (SSC) buffer (20X) (pH 7.0)

- 175.3g NaCl
- 88.2g $Na_3C_6H_5O_7$
- H_20 to make 1L

2.21.11 Hybridisation buffer (pH 7.0)

- 57.7ml 1M Na_2HPO_4
- 42.3ml 1M NaH_2PO_4
- 28g SDS
- H_20 to make 400ml

2.21.12 Protein Extraction Buffer (Protein EB)

- 20µl 1M TrisHCl (pH 7.5)
- 5µl 1M $MgCl_2$
- 60µl 5M NaCl
- 50µl 0.1M DTT
- 100µl 1% NP40
- 40µl 25X cOmplete™ protease inhibitor cocktail (Roche, Catalogue# 11697498001)
- 725µl H_20

2.21.13 Tris Buffered Saline Tween Triton (TBSTT) 5% Milk

- 1.2g Tris
- 4g NaCl
- H_20 to make 500ml
- 0.5ml Tween
- 0.5 ml Triton X
- 25g milk powder

2.21.14 2,4-diclorophenoxyacetic acid (2,4-D) selection media (pH 5.7)

- 1g Phyto-agar
- 0.215g MS salts
- H_20 to make 100ml
- 27.63µl 2,4-diclorophenoxyacetic acid solution (1mg/ml in 100% ethanol)

Chapter 3: Environmental influences on viral RNA silencing

3.1 Chapter introduction

3.1.1 The importance of environmental conditions on the outcome of plant diseases.

The importance of environmental conditions as major determinants of disease was first recorded by Theophrastus in *Historia Plantarum* (350-287BC). These astute observations of the influences of the environment on plant diseases were the genesis of the so-called miasma theory, in which 'bad air' was postulated to be the causal agent of disease. This theory was expanded by Mathieu Tillet, in the context of plant diseases, where he described the effects of soil moisture, soil texture, wind, and nutrient deprivation on the onset of disease (Tillet, 1755). In the mid-nineteenth century, research by John Snow, Louis Pasteur, and Robert Koch collectively established the germ theory of disease, in which pathogens are recognised as the causal determinants. While the miasma and germ theories are not mutually exclusive, the popularity of the latter resulted in many plant pathologists focussing on identifying the pathogens associated with diseases. In contrast, research into the influence of the environment on plant diseases was scarce for many years. Lewis Jones strongly advocated that both biotic and abiotic factors should be considered when studying plant diseases (Jones, 1924). In the context of plant virology, James Johnson pioneered research into environmental influences of diseases, giving particular attention to the role of temperature on the onset of viral symptoms (Fulton, 1984). Specifically, it was observed that symptoms caused by Tobacco Mosaic Virus (TMV) and Potato Virus X (PVX) on Tobacco and Potato, respectively, were less severe at warmer growth temperatures (Johnson, 1921, Johnson, 1922). The term 'heat masking' was introduced to describe this negative correlation between viral symptoms and increasing growth temperatures. Similar heat masking effects were later observed for field-grown Sorghum infected with Maize Dwarf Virus (Hine *et al.*, 1970).

3.1.2 Thermotherapy as a method for cleaning virus infected plant material

In parallel to the field studies describing the heat masking phenomenon, experiments concerning the effects of temperature on virus recovery under controlled laboratory

conditions yielded similar results. It was found that peach trees (Kunkel, 1936) and periwinkle plants (Kunkel, 1941) could be cured from virus infections by incubating them at 35^{o}C or above for at least two weeks. Likewise, incubating virus infected plants at 36^{o}C for three to four weeks resulted in recovery from a wide range of viruses on a range of different hosts (Kassanis, 1952, Kassanis, 1954). This phenomenon, now known as thermotherapy, is widely utilised as a phyto-sanitation strategy and is particularly useful for eliminating viruses from vegetatively propagated and/or perennial crops (Mink, 1998, Panattoni *et al.*, 2013). While the early descriptions of thermotherapy involved incubating whole plants at different temperatures (Kassanis, 1952, Kassanis, 1954, Kunkel, 1936, Kunkel, 1941) current application of this method frequently involves incubating dissected shoots and meristems on plant tissue culture media (Chen & Sherwood, 1991, Manganaris *et al.*, 2003, Tan *et al.*, 2010, Torres *et al.*, 2000). Table 3.1 summarises the diversity of crop diseases that are commonly treated by thermotherapy.

Plant Type	Plant Name	Virus Families
Woody perenials	Grape	*Bromoviridae, Closteroviridae, Comoviridae, Flexiviridae, Secoviridae, Tymoviridae*
	Apple	*Betaflexiviridae, Bromoviridae, Flexiviridae, Potyviridae*
	Peach	*Bromoviridae*
	Pear	*Flexiviridae, Tombusviridae*
	Apricot	*Potyviridae*
	Plum	*Bromoviridae*
	Raspberry	*Idaeoviruses*
Herbaceous annuals	Garlic	*Flexiviridae, Potyviridae*
	Potato	*Alphaflexiviridae, Flexiviridae Potyviridae*
	Sweet potato	*Geminiviridae, Potyviridae*
	Sugarcane	*Potyviridae*
	Artichoke	*Comoviridae, Potyviridae*
	Chrysanthemum	*Bromoviridae, Flexiviridae*

Table 3.1 List of commercially relevant virus infections which are treated by thermotherapy. This table has been adapted from the following review (Panattoni *et al*., 2013) wherein references to the primary literature can be found.

The mechanism of thermotherapy was originally proposed to be a result of heat directly de-stabilising virus particles (Kunkel, 1936). This direct thermal inactivation of virus particles can be quantified by heating purified viruses at different temperatures and assessing the viability of the heated virions by inoculating indicator plants, which form lesions upon virus infection. As the logarithm of the number of lesions directly correlates with the logarithm of the concentration of viable virions in the inoculum (Holmes, 1929), this assay can be used to calculate the so-called thermal inactivation point which is defined as the temperature required to inactivate a virus, for a heating period of ten minutes (Dijkstra & de Jager, 1998). Interestingly, when thermal inactivation points were compared to the efficacy of *in planta* thermotherapy for a wide range of virus-host examples, no correlation was found. Furthermore, the

temperatures at which thermotherapy becomes effective were found to be consistently lower than the thermal inactivation points for the viruses tested. This led to the conclusion that direct thermal inactivation of virus particles was not likely to be the mechanism of thermotherapy. Instead, it was proposed that virus replication and host defence are in a state of dynamic equilibrium which is strongly influenced by temperature (Kassanis, 1954). While it was recognised that the reduction of viral titres at higher temperature could be the result of either suppression of viral synthesis or induced host-mediated virus degradation, the former was generally favoured (Harrison, 1956, Kassanis, 1954). Recently however, there is mounting evidence that thermotherapy is more likely to be a result of host defence becoming more active at higher temperatures (Chellappan *et al.*, 2005, Havelda *et al.*, 2005, Ma *et al.*, 2016, Szittya *et al.*, 2003, Velázquez *et al.*, 2010, Qu *et al.*, 2005).

3.1.3 Virus-induced RNA silencing is promoted by increasing temperatures

Early insight into the molecular mechanisms underlying the phenomena of heat masking and thermotherapy came from studies on the *Tombusviridae* family of viruses. Like *Potyviridae*, virions of *Tombusviridae* consist of a solitary (+)ssRNA genomic sequence encapsidated by a multimeric structure of virus-encoded coat proteins (Yamamura & Scholthof, 2005). The type member Tombusvirus, Tomato Bushy Stunt Virus (TBSV), was found to be the causative agent for the devastating tomato plant decline (TPD) disease (Gerik *et al.*, 1990). Moreover, it was shown that both symptoms and viral titres were significantly reduced at high temperatures when host plants were grown over a range of temperatures from 16^{o}C to 28^{o}C (Gerik *et al*., 1990). This report was one of the first to directly link the temperature-mediated recovery of virus infection to a reduction in viral titres. An Enzyme Linked Immuno-Sorbant Assay (ELISA) was used to quantify the viral titres, giving a more direct quantification of viral titres to previous studies relying on symptom development and lesion assays on indicator plants to estimate viral titres (Kassanis, 1952, Kassanis, 1954). Another study characterised this in more molecular detail by infecting *N. benthamiana* protoplasts with TBSV and incubating them at 27^{o}C or 32^{o}C (Jones *et al.*, 1990). Genomic and sub-genomic TBSV RNAs were quantified by Northern blotting, Ethidium Bromide staining of electrophoretically separated RNA, and measuring incorporation of tritiated uridine. Using these methods, it was shown that

all TBSV RNAs accumulated at lower levels when protoplasts were incubated at 32ºC compared to 27ºC. The first direct evidence that virus-induced PTGS could be the main molecular mechanism responsible for temperature-activated virus recovery, was obtained from experiments using another Tombusvirus, Cymbidium Ringspot Virus (CymRSV) (Havelda *et al*., 2005, Szittya *et al*., 2003). Accumulation of CymRSV siRNAs was found to increase as growth temperatures increased, over a range of 15ºC to 27ºC. Both symptoms and viral titres were negatively correlated with the accumulation of siRNAs across the temperature range, indicating that temperature-activated viral PTGS is likely to be the main mechanism behind thermal recovery of virus infection. It was suggested that the DCL proteins responsible for generating vsiRNAs must become more active at higher growth temperatures. Interestingly, siRNAs derived from both viruses and transgenes were found to hyperaccumulate as temperature increased, whereas miRNA abundance was unaffected by temperature. This suggests that specific PTGS pathways are affected by temperature (Szittya *et al*., 2003).

3.1.4 Genetic components of temperature-activated viral RNA silencing

Temperature activation of viral PTGS leading to recovery from viral infection is not limited to the family *Tombusviridae*, but appears to be a general phenomenon. Reports linking thermal recovery to an induction of viral PTGS have been made for *Geminiviridae* (Chellappan *et al*., 2005), *Vigraviridae* (Qu *et al*., 2005), *Ophioviridae* (Velázquez *et al*., 2010) *Alphaflexiviridae* (Ma *et al*., 2016), and *Potyviridae* (Ma *et al*., 2016). Moreover, some of the genetic determinants for this response are beginning to emerge. By knocking-down expression of *RDR6* in *N. benthamiana* it was shown that a functional RDR6 is required for full temperature-induced silencing of TMV and Turnip Crinkle Virus (TCV) (Qu *et al*., 2005). Similarly, *Arabidopsis* plants harbouring mutations in various genes involved in RNA silencing pathways were used to test the genetic components necessary for temperature-activated PTGS of TCV. DCL2 was found to be necessary for the hyperaccumulation of TCV-induced siRNAs at 26ºC compared to 18ºC, and DCL2, AGO2 and HEN1 were all necessary for host survival when plants were infected with TCV and shifted to 26ºC (Zhang *et al*., 2012). Recently, the *RDR1* orthologue in *Nicotiana glutinosa* (*NgRDR1*) was shown to be transcriptionally activated by raising growth temperatures from 22ºC to 30ºC, resulting

in higher accumulation of vsiRNAs and consequently an active defence against PVY and PVX (Ma *et al*., 2016). It has previously been shown that the *NgRDR1* promoter contains a heat shock element (HSE) which is likely to be the cis-regulatory element responsible for the observed heat-activated transcription of this gene (Liu *et al.*, 2009). Intriguingly, *NgRDR6* did not show the same transcriptional activation by raising temperature (Ma *et al*., 2016). Taken together, various components of viral RNA silencing pathways have been implicated in temperature-activated PTGS though the most critical component may vary depending on the specific host-virus interaction. This may be partly due to differential sub-functionalisation of paralogues between different host species. For example *RDR6* and *RDR1* appear to be the major temperature sensitive components in *N. benthamiana* and *N. glutinosa*, respectively. This differential sub-functionalisation could be explained, at least partly, by the fact that *N. benthamiana* plants do not possess a functional *RDR1* homologue due to a natural loss-of-function mutation at this locus (Bally *et al.*, 2015, Yang *et al.*, 2004a). Additionally, different viruses may subvert temperature-sensitive host defence genes to varying degrees. This is likely to occur, as viral suppressors of RNA silencing (VSRs) encoded by different viruses are known to interfere with different components of RNA silencing pathways and with varying efficiencies (Burgyan & Havelda, 2011). Moreover, there have been suggestions that certain VSRs, such as the potyviral HC-Pro may exhibit temperature dependent activity (Del Toro *et al.*, 2015). Hence, recent reports have added further mechanistic detail to support the notion that thermal recovery from virus infection is largely mediated by viral PTGS, but the exact mechanism may vary for different virus and host species.

3.1.5 Effects of temperature on potyvirus infection

Reports on the influence of temperature on potyvirus infections have been seemingly contradictory. For example, despite the fact that thermotherapy has been applied successfully to several potyvirus infections (López-Delgado *et al.*, 2004, Jeeva *et al.*, 2004, Manganaris *et al*., 2003, Mangrauthia *et al.*, 2009, Ramgareeb *et al.*, 2010) , a number of studies have suggested that the severity of PVY infection is worsened by increasing temperatures (Kerlan *et al.*, 2011, Le Romancer & Nedellec, 1997). These caveats may be consolidated when various additional factors are taken into consideration. Firstly, quantifying potyvirus infection rates at the field level generally

gives a positive correlation between infection rate and temperature increases (Hasan & Rashid, 2015). It is likely, however, that this is mainly due to modest rises in temperature positively affecting vector-mediated transmission (Broadbent & Hollings, 1951, Newman, 2004, Singh *et al.*, 1988). Secondly, studies which focus on the effects of temperature on the severity of potyvirus-induced symptoms alone (Le Romancer & Nedellec, 1997) may fail to consider that symptom severity and viral titre may not always correspond. Indeed, several studies have shown that PVY titres significantly decrease at raised growth temperatures (Chung *et al.*, 2016, Del Toro *et al*., 2015) and that it can at least partly be explained by increased efficacy of virus induced PTGS (Ma *et al*., 2016). A third consideration is the temperature range which is applied in the study. For example, it was recently shown for TuMV-infected *Brassica rapa*, that both symptom severity and viral titres steadily increase when plants were grown over a range from 13^{o}C to 28^{o}C in 5^{o}C increments. However, plants grown at 33^{o}C exhibited a dramatic recovery phenotype and a concurrent decrease in viral titre (Chung *et al*., 2015). A similar trend was observed for *N. benthamiana* plants infected with PVY (Chung *et al*., 2016). This is best conceptualised by the theory of a dynamic equilibrium between viral synthesis and degradation originally proposed by Kassanis (Kassanis, 1954) and Harrison (Harrison, 1956). It is possible that gradual increases in ambient temperature promote viral replication and/or translation resulting in higher viral titres as temperature increases over a narrow ambient range, and at early time points. At later time points and higher growth temperatures, activation of viral PTGS may eventually shift the equilibrium towards a net reduction in viral titre and hence recovery from the infection. Failure to consider these temporal dynamics may have led to some erroneous interpretations from previous studies.

3.1.6 Effects of light on plant virus infections

While the effects of temperature on viral infections have been relatively well characterised, the effects of light (another important environmental input) on resistance to viruses, and the associated mechanisms remain obscure. It has long been noted that there are significant seasonal effects on virus resistance. Specifically, plants tend to be more susceptible in winter months when both the day length and light intensity are reduced (Smith & Bald, 1935). To determine whether this seasonal effect could be partly explained by differences in light intensity, *N. tabacum* and *N. glutinosa*

plants were grown under high and low intensity light before inoculations with Tobacco Necrosis Virus (TNV), TMV and TBSV. Plants were found to be more susceptible to each of these viruses when grown under low light intensities (Bawden & Roberts, 1947). However, as plants were maintained under different light intensities for several weeks prior to inoculation, the observed differences in viral resistance could be explained by indirect effects of the light treatment, such as differing physiological age of the inoculated leaves or differences in leaf toughness. To rule out these indirect effects, similar experiments were conducted where plants were exposed to high light intensities, low light intensities, or darkness for short periods ranging from 1-48 hours prior to inoculation. Again, low light intensities and darkness resulted in higher rates of infection compared to plants treated with high light intensities (Bawden & Roberts, 1948, Helms & McIntyre, 1967, Matthews, 1953b). Attempts to link these phenomena to differences in photosynthetic activity were made by infusing detached leaves with a variety of sugars (Yarwood, 1952b) or acidic by-products of photosynthesis (Wiltshire, 1956) prior to viral inoculation. These experiments were mainly inconclusive, but a link between compromised photosynthesis and increased viral susceptibility was made by the observation that subjecting plants to low CO_2 levels increased viral susceptibility to a similar extent as dark treatments prior to inoculation (Wiltshire, 1956). More recently, this association between photosynthetic activity and viral resistance was reinforced by observations that PTGS-mediated disruption of core photosynthetic components (including PbsO/Q, FtsH, and RuBisCO) and in *N. benthamiana* resulted in an increased susceptibility to TuMV (Manfre *et al*., 2011).

Curiously, while treating plants to low light intensities or darkness immediately before viral inoculation resulted an increase in susceptibility, subjecting plants to these conditions after viral inoculation appears to have the opposite effect (Bawden & Kleczkowski, 1955, Bawden & Roberts, 1948, Helms & McIntyre, 1967, Matthews, 1953b). Interestingly, a similar effect was observed for temperature treatments, whereby heat treatments prior to inoculation resulted in greater viral susceptibility whereas post-inoculation heating led to increased resistance (Kassanis, 1954).

Taken together, it seems that some of the seasonal differences in viral resistance can be explained by differing light intensities which in turn may be a result of indirect effects concerning plant development and physiology, and possibly also direct effects

on photosynthetic activity. To evaluate how light intensities directly affect the equilibrium between viral synthesis and degradation however, it may be more pertinent to consider the effects of different light intensities after viral inoculation. While this experimental design does not reflect the biological scenario of seasonal effects on viral infection, it may be useful for interpreting differences in viral resistance across day-night cycles (Helms & McIntyre, 1967, Matthews, 1953a) as in these conditions the light intensity shortly after viral inoculation, as the virus begins to replicate within the host, may be a key determinant for the outcome of infection. In addition, the relevance of shifting infected plants to different light intensities may be interesting in artificial scenarios. By analogy to thermotherapy, if treatment with low light intensities after viral inoculation can reduce viral titres, control of light levels could be used in combination with heat treatment to efficiently cure plants from viral diseases. Despite these potential applications, the mechanisms by which different light intensities affect viral establishment and recovery after viral inoculation (as opposed to pre-conditioning effects) are unknown. As it was recently reported that light intensity can dramatically affect the efficacy of transgene-induced RNA silencing (Kotakis *et al.*, 2010, Kotakis *et al.*, 2011) it is tempting to speculate that some of the effects of light intensity on viral resistance could be explained by differential activities of viral PTGS under these different growth environments.

3.2 Chapter aims

This chapter aims to investigate whether the mild heat treatments, analogous to the practice of thermotherapy, are effective for inducing recovery from TuMV-GFP infections. Additionally, the potential for inducing viral recovery by incubating plants under different light intensities will be investigated. Furthermore, the potential role of TuMV-vsiRNAs, and the genetic determinants underlying these putative responses will be considered.

3.3 *In vivo* monitoring of early TuMV-GFP infection at different temperatures in local leaves of *Nicotiana benthamiana*

To investigate the effects of temperature on the early stages of TuMV infection, *in vivo* monitoring of TuMV-GFP infection foci was conducted for rub-inoculated (local) leaves. *N. benthamiana* plants were chosen as a host because their large and flat leaves make them suitable for non-destructive imaging of fluorescently tagged viruses. Plants were grown at 21^{o}C for 4 weeks prior to inoculation, and held at 21^{o}C for one day after inoculation before shifting the plants to either 15^{o}C, 21^{o}C or 27^{o}C. The reason for this one day of incubation at 21^{o}C after rub-inoculation was to allow the inoculated leaves to recover from the tissue abrasion, thereby limiting the effects of the temperature shift on leaf survival due to tissue damage. The rationale for first studying the effects of a temperature shift on early infection foci in local leaves was to enable a direct analysis of the effects of temperature on viral replication and host defence, without needing to consider potentially confounding factors such as phloem transport rate, which would likely affect the rate of systemic infection at different temperatures. Infection foci became visible by two days post inoculation (dpi), corresponding to one day post shift (dps). At this time-point (2dpi/1dps), few and barely visible infection foci were present in the 15^{o}C samples, whereas many bright foci were observed at 21^{o}C and 27^{o}C. Furthermore, the rate of foci development (assessed by foci number and size) was lowest at 15^{o}C and highest at 21^{o}C. At 27^{o}C necrotic lesions resembling a hypersensitive response (HR) became evident from 3dpi/2dps (Figure 3.1).

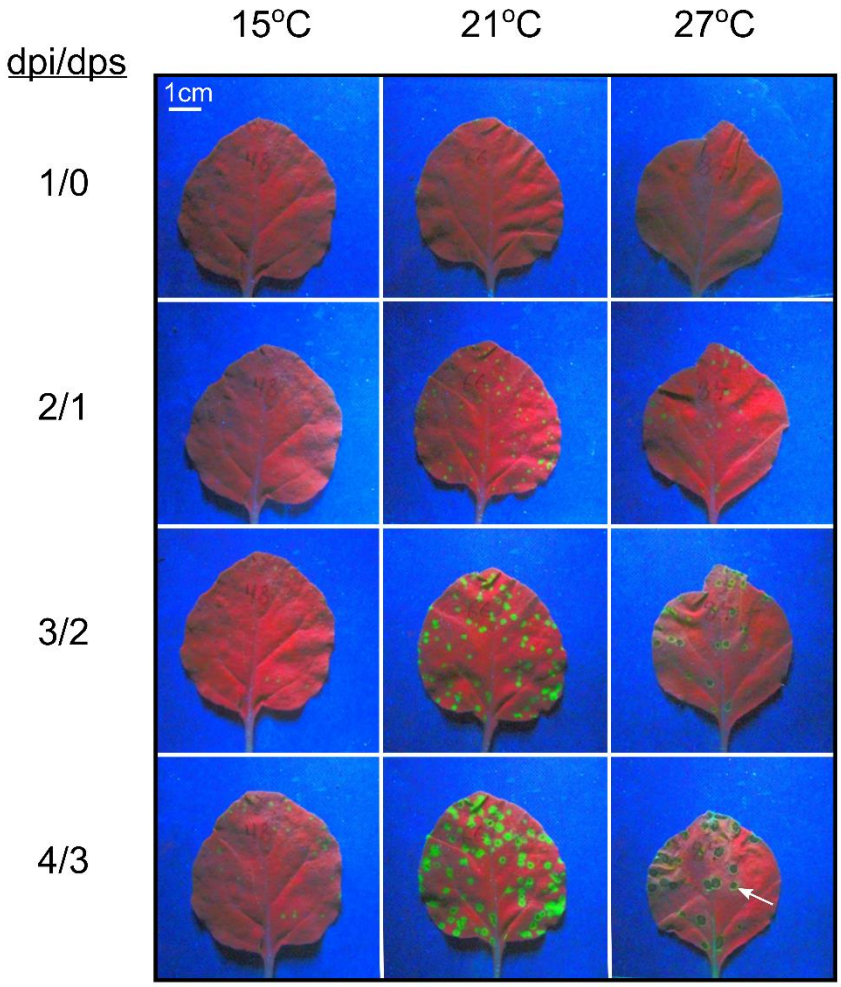

Figure 3.1 Time course UV imaging of *N. benthamiana* leaves rub-inoculated with TuMV-GFP infectious sap. Plants were shifted to 15°C, 21°C, or 27°C at 1 day post inoculation (dpi). Hence, days post shift (dps) are one less than dpi for each time interval. Plants were maintained at 21°C until 1dpi/0dps and grown under 50µE white light throughout. White arrow indicates a necrotic lesion.

To quantify the qualitative results in Figure 3.1, a customised software was created to measure several parameters of the infection foci. Briefly, an image processing script was written in MatLab^{TM} to separate the different colours of the inoculated leaf (blue background, red leaf, and green GFP infection foci). Information from the three colour channels was used to create binary images of the leaf area (Figure 3.2B) and infection foci (Figure 3.2C) These binary images were then used to calculate the leaf area, total GFP area, and number of infection foci (Figure 3.2D). These methods were used to quantify the qualitative data presented in Figure 3.1, using total GFP content (Figure 3.3A) and number of foci per cm^2 of leaf surface (Figure 3.3B) to assess the establishment of TuMV-GFP infection. Samples were collected at the end of the time-series (5dpi/4dps) to measure the abundance of viral RNA by qRT-PCR (Figure 3.3C). The results for each of these quantitative parameters indicate that the establishment of infection is promoted at 21^{o}C compared to 15^{o}C (Figure 3.3), possibly due to a greater rate of viral replication and/or translation. At 27^{o}C, foci became visible earlier than at 15^{o}C (Figure 3.1) but at later time points the level of infection was reduced at 27^{o}C compared to 21^{o}C. This indicates that both 21^{o}C and 27^{o}C promote the early establishment of infection, but growth at 27^{o}C results in a reduction of viral RNA, suggestive of a temperature-stimulated defence response. The presence of necrotic lesions in the plants shifted to 27^{o}C, along with the reduction in viral RNA in these plants suggests that part of the recovery from TuMV-GFP infection could result from a temperature stimulation of HR.

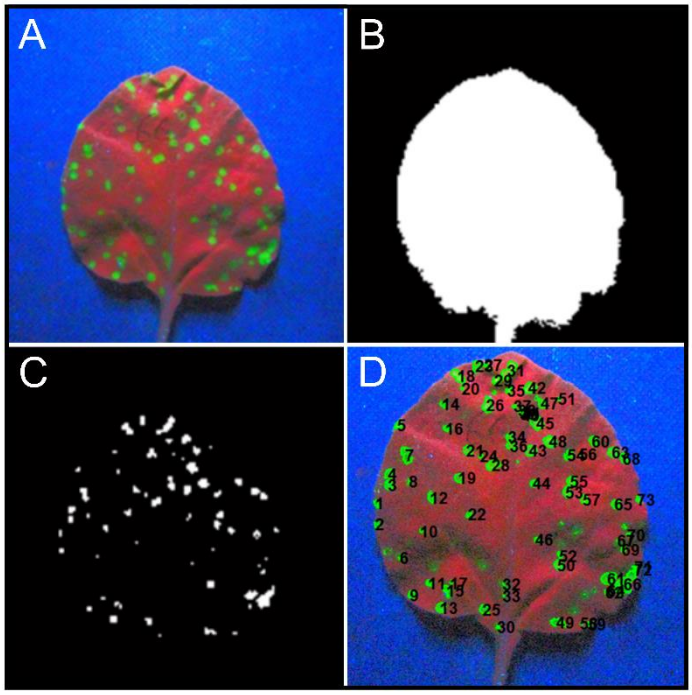

Figure 3.2 Automated quantification of TuMV-GFP infection foci. A Matlab script was created to parse the red-green-blue pixel data (**A**) to create binary image template of the leaf (**B**) and TuMV-GFP foci (**C**). A pattern recognition function enabled infection foci to be automatically counted and numbered as an overlay on the original image file (**D**). Using a known pixel/cm^2 conversion factor, the number of foci and total TuMV-GFP coverage per cm^2 of leaf could be calculated.

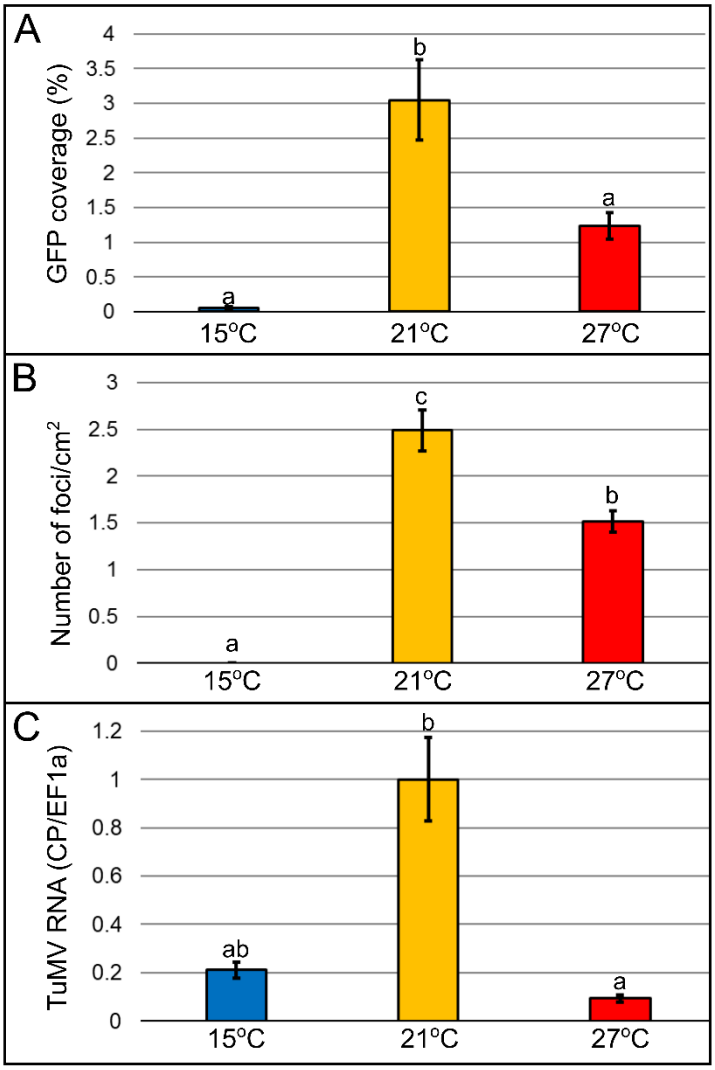

Figure 3.3 Quantification of TuMV-GFP infection for plants maintained at 15ºC, 21ºC, and 27ºC. Plants were maintained at 21ºC until 1dpi/0dps and grown under 50µE white light throughout. The Matlab image processing script was used to measure: **A** total TuMV-GFP coverage, and **B** number of infection foci per cm^2 leaf area. Bars represent the arithmetic mean of 6 independently infected leaves at 3dpi/2dps. **C** qRT-PCR to measure the viral RNA in the same leaves shown in **A**/**B**. Samples were collected at 5dpi/4dps and 3 leaves were pooled into one sample. Hence bars show the arithmetic mean of 2 samples. The *EF1a* 'housekeeping gene' was used for ΔΔCp calculations, and values are plotted relative to the mean TuMV signal at 21ºC. Error bars on all charts show the standard error in the arithmetic mean (SEM). Tukey's Honest Significant Difference (HSD) pairwise comparison of means revealed statistically significant groupings, as indicated by letters above each bar.

3.4 Analysis of early TuMV-GFP infection at different temperatures in local leaves of *Arabidopsis thaliana*

While using *N. benthamiana* as a host for TuMV-GFP infection was a useful system to assess the early dynamics of infection, using non-destructive time-course imaging, the presence of a strong (and potentially temperature sensitive) HR would complicate analysis of the role of PTGS in temperature dependent viral recovery. For this reason, subsequent experiments were conducted using *Arabidopsis*. The main advantage of using *Arabidopsis* is that the Columbia ecotype (Col-0) which was used for experiments does not exhibit a necrotic HR to the TuMV UK1 strain which was used in this study (Manacorda *et al.*, 2013). A further advantage of using *Arabidopsis* as a susceptible host is that it allows the well characterised RNA silencing mutants to be utilised to test the hypothesis that temperature regulation of viral PTGS might influence the outcome of infection. *Arabidopsis* plants were grown at 21ºC for 4 weeks and then rub-inoculated with TuMV-GFP sap. Plants were maintained at 21ºC for 4 days after rub-inoculation before shifting them to 15ºC, 21ºC, or 27ºC. The reason for this longer incubation time compared to the experiments with *N. benthamiana* is that infection foci in *Arabidopsis* only became visible from 4dpi, and it was important to verify that rub-inoculations were successful before shifting to different temperatures, as *Arabidopsis* leaves were more difficult to successfully inoculate than *N. benthamiana*. By 7dpi/3dps the different temperature treatments had noticeable effects on the establishment of infection. The GFP fluorescence resulting from TuMV-GFP infection was highest at 27ºC and lowest at 15ºC (Figure 3.4A). qRT-PCR to directly measure viral RNA from the same samples revealed an approximately 2-fold increase in viral RNA at 21ºC compared to 15ºC and a further ~2-fold increase at 27ºC compared to 21ºC (Figure 3.4B). These results confirm the previous findings that the higher growth temperatures of 21ºC and 27ºC promote the early establishment of TuMV infection. As previous reports suggested that vsiRNA accumulation increases as the temperature is raised over a range of 15ºC to 27ºC (Chellappan *et al*., 2005, Havelda *et al*., 2005, Szittya *et al*., 2003), RNA from the same 7dpi/3dps samples was used to detect vsiRNAs by Northern blotting (Figure 3.4C).

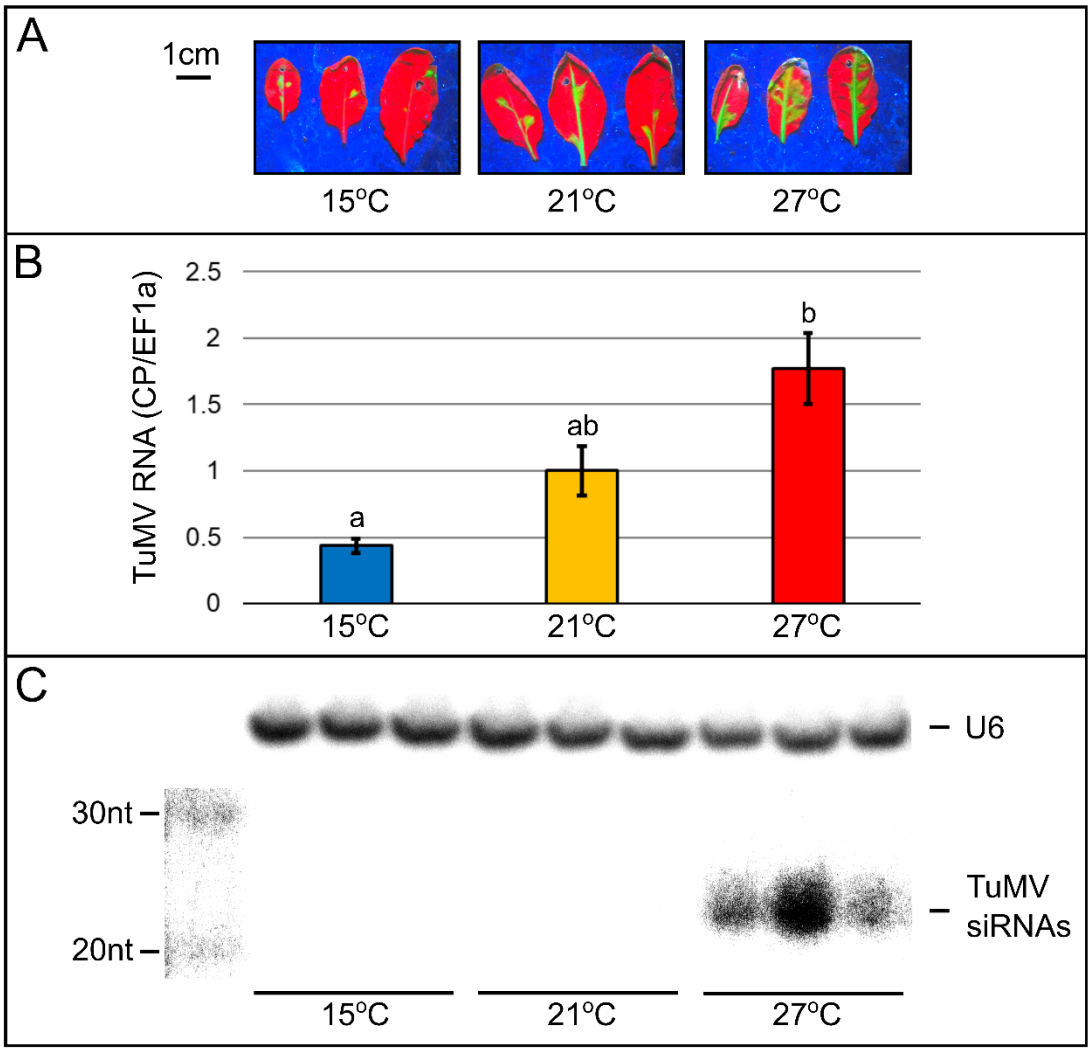

Figure 3.4 Relative abundance of TuMV-GFP RNA and vsiRNA abundance in local tissue of rub-inoculated *Arabidopsis* plants maintained at 15°C, 21°C, and 27°C. Samples were collected at 7dpi/3dps. Plants were maintained at 21°C until 4dpi/0dps and grown under 100µE white light throughout. **A** UV imaging of TuMV-GFP in locally infected leaves. **B** qRT-PCR to measure the TuMV-GFP RNA. 3 infected leaves from the same plant were pooled into one sample and 3 samples from 3 independent plants were used as biological replicates. Hence, bars represent the arithmetic mean of 3 samples ± SEM. Data are plotted relative to the mean of the 21°C samples. Tukeys HSD pairwise comparison of means revealed statistically significant groupings, as indicated by letters above each bar. **C** Northern blot of TuMV-derived siRNAs using the same samples as in **A**/**B**. A U6-transcript probe was used as a loading control and a small RNA ladder (Thermo Fishe, Catalogue# AM7778) was used as a size marker.

In accordance with previous studies, vsiRNAs were abundant in plants shifted to 27ºC but were not detectable for plants kept at 15ºC or 21ºC. The fact that viral RNA was greatest at 27ºC despite a high abundance of siRNAs could result from a few scenarios. One hypothesis is that the TuMV VSR, Hc-Pro, could disable vsiRNAs so efficiently (and possibly more efficiently at higher temperatures) that the virus becomes effectively immune to the hyper-accumulation of siRNAs at 27ºC. Alternatively, the vsiRNAs detected at 7dpi may be biologically active, but the reduction in viral RNA may only become apparent at later time-points. To test the latter hypothesis, samples were taken at 17dpi/10dps. At this more advanced stage of infection, GFP was clearly visible in local leaves maintained at 15ºC and 21ºC, while at 27ºC the GFP signal was very weak, suggesting an active antiviral defence response (Figure 3.5A). qRT-PCR confirmed that the viral RNA was significantly reduced at 27ºC compared to 21ºC, though still higher than the viral RNA of plants kept at 15ºC (Figure 3.5B). Furthermore, detection of vsiRNAs by Northern blotting confirmed that virus-induced PTGS becomes more active at 27ºC compared to lower temperatures (Figure 3.5C). Taken together, these data suggest that increasing temperature over a range of 15ºC to 27ºC promotes both the synthesis and degradation of TuMV. At 21ºC in locally infected *Arabidopsis* tissue, this thermal equilibrium results in a net increase in viral RNA, compared to lower temperatures. Increasing the temperature to 27ºC results in a further net increase in viral RNA at early stages of infection, despite a large increase in vsiRNA abundance. At later points in infection however, the hyperaccumulation of vsiRNAs at 27ºC shifts the thermal equilibrium towards a net reduction of viral RNA. This model, explaining the effects of temperature shifts on TuMV infection in local tissue, is outlined diagrammatically below (Figure 3.6B).

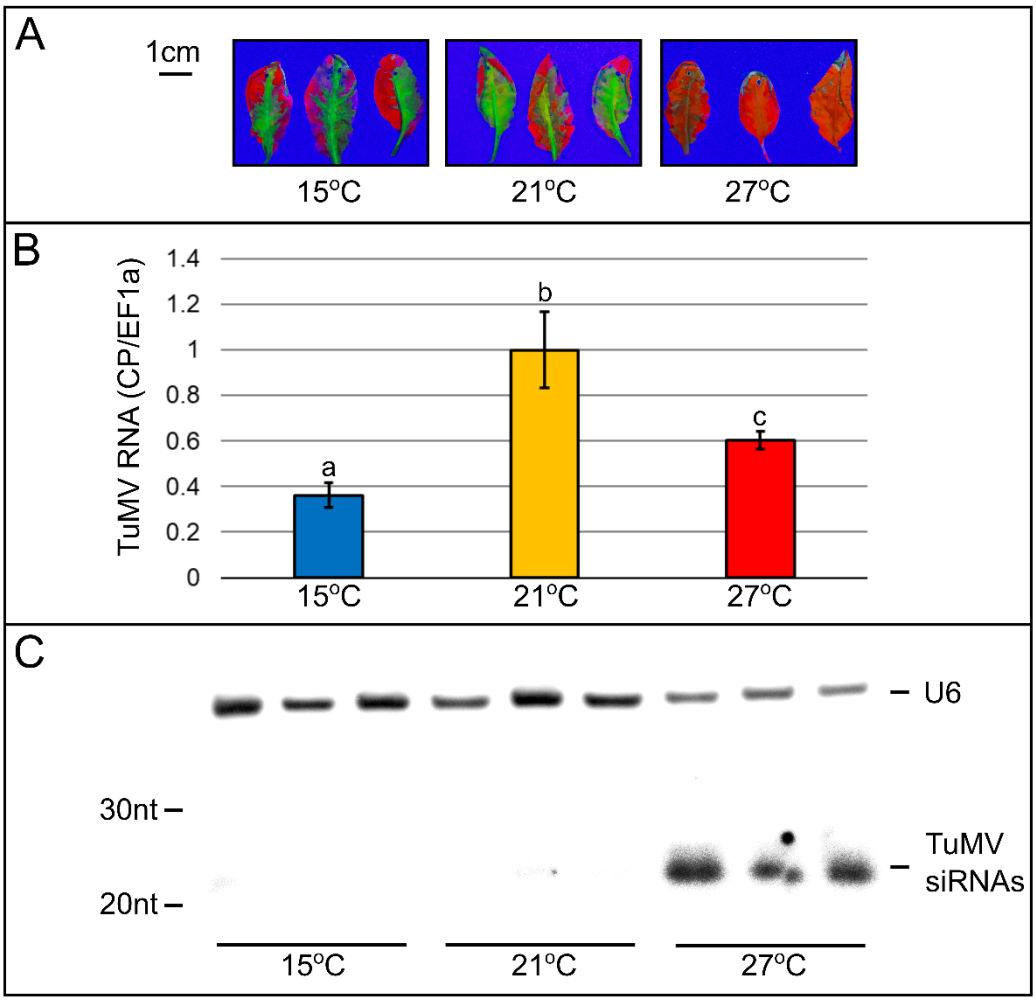

Figure 3.5 Relative abundance of TuMV-GFP RNA and vsiRNA abundance in local tissue of rub-inoculated *Arabidopsis* plants maintained at 15ºC, 21ºC, and 27ºC. Samples were collected at 17dpi/10dps. Plants were maintained at 21ºC until 7dpi/0dps and grown under 100µE white light throughout. **A** UV imaging of TuMV-GFP in locally infected leaves. **B** qRT-PCR to measure the TuMV-GFP RNA. 3 infected leaves from the same plant were pooled into one sample and 3 samples from 3 independent plants were used as biological replicates. Hence, bars represent the arithmetic mean of 3 samples ± SEM. Data are plotted relative to the mean of the 21ºC samples. Tukeys HSD pairwise comparison of means revealed statistically significant groupings, as indicated by letters above each bar. **C** Northern blot of TuMV-derived siRNAs using the same samples as in **A**/**B**. A U6-transcript probe was used as a loading control.

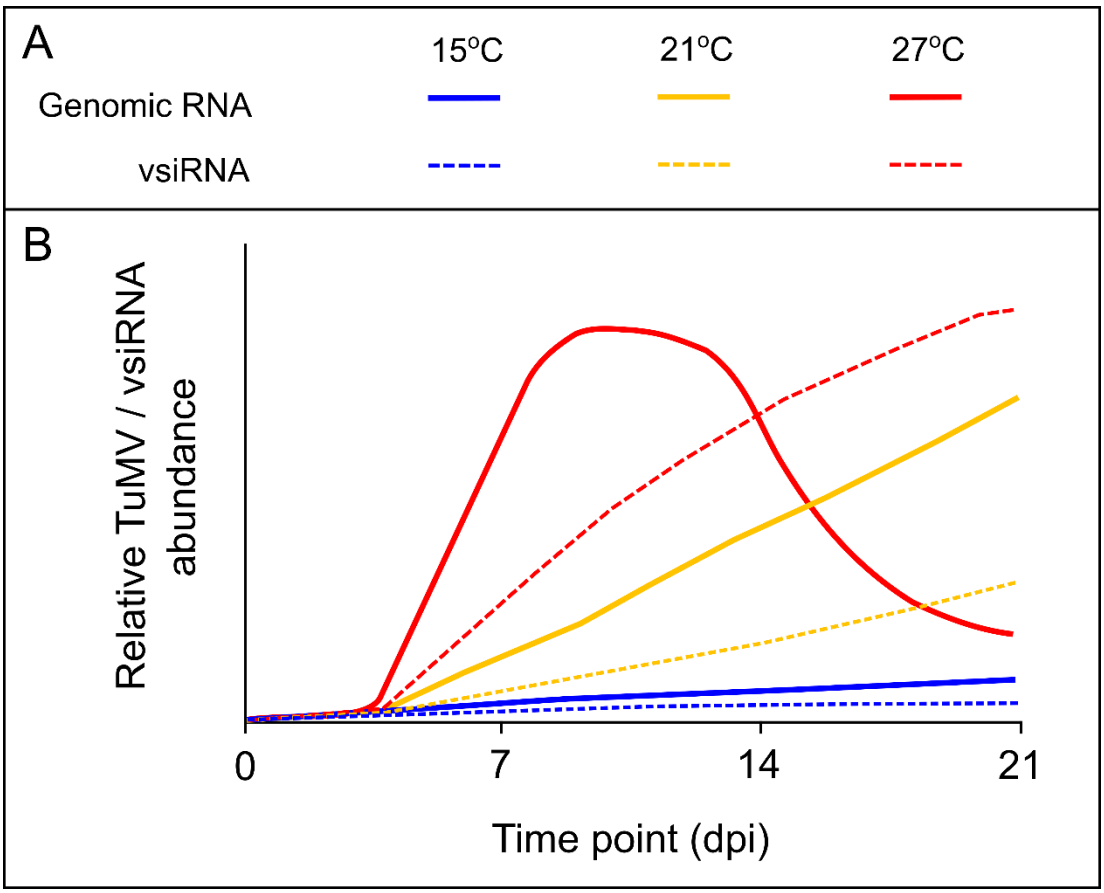

Figure 3.6 'Thermal-equilibrium' model to explain the effects of temperature shifts on the accumulation of genomic RNA and vsiRNAs during TuMV infections in *Arabidopsis* local tissue. **A** Figure key: Genomic RNA and vsiRNAs are indicated by solid and dashed lines, respectively. Blue/orange/red lines indicate 15°C/21°C/27°C growth temperatures, respectively. **B** Progression of TuMV genomic RNA and vsiRNA accumulation at different growth temperatures. At early time-points (~7-14dpi), the accumulation of TuMV genomic RNA is greatest at 27°C due to temperature-stimulated synthesis, despite a concurrent increase in vsiRNA production. The hyperproduction of vsiRNAs at 27°C eventually results in an equilibrium shift towards net degradation (~14-21dpi), resulting in lower TuMV RNA compared to 21°C. At this time point, viral RNA is lowest at 15°C because of the very low rate of virus synthesis at this low temperature.

3.5 Analysis of TuMV-GFP infection and recovery at different temperatures in systemic tissue of *Arabidopsis thaliana*

As incubation at 27ºC results in a reduction of viral RNA in local tissue at advanced stages of infection, the next logical question is whether a similar high-temperature induced recovery mechanism is active in systemic tissue. To address this question, *Arabidopsis* plants were infected with TuMV-GFP and maintained at 21ºC for 7 days before shifting them to either 15ºC, 21ºC, or 27ºC. The reason for incubating plants at 21ºC for 7 days was to allow the virus to move into systemic tissue before shifting to the different temperatures. This reduces the effects of differential unloading of the virus into the systemic tissue at different temperatures due to altered phloem transport rates. By performing the temperature shift after the virus had been extensively unloaded into systemic tissue, any observed differences in viral RNA at later time points are likely to be mainly due to the effects of temperature on the synthesis-degradation equilibrium, though it is impossible to completely exclude addition effects of phloem transport and cell to cell movement. By 21dpi/14dps plants maintained at both 15ºC and 21ºC exhibited severe stunting and leaf curling symptoms associated with TuMV infection. Additionally, the GFP produced by TuMV-GFP showed a very strong signal at these lower growth temperatures. In contrast, plants maintained at 27ºC exhibited much reduced viral symptoms and much of the systemic tissue did not show any GFP fluorescence, indicating that TuMV-GFP had been excluded from the systemic tissue (Figure 3.7A, inset). Quantification of the viral RNA by qRT-PCR revealed no differences between the plants maintained at 15ºC and 21ºC, but a more than 2-fold decrease in viral RNA at 27ºC compared to both lower temperatures (Figure 3.7B). Once again, vsiRNA abundance was found to be highest at 27ºC supporting the notion that viral PTGS is more active at this higher temperature resulting in recovery from the virus infection (Figure 3.7C).

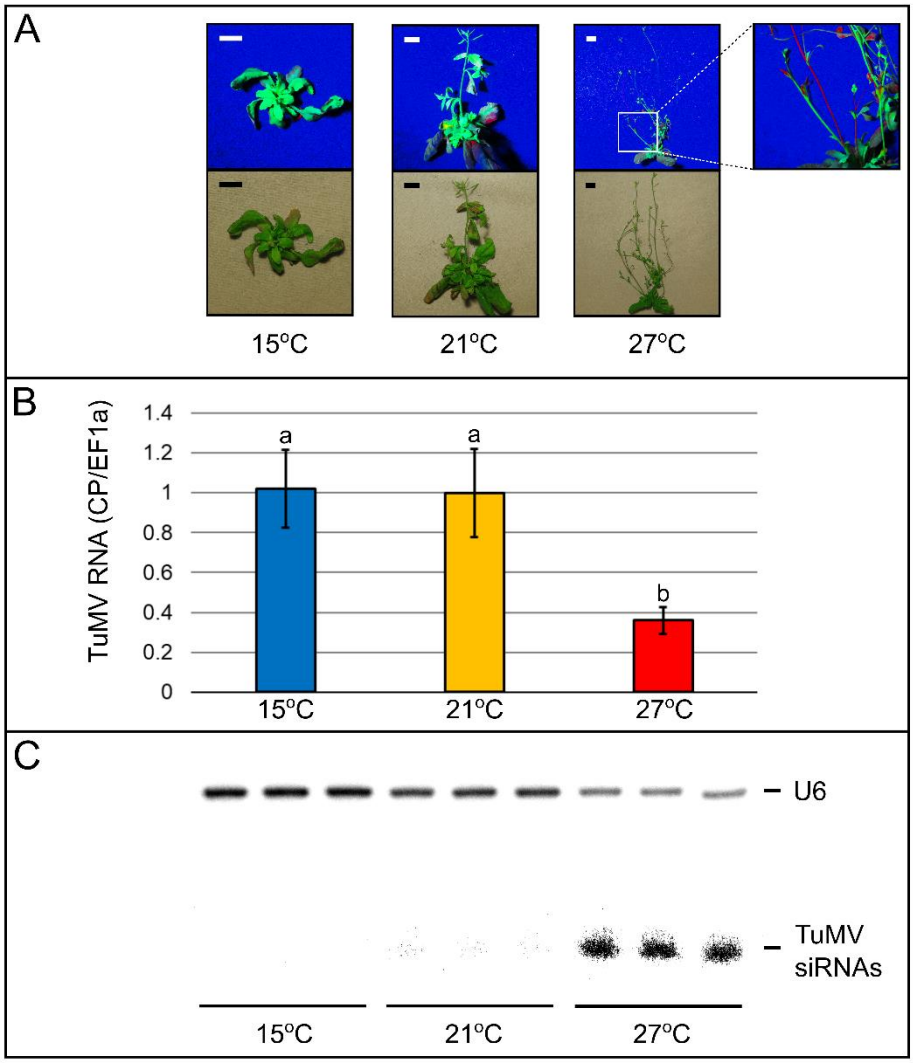

Figure 3.7 Relative abundance of TuMV-GFP RNA and vsiRNA abundance in systemic tissue of rub-inoculated *Arabidopsis* plants shifted to 15°C, 21°C, and 27°C. Samples were collected at 21dpi/14dps. Plants were maintained at 21°C until 7dpi/0dps and grown under 100µE white light throughout. **A** UV imaging of TuMV-GFP in systemically infected tissue. Scale bars indicate 1cm. An inset is included for the 27°C image to show the loss of GFP signal (red sectors) indicating recovery from TuMV-GFP infection. **B** qRT-PCR to measure the TuMV-GFP RNA. The total above-soil plant tissue from one plant was sampled and 3 independent plants were used as biological replicates. Hence bars represent the arithmetic mean of 3 samples ± SEM. Data are plotted relative to the mean of the 21°C samples. Tukeys HSD pairwise comparison of means revealed statistically significant groupings, as indicated by letters above each bar. **C** Northern blot of TuMV-derived siRNAs using the same samples as in **A**/**B**. A U6-transcript probe was used as a loading control.

3.6 Genetic factors responsible for activation of viral PTGS at high temperatures

It has previously been shown that DCL4, and DCL2 are the most important DCL proteins responsible for generating primary TuMV-vsiRNAs in *Arabidopsis* (Garcia-Ruiz *et al*., 2010), while DCL3 constitutes the main anti-TuMV DICER in the *dcl2-1dcl4-2* double mutant. Hence, the production of TuMV-vsiRNAs is severely hampered in the *dcl2-1dcl3-1dcl4-2* triple mutant. However, small-RNA sequencing revealed that TuMV-vsiRNAs can accumulate to low levels in *dcl2-1dcl3-1dcl4-2* mutants, likely due to the action of DCL1 (Garcia-Ruiz *et al*., 2010). Production of a stable mutant line which is completely DCL-defective is not possible as the hypomorphic *dcl1-9* mutant allele results in infertility when in a homozygous state. Taken together, the *dcl2-1dcl3-1dcl4-2* is the most feasible genetic background to achieve a maximal loss of anti-TuMV DCL function. A similar functional hierarchy exists for the RDR proteins responsible for producing secondary TuMV-vsiRNAs whereby RDR1 is most critical, and RDR6 is additionally required for full anti-TuMV silencing (Garcia-Ruiz *et al*., 2010). To investigate the molecular components of viral PGTS pathways which may be responsible for temperature-activated recovery from TuMV infection, analogous temperature shift experiments to those previously described were conducted using wild type (Col-0) *Arabidopsis* alongside mutants deficient in various components of TuMV-vsiRNA production. As previous results indicated that the shift in equilibrium from net viral synthesis to net degradation occurred between 21^{o}C and 27^{o}C, only these temperatures were used. At 28dpi/21dps a mild recovery phenotype (manifested the absence of GFP in the peduncle, sepals and pistils of inflorescences) was visible in the Col-0 genotype at 27^{o}C but not at 21^{o}C (Figure 3.8). Interestingly, no recovery phenotype was observed in the *dcl2-1dcl3-1dcl4-2*, *rdr1-1rdr6-15*, or *rdr1-1* genetic backgrounds at either 21^{o}C or 27^{o}C and only a very mild recovery phenotype was visible at 27^{o}C in the *rdr6-15* mutant (Figure 3.8). These results are suggestive that DCL2, DCL3 or DCL4 and RDR1 may be required for the temperature-induced recovery from TuMV-GFP infection. However, visual monitoring of GFP alone is not ideal for assessing the recovery from TuMV-GFP as the GFP protein may persist for some time after the TuMV-GFP RNA has been degraded.

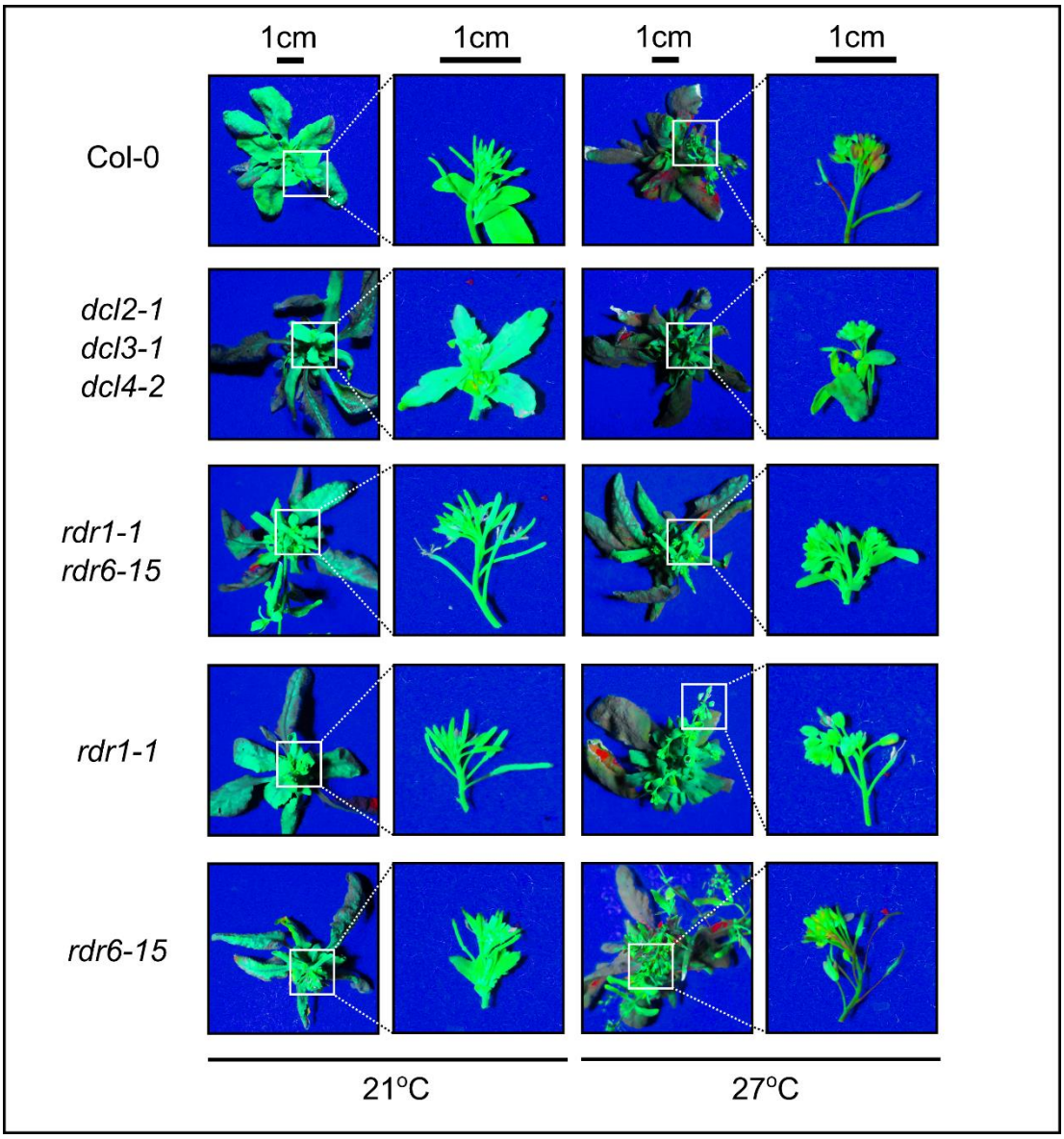

Figure 3.8 Visual assessment of TuMV-GFP, by UV imaging, in systemic tissue in different genetic backgrounds maintained under 21°C and 27°C white light. Photographs were taken at 28dpi/21dps. Plants were maintained at 21°C and 100µE white light until 7dpi/0dps, then shifted to either 21°C or 27°C under 200µE white light. Insets show a representative inflorescence for each genotype under the different growth conditions. Red patches are indicative of a TuMV-GFP recovery response.

To more accurately assess the abundance of viral RNA, qRT-PCR was performed for samples collected at 28dpi/21dps at both 21ºC and 27ºC, for the different genetic backgrounds. In agreement with previous results (Figure 3.7B), wild type (Col-0) plants shifted to 27ºC exhibited a roughly 2-fold reduction in TuMV-GFP RNA compared to plants maintained at 21ºC (Figure 3.9A). This ~2-fold reduction in TuMV-GFP at 27ºC was also observed in the *dcl2-1dcl3-1dcl4-2*, *rdr1-1*, and *rdr6-15* mutants (Figure 3.9 B,D,E). This indicates that the DCL2, DCL3, and DCL4 proteins are not strictly necessary for the recovery from TuMV-GFP at 27ºC. Similarly, these data indicate that neither RDR1 nor RDR6 is solely responsible for the temperature-induced recovery. Interestingly however, the temperature-induced recovery from TuMV-GFP infection was lost in the *rdr1-1rdr6-15* double mutant background, as the TuMV-GFP RNA was 4-fold higher at 27ºC than at 21ºC in this genotype (Figure 3.9C). These data suggest that RDR1 and RDR6 together are responsible for the temperature-induced recovery. Furthermore, the fact that magnitude of the reduction in TuMV-GFP RNA at 27ºC relative to 21ºC was almost identical in *rdr1-1*, *rdr6-15* and Col-0 (Figure 3.9A,D,E) suggests full functional redundancy between RDR1 and RDR6, with respect to the temperature-stimulated defence response. The increase in TuMV-GFP RNA at 27ºC compared to 21ºC in *rdr1-1rdr6-15* confirms the notion that TuMV synthesis is stimulated by raised ambient temperatures, and implies that RDR1 and RDR6 together comprise a necessary part of the temperature-stimulated defence which can result in a shift of the viral synthesis-degradation equilibrium to net degradation at 27ºC. These results are concordant with previous reports which identified RDR1(Ma *et al*., 2016) and RDR6 (Qu *et al*., 2005) as necessary components of temperature-stimulated defence against several viruses in *N. glutinosa* and *N. benthamiana*, respectively. The results presented herein imply that the roles of RDR1 and RDR6 in viral thermotherapy are likely to be involved in a mechanism that is broadly conserved across different plant families and which is functional against a variety of virus species.

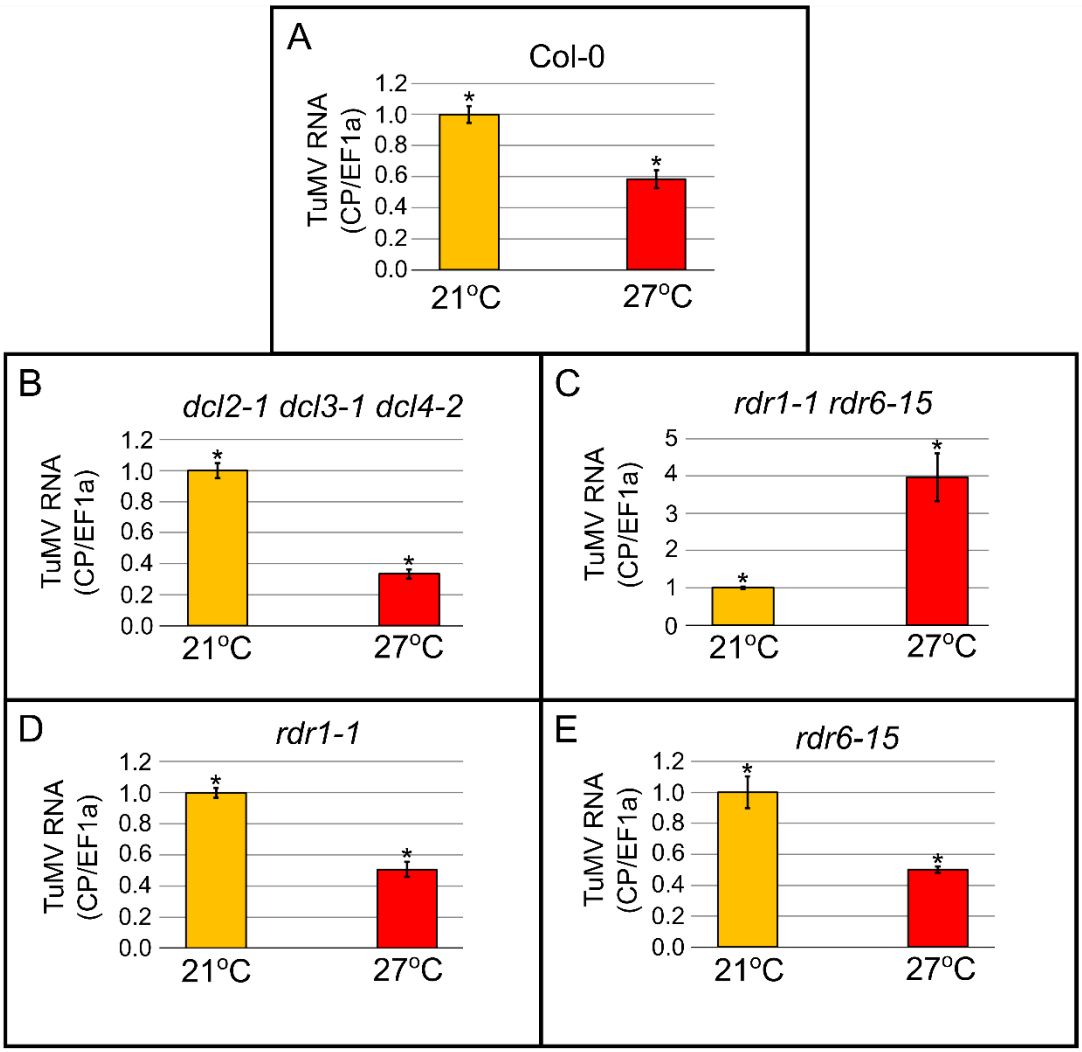

Figure 3.9 Quantification of TuMV-GFP RNA by qRT-PCR for plants shifted to 21ºC or 27ºC. Samples were collected at 28dpi/21dps. Plants were maintained at 21ºC and 100µE white light until 7dpi/0dps, then shifted to either 21ºC or 27ºC under 200µE white light. The *EF1a* 'housekeeping gene' was used for ΔΔCp calculations, and values are plotted relative to the mean TuMV signal at 21ºC. Error bars represent the SEM of 3 biological replicates. '*' indicates statistically significant difference in the means, assessed by a Student's T-test, $p<0.05$.

3.7 In *vivo* monitoring of early TuMV-GFP infection at different light intensities in local leaves of *Nicotiana benthamiana*

While the mechanisms by which temperature can influence recovery from infection has been reported for various viruses, details for how different light intensities can affect the equilibrium between viral synthesis and degradation remain scarce. A similar experimental design to that presented in Figure 3.1 was used to monitor the effects of exposing plants to different light intensities after inoculation with TuMV-GFP in *N. benthamiana* leaves. Plants were grown at 21^{o}C under 50µE (µE = µMolesm^{-2}s^{-1}) light. They were kept in these conditions for one day after inoculation with TuMV-GFP and then shifted to either 10µE, 50µE or, 100µE. A 21^{o}C growth temperature was maintained for each of these conditions. Infection foci became clearly visible under all light intensities by 2dpi/1dps, though the number of foci was much greater at 50µE and 100µE compared to 10µE. Moreover, the rate of foci development over time appeared highest at 100µE, followed by 50µE, and least for 10µE (Figure 3.10). The image processing software previously described was used to quantify these qualitative data. Consistently, TuMV-GFP infection was more advanced at higher light intensities when infection rate was scored by GFP content of leaves (Figure 3.11A), foci density (Figure 3.11B), or qRT-PCR (Figure 3.11C). These novel findings suggest that under these growth conditions in inoculated *N. benthamiana* leaves, the rate of TuMV infection is accelerated by high light intensities, or impeded at a very low light intensity. Though less pronounced than the HR response seen at 27^{o}C, necrotic lesions were visible to a greater extent at 10µE compared to higher light intensities, indicating that lower viral RNA at this low light intensity could be at least partially explained by an activation of HR.

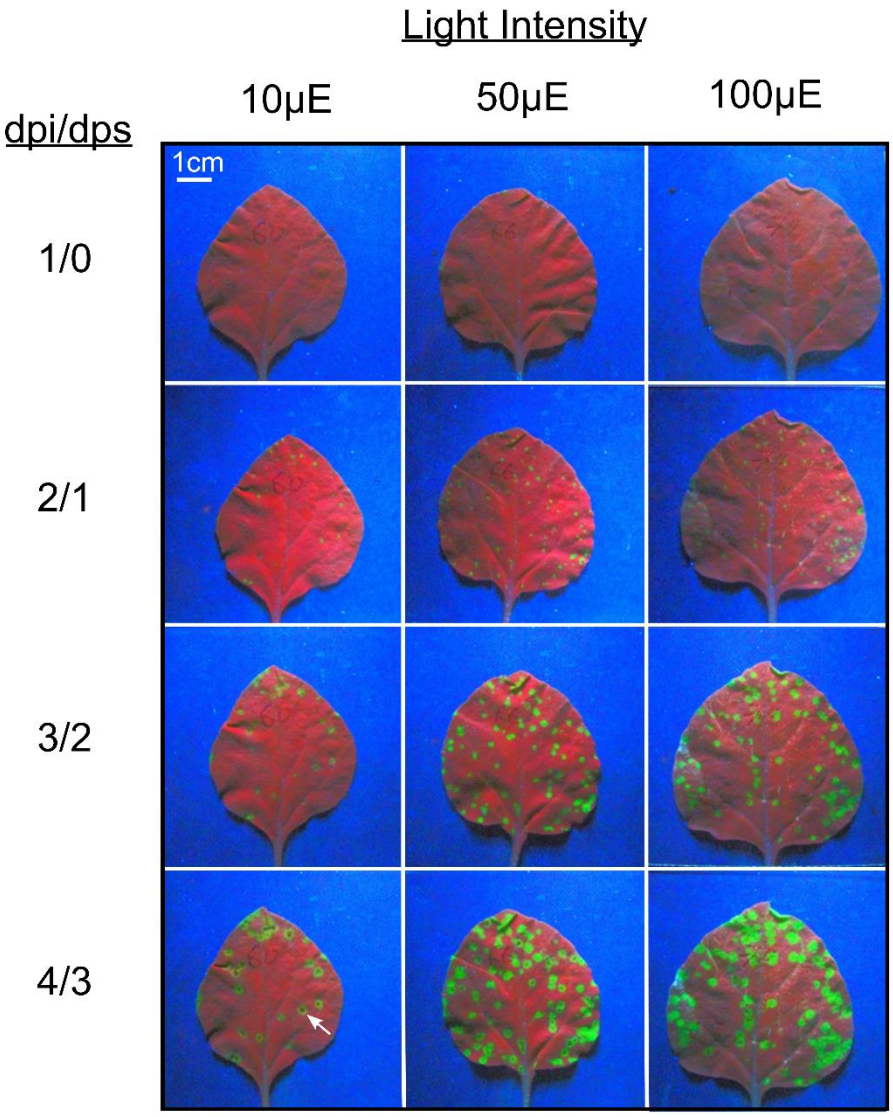

Figure 3.10 Time course UV imaging of *N. benthamiana* leaves rub-inoculated with TuMV-GFP infectious sap. Plants were shifted to 10µE, 50µE, and 100µE at 1 day post inoculation (dpi). Hence days post shift (dps) are one less than dpi for each time interval. Plants were maintained under 50µE white light until 1dpi/0dps and 21°C throughout. White arrow indicates a necrotic lesion.

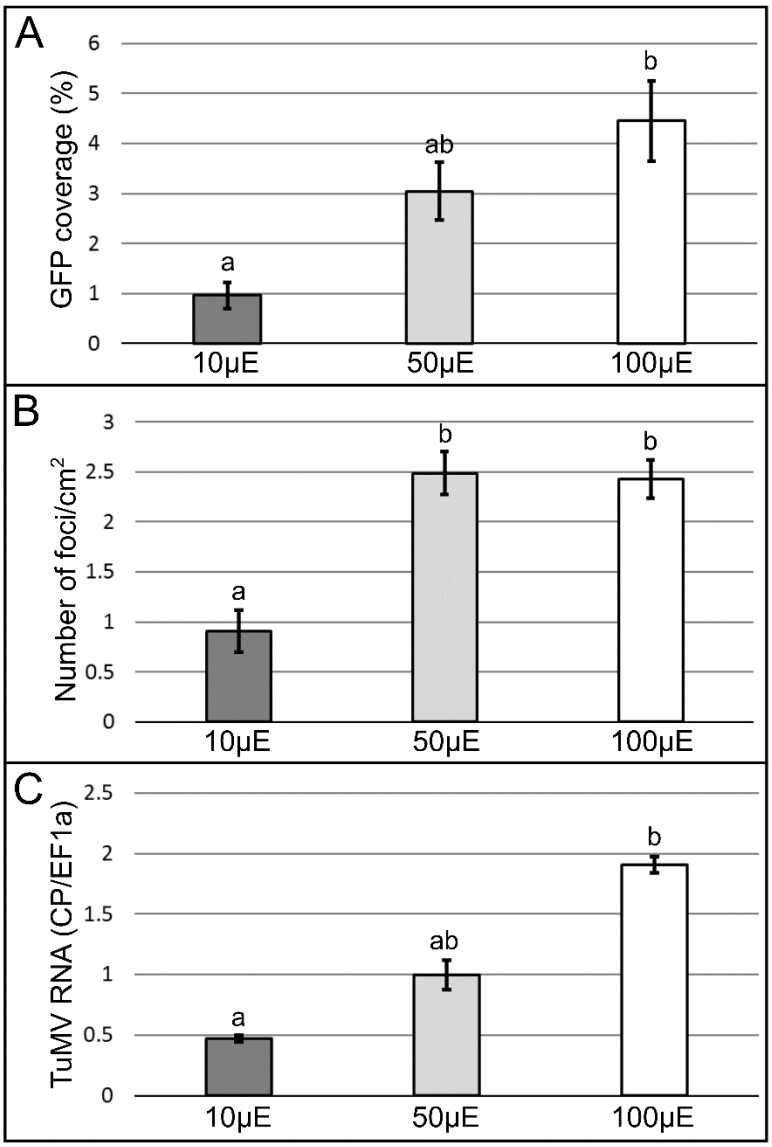

Figure 3.11 Quantification of TuMV-GFP infection for plants maintained under 10µE, 50µE, and 100µE white light. Plants were maintained under 50µE white light until 1dpi/0dps and 21°C throughout. The Matlab image processing script was used to measure: **A** total TuMV-GFP coverage, and **B** number of infection foci per cm^2 leaf area. Bars represent the arithmetic mean of 6 independently infected leaves at 3dpi/2dps. **C** qRT-PCR to measure the viral RNA of the same leaves shown in **A**/**B**. Samples were collected at 5dpi/4dps and 3 leaves were pooled into one sample. Hence, bars show the arithmetic mean of 2 samples. The *EF1a* 'housekeeping gene' was used for ΔΔCp calculations, and values are plotted relative to the mean TuMV signal for samples at 50µE. Error bars on all charts show the SEM. Tukeys HSD pairwise comparison of means revealed statistically significant groupings, as indicated by letters above each bar.

3.8 Analysis of early TuMV-GFP infection at different temperatures in local leaves of *Arabidopsis thaliana*

To avoid complications of an active HR, and to make use of RNA silencing mutants, subsequent experiments to monitor the effects of light intensity on TuMV-GFP infection were conducted using *Arabidopsis* as a susceptible host plant. Plants were grown at 21^{o}C and 100µE light intensity for four weeks, then infected with TuMV-GFP and kept under these conditions for four days before shifting them to 50µE, 100µE or 200µE. Because previous results suggested that TuMV induced siRNAs are only detectable at 27^{o}C in local tissue (Figure 3.4C), the growth temperature was increased from 21^{o}C to 27^{o}C at the same time as shifting to the different light intensities. Consistent with the observations made for local infection on *N. benthamiana*, the abundance of TuMV-GFP was lower at 50µE than the higher light intensities, scored qualitatively by GFP fluorescence (Figure 3.12A) or quantitatively by qRT-PCR (Figure 3.12B) at 7dpi/3dps. Interestingly, the accumulation of vsiRNAs was much higher at 50µE than for either of the higher light intensities (Figure 3.12C). This suggests that in addition to high temperature, low light intensities can promote antiviral RNA silencing, leading to a reduction in viral RNA.

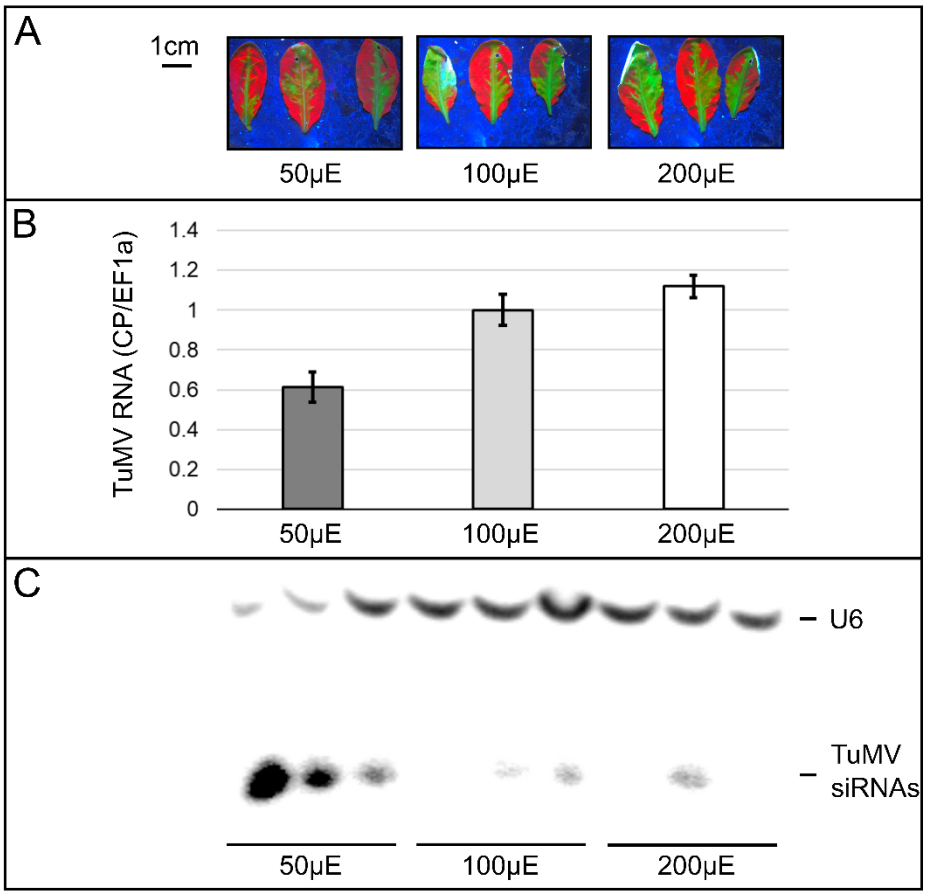

Figure 3.12 Relative abundance of TuMV-GFP RNA and vsiRNA abundance in local tissue of rub-inoculated *Arabidopsis* plants maintained maintained under 50µE, 100µE, and 200µE white light. Samples were collected at 7dpi/3dps. Plants were maintained at 100µE and 21ºC until 4dpi/0dps, and shifted to 27ºC from 0dps onwards. **A** UV imaging of TuMV-GFP in locally infected leaves. **B** qRT-PCR to measure the TuMV-GFP RNA. 3 infected leaves from the same plant were pooled into one sample and 3 samples from 3 independent plants were used as biological replicates. Hence bars represent the arithmetic mean of 3 samples ± SEM. Data are plotted relative to the mean of the 100µE samples. **C** Northern blot of TuMV-derived siRNAs using the same samples as in **A**/**B**. A U6-transcript probe was used as a loading control.

3.9 Analysis of TuMV-GFP infection and recovery at different temperatures in systemic tissue of *Arabidopsis thaliana*

As low light levels intensified the temperature-induced accumulation of siRNAs in local tissue, it was hypothesised that low light intensities may also facilitate recovery from systemic infection. To investigate this, plants were grown at 21^{o}C with 100µE light for four weeks before TuMV-GFP infection. They were kept under these conditions for seven days to allow systemic infection to develop, before shifting them to either 50µE, 100µE or 200µE. As the accumulation of siRNA in systemic tissue is much higher than for local tissue, plants were maintained at 21^{o}C throughout the experiment to determine whether low light intensity alone could result in recovery without shifting to 27^{o}C. At 28dpi/21dps, viral symptoms were most severe for plants kept at 200µE, less severe at 100µE, and almost undetectable at 50µE (Figure 3.13A). Monitoring the same plants under UV light revealed red patches, indicating a loss of GFP signal due to virus exclusion, in the systemic tissue of plants kept at 50µE and 100µE but not those kept at 200µE (Figure 3.13A). qRT-PCR revealed an approximately 3- to 6-fold reduction in viral RNA for plants kept at 50µE and 100µE compared those at 200µE (Figure 3.13B). Interestingly, the accumulation of vsiRNAs was much higher for plants kept at the two lower light intensities compared to those at 200µE (Figure 3.13C). These results suggest that shifting plants to very low light intensities may promote RNA silencing-mediated recovery from TuMV infection in an analogous manner to shifting to higher temperature.

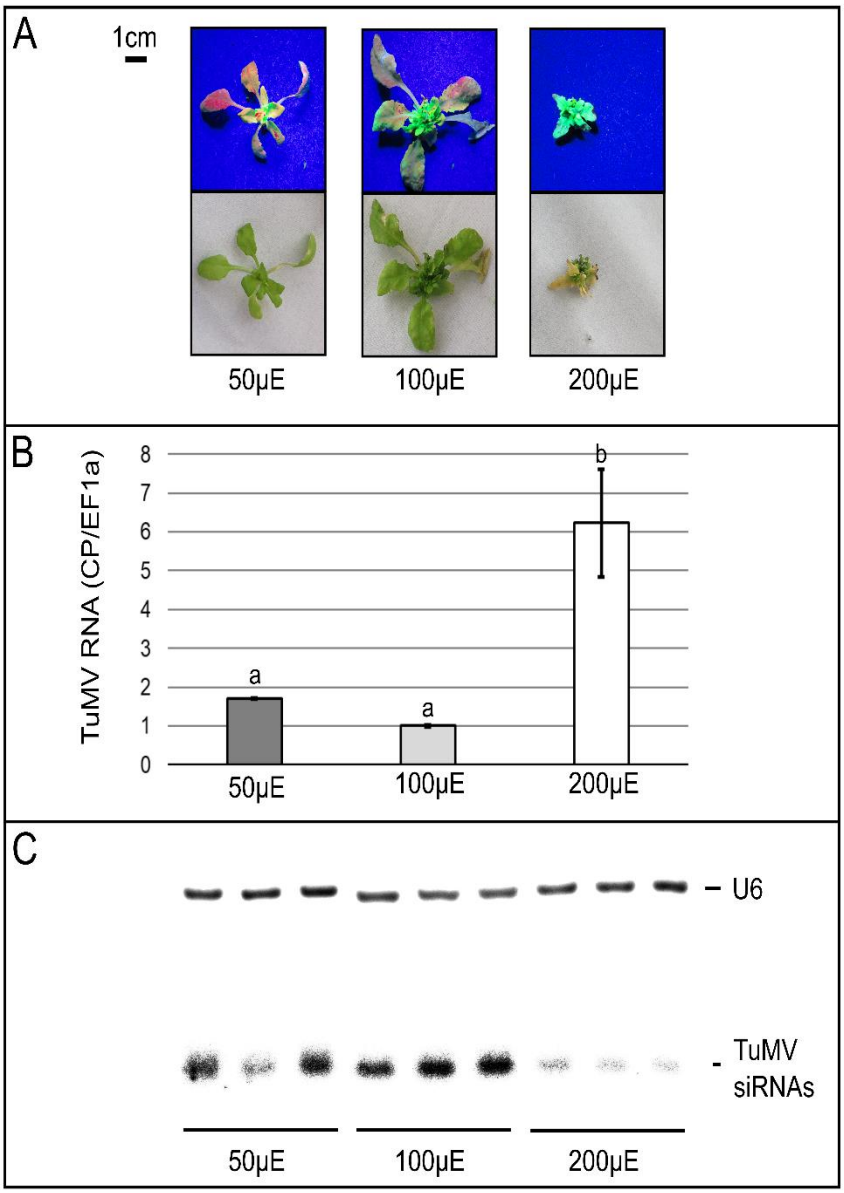

Figure 3.13 Relative abundance of TuMV-GFP RNA and vsiRNA abundance in systemic tissue of rub-inoculated *Arabidopsis* plants maintained under 50µE, 100µE, and 200µE white light. Samples were collected at 28dpi/21dps. Plants were maintained under 100µE white light until 7dpi/0dps, and 21ºC throughout. **A** UV imaging of TuMV-GFP in systemically infected tissue. Scale bar indicates 1cm. **B** qRT-PCR to measure the TuMV-GFP RNA. The total above-soil plant tissue from one plant was sampled and 3 independent plants were used as biological replicates. Hence bars represent the arithmetic mean of 3 samples ± SEM. Data are plotted relative to the mean of the 100µE samples. Tukeys HSD pairwise comparison of means revealed statistically significant groupings, as indicated by letters above each bar. **C** Northern blot of TuMV-derived siRNAs using the same samples as in **A**/**B**. A U6-transcript probe was used as a loading control.

3.10 Genetic factors responsible for activation of viral PTGS at low light intensity

To investigate the molecular components of viral PGTS pathways which may be responsible for enhanced recovery from TuMV infection at low light intensities, analogous light-shift experiments (at 21^{o}C) to those previously described were conducted using wild type (Col-0) *Arabidopsis* alongside mutants deficient in various components of TuMV-vsiRNA production. As previously observed, shifting Col-0 plants to 50µE resulted in TuMV-GFP recovery symptoms, manifested by the absence of GFP in stems and inflorescences. Such recovery phenotypes were not observed for plants maintained at 200µE (Figure 3.14). The TuMV-GFP recovery phenotype was also visible in the *dcl2-1dcl3-1dcl4-2* mutant genotype at 50µE but not at 200µE. In the *rdr1-1*, *rdr6-15*, and *rdr1-1rdr6-15* mutants, TuMV-GFP appeared to persist in the systemic tissue at both light intensities, suggesting that the low-light-mediated recovery from TuMV-GFP is diminished in these genotypes (Figure 3.14). To more accurately assess the accumulation of viral RNA, qRT-PCR was performed for the samples collected at 28dpi/21dps for the different genotypes maintained at 50µE and 200µE. In agreement with the recovery phenotypes observed under UV light, the Col-0 and *dcl2-1dcl3-1dcl4-2* mutant contained lower TuMV-GFP RNA at 50µE compared to 200µE (Figure 3.15A,B). There was no significant difference in TuMV-GFP RNA between the two light intensities for *rdr1-1* or *rdr6-15* (Figure 3.15D,E). This is consistent with the absence of recovery phenotypes at either light intensity for these genotypes (Figure 3.14). Interestingly, in the *rdr1-1rdr6-15* mutant the TuMV-GFP RNA was approximately 5-fold higher at 50µE compared to 200µE (Figure 3.15C). Taken together, these data suggest that RDR1 and RDR6 both contribute the enhanced resistance to TuMV-GFP at low light intensity. Furthermore, the trend of TuMV-GFP RNA accumulation at different light intensities for the *rdr1-1* and *rdr6-15* mutants is an intermediate of the trends observed in Col-0 and *rdr1-1rdr6-15*, implying that RDR1 and RDR6 function additively in this defence response.

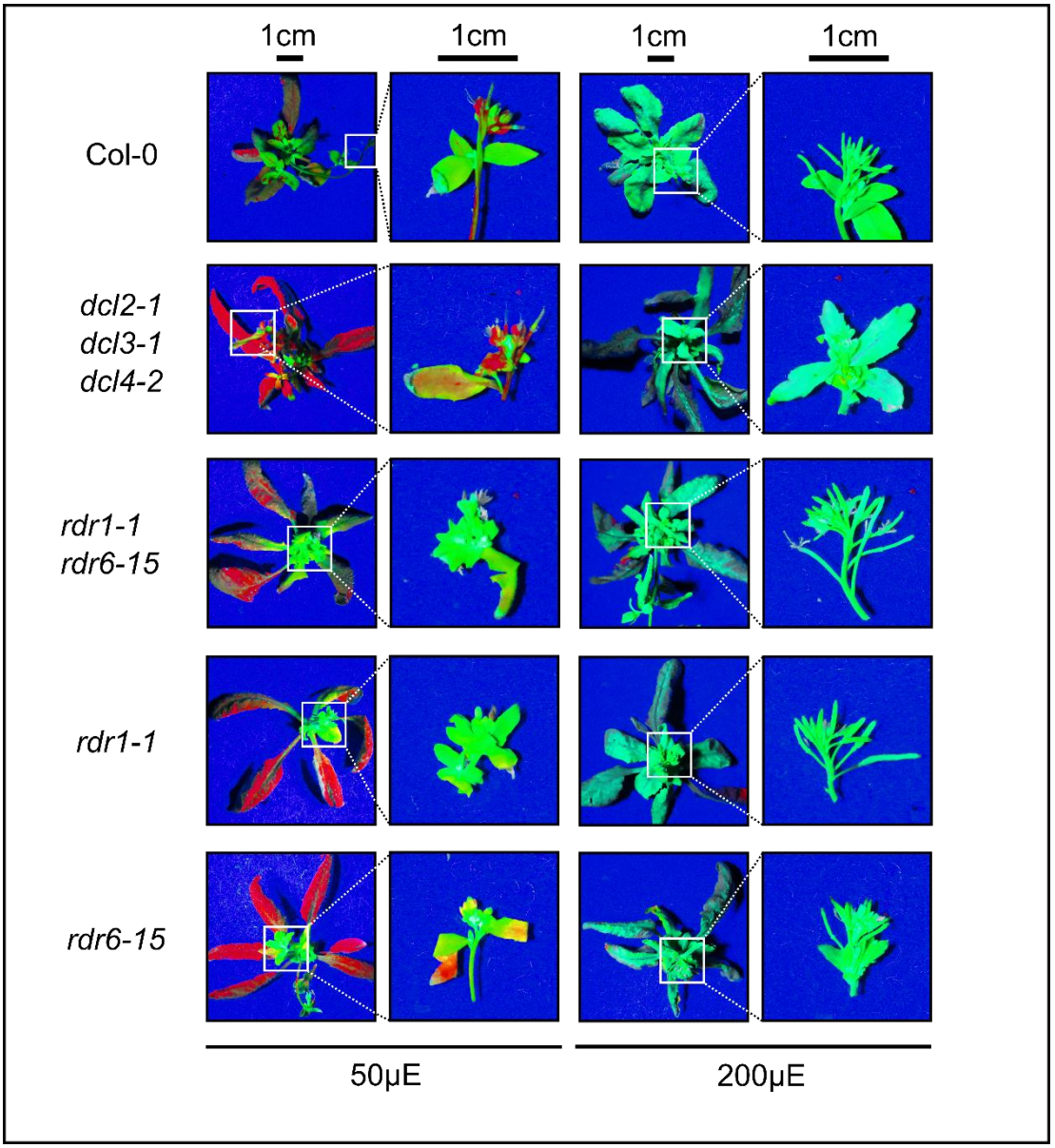

Figure 3.14 Visual assessment of TuMV-GFP, by UV imaging, in systemic tissue in different genetic backgrounds shifted to 50µE and 200µE white light. Photographs were taken at 28dpi/21dps. Plants were maintained under 100µE white light until 7dpi/0dps, and 21ºC throughout. Insets show a representative inflorescence for each genotype under the different growth conditions. Red patches are indicative of a TuMV-GFP recovery response.

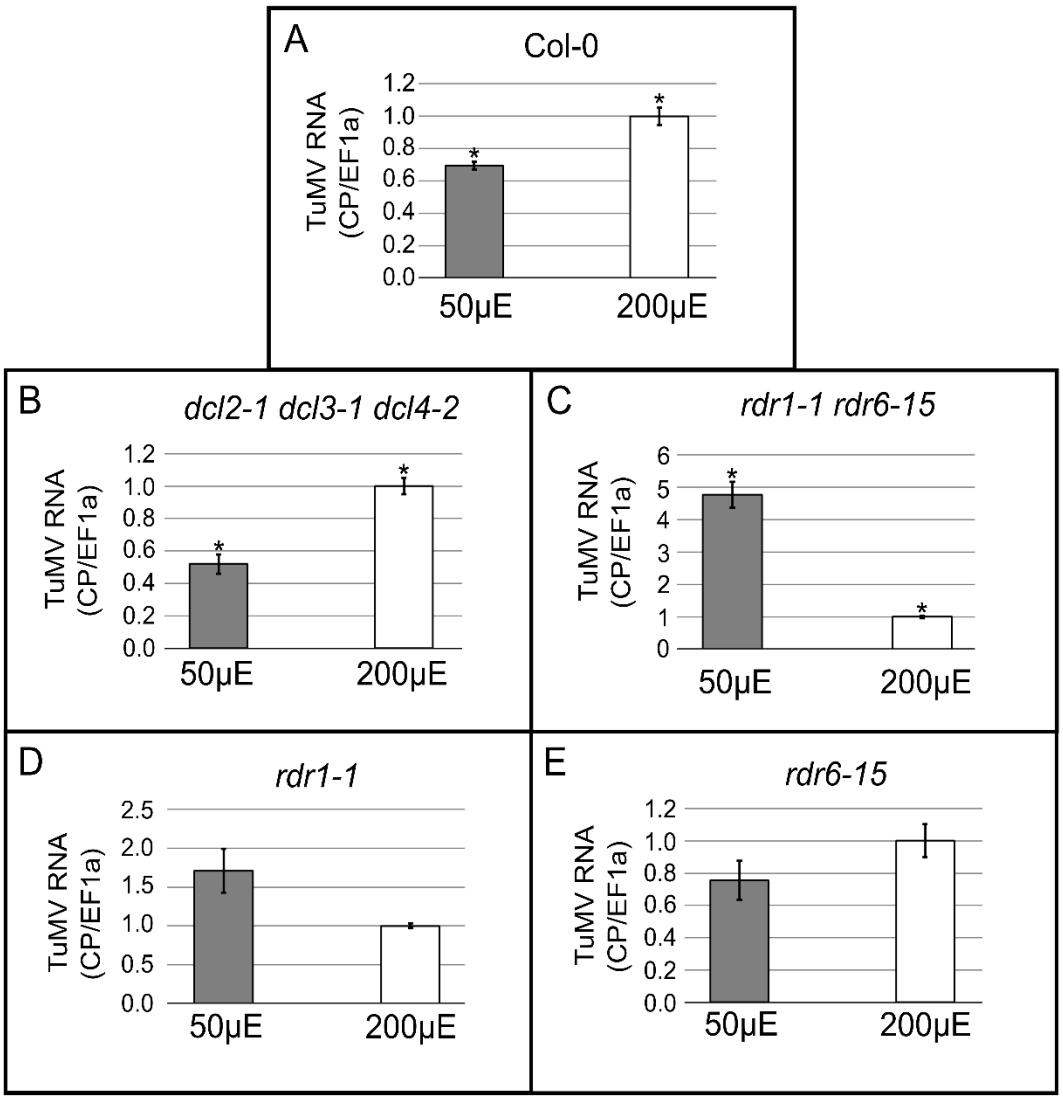

Figure 3.15 Quantification of TuMV-GFP RNA by qRT-PCR for plants shifted to 50µE or 200µE white light. Samples were collected at 28dpi/21dps. Plants were maintained under 100µE white light until 7dpi/0dps, and 21ºC throughout. The *EF1a* 'housekeeping gene' was used for ΔΔCp calculations, and values are plotted relative to the mean TuMV signal at 21ºC. Error bars represent the SEM of 3 biological replicates. '*' indicates statistically significant difference in the means, assessed by a Student's T-test, $p<0.05$.

3.11 Combined effects of temperature and light intensity shifts on the expression profiles of *RDR1* and *RDR6*

As RDR1 and RDR6 emerged as necessary components for both high temperature- and low light-induced recovery from TuMV-GFP infection, qRT-PCR was used to monitor the relative expression of *RDR1* and *RDR6* in wild type *Arabidopsis* plants infected with TuMV-GFP and shifted to 21^{o}C/27^{o}C and 50µE/200µE white light. The same samples taken at 28dpi/21dps, presented in Figure 3.9A and Figure 3.15A were used for this analysis. Additionally, 28dpi/21dps samples shifted to both 27^{o}C and 50µE were analysed to test the combinatorial effect of low light intensity and high temperature on *RDR1* and *RDR6* gene expression. 3-dimensional plots of the relative expression of *RDR1* and *RDR6* show that expression of both genes is highest in the 27^{o}C/50µE condition (Figure 3.16A,B). Moreover, 2-way ANOVAs indicated that light intensity significantly affects the expression of *RDR1* ($F_{1,8}$=20.33, p=0.002) and *RDR6* ($F_{1,8}$=23.14, p=0.001) at both temperatures. Conversely, the effect of temperature at both light intensities was only significant for *RDR1* ($F_{1,8}$=51.58, p=9.4e^{-5}). Furthermore, there was no significant interaction between the effects of light and temperature treatments on *RDR1* expression ($F_{1,8}$=1.30, p=0.29), indicating that the two environmental factors regulate transcription at this locus independently. Consequently, 27^{o}C and 50µE treatments have positive and additive effects on *RDR1* expression. In summary, shifting plants to 50µE results an increase of transcriptional activity at both *RDR1* and *RDR6*, irrespective of the growth temperature and shifting plants to 27^{o}C increases *RDR1* expression irrespective of the light intensity. The temperature induction of RDR6 expression is apparent only under 50µE but not 200µE light intensity. These data corroborate the phenotypic and genetic data suggesting that RDR1 and RDR6 are necessary components for heightened defence against TuMV-GFP resulting from raised growth temperature and lowered light intensity.

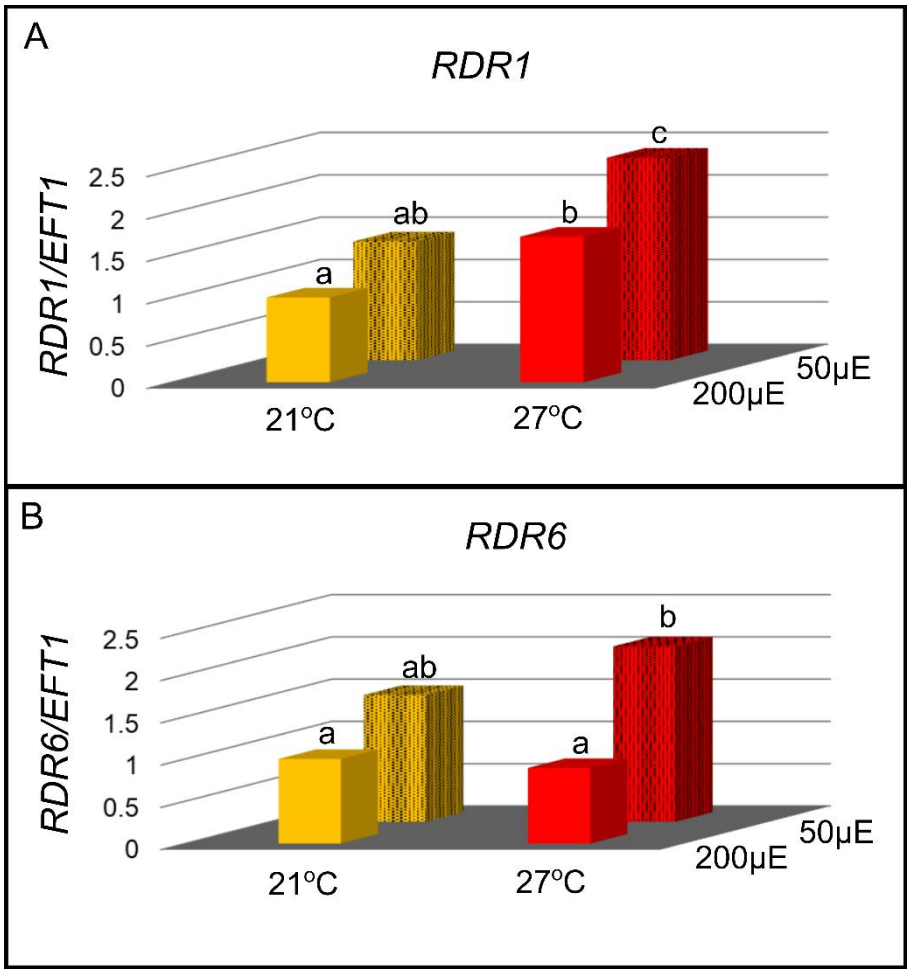

Figure 3.16 Quantification of *RDR1* and *RDR6* relative expression in TuMV-GFP infected Col-0 plants, by qRT-PCR. Plants were shifted to 50µE or 200µE white light, with a concomitant temperature shift to either 21°C or 27°C. Plants were maintained under 21°C/100µE white light until 7dpi/0dps. Samples were collected at 28dpi/21dps. The *EF1a* 'housekeeping gene' was used for ΔΔCp calculations, and values are plotted relative to the mean expression value at 21°C. Error bars represent the SEM of 3 biological replicates. Tukey's HSD pairwise comparison of means revealed statistically significant groupings, as indicated by letters above each bar.

3.12 Chapter discussion

3.12.1 Temperature increases are likely to affect multiple aspects of the potyvirus lifecycle and host defence responses.

Studies into the effects of temperature on potyvirus infections have resulted in some seemingly contradicting observations. For instance, disease incidence and severity of PVY infections in potato crops are often considered to have a positive correlation with modest temperature rises (Hasan & Rashid, 2015). This apparently conflicts with the widespread application of thermotherapy to treat potyvirus infections in a range of plant species (Table 3.1). These anomalies can be consolidated by taking a holistic, epidemiological approach to potyvirus infections. Firstly, vector dynamics are likely to be the main factor driving the positive correlation between disease incidence and mild temperature increases in fields (Broadbent & Hollings, 1951, Newman, 2004, Singh *et al*., 1988). Secondly, reports of positive correlations between temperature increases and potyvirus disease severity relate to observations of PTNRD on potato tubers (Kerlan *et al*., 2011, Le Romancer & Nedellec, 1997). It is likely that this phenomenon is mainly due to host necrotic defence responses becoming more active at higher temperatures, rather than viral titres reaching higher levels. This is supported by the observation of increased virus-induced tissue necrosis for TuMV-GFP infected *N. benthamiana* plants shifted to 27^{o}C compared to lower growth temperatures (Figure 3.1). However, a direct link between these observations remains speculative.

Even when confounding factors such as vector dynamics and necrotic responses are experimentally controlled, the effects of different temperature treatments on the potyvirus lifecycle and associated cellular immune responses appears to be complex. The concept of an equilibrium between virus synthesis (resulting from replication and translation of viral RNA) and degradation is useful to understand the apparently multi-faceted effects of temperature treatments on potyvirus titres. Time-course imaging of TuMV-GFP infection foci in locally inoculated *N. benthamiana* leaves indicate that TuMV-GFP synthesis is stimulated by warmer ambient temperatures (Figure 3.1). Furthermore, analysis of local TuMV-GFP infections at early time-points in *Arabidopsis* suggests that abundance of viral RNA increases with rising temperature despite antiviral PTGS becoming more active (Figure 3.4). This suggests that the rate

of synthesis exceeds the rate of degradation at warmer temperatures during early infection. However, by analysing later time-points of locally infected *Arabidopsis* leaves, it was shown that the enhanced antiviral PTGS at 27ºC can eventually result in the rate of degradation overcoming synthesis (Figure 3.5, Figure 3.6B). Furthermore, this shift towards net TuMV-GFP degradation at 27ºC also resulted in mild recovery from systemic infections, in an analogous manner to thermotherapy practices (Figure 3.7). These conclusions are concordant with similar experimental data relating to the effects of different ambient temperatures on the synthesis/degradation of TuMV and PVY (Chung *et al*., 2015, Chung *et al*., 2016, Del Toro *et al*., 2015, Ma *et al*., 2016).

The determinants for the initial promotion of TuMV synthesis at warmer temperatures remain to be elucidated. A simplistic scenario could involve enzymes involved in replication and translation of the TuMV genome becoming more active, or expressed to higher levels, at higher temperatures. An interesting candidate to test is HSP70, which plays an integral role in potyvirus replication/translation (Hafren *et al*., 2010) and has also been shown have a proportionate, positive correlation between its transcriptional activity and increasing temperature, over a 12ºC to 27ºC range (Kumar & Wigge, 2010). The putative role of HSP70 as a determinant for increased TuMV synthesis at warm ambient temperature could be evaluated using transgenic plants that express HSP70 under non-temperature-responsive promoters, or *hsp70* null-mutant plants.

The determinants of the enhanced PTGS-induced degradation of TuMV at 27ºC were characterised in this study. A loss of temperature-induced recovery from TuMV-GFP infection was observed the *rdr1-1rdr6-15* double mutant (Figure 3.9). However, both the *rdr1-1* and *rdr6-15* mutants displayed a temperature-enhanced recovery that was indistinguishable from wild type plants, suggesting full functional redundancy between these loci, with respect to the temperature-induced recovery phenotype. Moreover, the TuMV-GFP RNA in the *rdr1-1rdr6-15* double mutant was much higher at 27ºC compared to 21ºC, likely reflecting temperature-enhancement of viral synthesis. These data imply that the increased synthesis of TuMV is curtailed by an enhanced antiviral PTGS at 27ºC, mediated redundantly by RDR1 and RDR6 proteins. Moreover, the transcriptional activities of both *RDR1* and *RDR6* are enhanced by a temperature shift from 21ºC to 27ºC indicating that the temperature-dependence of

TuMV-vsiRNA production is at least partly explained by transcriptional regulation at these loci. A similar model was previously suggested to explain the enhanced resistance to several viruses in *Nicotiana* species (Ma *et al*., 2016, Qu *et al*., 2005), suggesting that thermal induction of *RDR1* and/or *RDR6* may constitute a conserved mechanism for enhanced virus resistance at warm ambient temperatures.

While the genetic data support the conclusion that RDR1 and RDR6 are involved in a temperature-stimulated RNA silencing defence response, it is somewhat unexpected to find that the *dcl2dcl3dcl4* mutant was indistinguishable from the wild type genotype for this response. This is counter-intuitive due to the fact that DCL2, DCL3, and DCL4 act upstream of the RDRs in the vsiRNA-generating pathway. This highlights a critical limitation of mutational genetic analysis, namely redundancy. It is likely that the functional DCL1 in the *dcl2dcl3dcl4* genetic background could generate a basal layer of TuMV-vsiRNAs (as previously shown (Garcia-Ruiz *et al.*, 2010)) which would provide substrate for temperature-stimulated vsiRNA amplification by RDR1 and RDR6. Therefore, a quadruple *dcl1dcl2dcl3dcl4* mutant would be required to assess the role of DCLs in this response. As complete loss-of-function *dcl1* mutations are difficult to obtain due to severe developmental phenotypes, experiments involving a quadruple *dcl1dcl2dcl3dcl4* mutant would be complex and were beyond the scope of this study.

3.12.2 Low light intensity appears to reduce TuMV synthesis and enhance anti-TuMV PTGS by induction of RDR1 and RDR6

The effects of different light environments on virus infections have been poorly explored and, arguably, remain scarcely characterised. In this study, monitoring of TuMV-GFP infection foci in locally infected *N. benthamiana* leaves at early time-points revealed a nearly linear, positive correlation between the TuMV-GFP titre and increasing light intensity (Figure 3.10, Figure 3.11). This is conceptually intuitive as cellular resources required for viral synthesis are likely to be limiting under such extremely low light intensity and, consequently, could limit viral replication and/or translation. Furthermore, there have been numerous reports to suggest that SA-mediated pathogen defence responses are light dependent (Chandra-Shekara *et al.*, 2006, Genoud *et al.*, 2002, Griebel & Zeier, 2008, Zeier *et al.*, 2004). Therefore, under

conditions of high light intensity, SA-mediated defence pathways may operate to curtail the presumed increased rate of virus synthesis.

By analysing the production of TuMV-vsiRNAs under different light intensities in locally infected *Arabidopsis* leaves, a curious observation was made suggesting that antiviral PTGS may become more active at low light intensity (Figure 3.12). This increased production of TuMV-vsiRNA at low light intensity was associated with a reduction in TuMV-GFP RNA in both locally and systemically infected tissue (Figure 3.12, Figure 3.13), indicating that the enhanced anti-TuMV PTGS could induce recovery from TuMV infection. Interestingly, this low-light-induced recovery was lost in in the *rdr1-1* and *rdr6-15* mutants, implicating these loci in the response (Figure 3.15). Moreover, the TuMV-GFP RNA was higher at 50µE compared to 200µE in the *rdr1-1rdr6-15* genotype, indicating a more extreme loss of silencing function in the double mutant compared to either single mutant. This suggests that RDR1 and RDR6 function additively in the hyper-production of TuMV-vsiRNAs at low light intensity, which is further corroborated by transcriptional induction of both RDR1 and RDR6 at 50µE compared to 200µE. While the increased production of vsiRNAs and the loss of the recovery response in the *rdr1rdr6* mutant at low light intensities argue for a role of antiviral RNA silencing in this response, it is conceivable that some of the reduction of TuMV-GFP RNA could result from a general reduction in cellular activity (relating to transcription and translation, for example). The observation that the signal for the U6 transcript, used as a reference probe for Northern blotting, was consistently weaker for low-light intensity samples suggests that such a general reduction in cellular activities may be an additional cause for the reduction of TuMV-GFP RNA at extremely low light intensities.

3.12.3 Potential for antiviral 'light therapy' as an alternative or complementary approach to thermotherapy

It is likely that the low-light-intensity-induced recovery is a manifestation of both reduced viral synthesis and increased vsiRNA-mediated degradation resulting in an equilibrium shift towards net reduction in virus RNA. This observation has potential application for improved phyto-sanitation techniques, by optimising light conditions to improve the efficiency of thermotherapy practices. Thermotherapy is often

conducted using explants which are grown on media containing growth nutrients and exogenous carbon sources such as sucrose. Therefore, the exogenous carbon source could allow explants to be cultured under extremely low light intensities or even complete darkness to improve the induction of antiviral PTGS. However, this relies on the assumption that these tissue-culturing conditions would not disrupt the low-light-intensity-induction of antiviral-PTGS. Hence the possibility of 'skoto-therapy' to eliminate viruses from germplasm is an interesting avenue which could emerge from the results presented herein, but requires further research to validate its translation to different virus/host combinations and to different growth systems.

3.12.4 Different antiviral defence responses may be alternately activated/repressed in different environmental conditions

It is surprising that a higher RNA of TuMV-GFP was observed at the lower light intensity in the *rdr1-1rdr6-15* mutant, as previous results suggested that TuMV synthesis is promoted by high light intensity (Figure 3.10, Figure 3.11). A possible explanation for this is that long-term exposure to such low light intensity could severely hamper SA-mediated defence pathways. Thus, a model could be proposed whereby SA-mediated defence forms the predominant antiviral response under conditions of high light intensity, and vsiRNA-mediated defence constitutes the major antiviral response in low light intensity environments. Hence, the loss of the low-light-intensity-induced PTGS in the *rdr1-1rdr6-15* mutant, together with the low activity of SA-mediated defence, would render plants extremely susceptible to virus infection. The notion that plants may alternate, or even mutually repress, different defence pathways is not unprecedented. A well characterised example is the mutual antagonism between defence pathways regulated by SA and jasmoic acid (JA). A somewhat simplified explanation for this phenomenon posits that it allows plants to mount a defence response that is appropriate to the type of invading pathogen, as SA-pathways mainly defend against biotrophic pathogens, whereas JA-pathways are involved in necrotroph and herbivore defence (Rojo *et al.*, 2003, Thomma *et al.*, 2001). Such a simplistic model of mutual antagonism between SA- and vsiRNA- based virus resistance is unlikely, particularly as there are examples of synergism between the two, such as the SA-mediated induction of *RDR1* expression (Xie *et al.*, 2001, Yu *et al.*, 2003). In the context of results presented herein, it seems more likely that SA- and

vsiRNA-defence pathways may exhibit alternating peak activities that are conditional on external environmental signals. One possible advantage to this could be alleviation of the so-called growth vs. defence paradigm (Fan *et al.*, 2014, Robert-Seilaniantz *et al.*, 2011), which suggests that many defence pathways are restrictive to plant growth, possibly to allow cellular resources to be invested primarily to immune responses. It is conceivable that SA-based defence pathways may be more antagonistic to growth than those involving vsiRNAs, which are more precisely directed to the invading pathogen. Hence, under severely growth-limiting conditions, such as low light intensity, further restriction to growth by activation of SA-defence could be detrimental to the plants. Activation of vsiRNA-based defence could provide a vital, alternative defence pathway when cellular resources for growth are limiting. This model could be further tested by assessing the relative contribution of vsiRNA-based defence under different growth-limiting conditions. Another reason for the apparent alternation between SA- and vsiRNA-based immunity could be to allow temporal separation of the two defence pathways to allow for a continuous defence response, for instance across day-night cycles. This hypothesis is explored further in the subsequent chapter.

3.12.5 Biological relevance of the light and temperature shift experiments

It is noteworthy that the experimental design of shifting virus infected plants to different light intensity and temperature environments represents an artificial scenario which would not be experienced by plants in nature. However, the rationale for this experimental design was firstly to assess how these environmental signals directly affect the virus synthesis/degradation equilibrium, by limiting as far as possible the effects of these altering conditions on plant development. Furthermore, the experimental design can be viewed as a model for the artificial practice of thermotherapy, therefore providing potential commercial relevance to the results obtained. Nevertheless, it is worthwhile to consider the biological phenomena which may underlie the observed effects of altered environmental conditions on antiviral defences. Mild temperature increases are known to positively affect virus transmission (Broadbent & Hollings, 1951, Newman, 2004, Singh *et al.*, 1988), phloem transport rate (Wardlaw & Bagnall, 1981), and (for potyviruses at least) the rate of virus synthesis. Hence increased production of vsiRNAs as temperatures increase may have

evolved as an important defence strategy to curtail virus infections in these otherwise virus-favourable conditions. Interestingly, this phenomenon appears to not be specific to plants, as thermally-enhanced siRNA-mediated resistance to viruses has also been reported in mosquitos (Adelman *et al.*, 2013) and a recent report has linked a dsRNA-dependent pathway to warm-temperature-induced recovery to rhinovirus in human cells (Foxman *et al.*, 2016).

While it may be tempting to draw parallels between the observed effects of different light intensities on virus defence and natural conditions of shade, such comparisons should be treated with caution. This is mainly because natural shading from neighbouring plants involves a reduction in light intensity but also a spectral shift in the light quality. The experimental conditions used herein resulted in a reduction in light intensity without altering light quality and hence do not constitute a valid model for natural shade conditions. It would be interesting to test whether the observed effects of low light intensity on antiviral PTGS could have evolved as an adaptation for improved virus resistance under natural shade conditions. However, technical limitations did not allow this hypothesis to be pursued in this study. Day/night cycles represent a different natural scenario that could represent an evolutionary pressure for adapting biological responses to different light intensities. Hence, the increased vsiRNA production under the artificially-induced low light intensity conditions could reflect differences in vsiRNA production that are naturally influenced by the changing light intensity under day/night cycles. The subsequent chapter is dedicated to investigating hypothesis.

Chapter 4: The impact of inoculation time on potyvirus susceptibility

4.1 Chapter Introduction

4.1.1 General features of circadian rhythms

The daily rotation of the earth about its axis subjects organisms to dramatically different environments associated with day and night. Consequently, most organisms display differences in behaviour, physiology, biochemistry, and gene expression across day/night cycles. Collectively these are known as diurnal rhythms. One of the earliest recorded diurnal rhythms was that of leaf movements in various plant species, whereby leaves move towards the stem (hyponasty) at night and move outwards from the stem (epinasty) during the day (de Mairan, 1729). Intriguingly, during early experiments it was observed that hyponasty/epinasty rhythms persisted under constant growth conditions (in the absence of external environmental cues) (de Mairan, 1729, Zinn, 1759). The term circadian (derived from Latin: circa = about, dies = day) was coined based on the observation that these rhythms persisted with a period of approximately 24 hours (reviewed in (McClung, 2006)). The majority of diurnal rhythms which have been described also display circadian characteristics. While circadian rhythms can persist independently from external cues in artificial scenarios, the two naturally act in concert. This is best illustrated by the process of entrainment in which the phase of a circadian rhythm is set by external temporal cues, or zeitgebers (time givers). Entrainment is a logical necessity for accurate circadian time-keeping for several reasons. Firstly, at non-equatorial latitudes the time of dawn and dusk change throughout the year. This would result in asynchrony between circadian rhythms and diurnal cycles if it were not for frequent entrainment to the changing diurnal conditions. Secondly, most circadian rhythms have a period close to, but not exactly, 24 hours. Even small period differences between the circadian/diurnal cycles would eventually lead to asynchrony, without an entrainment mechanism to repeatedly couple them. In plants, light plays a very important role in the entrainment of circadian rhythms. Consequently, light is a potent zeitgeber in natural conditions, and is the most frequently used zeitgeber in experimental conditions. By convention, the timing of diurnal/circadian cycles is expressed in zeitgeber time (ZT) units whereby ZT0 occurs at dawn and subsequent ZT values represent the relative time after dawn.

4.1.2 Molecular mechanisms of the plant circadian clock

The ability for circadian rhythms to persist without external cues implies the existence of an endogenous time-keeping device, which is often referred to as the circadian oscillator or clock. In most cases, the core oscillator of eukaryotic cells is composed of transcriptional/translational feedback loops. Interestingly, while circadian clocks are ubiquitous in eukaryotes, the lack of genetic synteny for core clock components implies that circadian oscillators have evolved independently in plants, fungi, and animals (McClung, 2006, Merrow *et al.*, 2005, Rosbash, 2009, Saunders, 2009). Decades of research into the circadian oscillator of plants indicates that it operates by a complex web of interactions involving a plethora of genetic loci (Pokhilko *et al.*, 2012). The core oscillator function, however, is often simplified to a few key components (Figure 4.1). These core components can be loosely compartmentalised into three groups. The first group is formed by two Myb (myoblastosis) transcription factors called LHY (LATE ELONGATED HYPOCOTYL) and CCA1 (CIRCADIAN CLOCK ASSOCIATED 1) which form heterodimers that regulate gene expression by binding to promoter motifs such as the 'Evening Element' (EE) and 'CCA1 Binding Site' (CBS). The second group consists of another class of transcription factors in the PRR (pseudo response regulator) family. The third group is the so-called evening complex (EC) which is a multimeric protein complex that operates, at least partially, by transcriptionally regulating other 'core clock' genes (Huang & Nusinow, 2016). *LHY* and *CCA1* are transcriptionally activated by light, resulting in a peak of *LHY/CCA1* expression around dawn and early morning. LHY/CCA1 promote expression of *PRR9*, *PRR7*, and *PRR*5 in a sequential manner by binding to CBS and EE motifs in their promoters (Farre *et al.*, 2005, Pokhilko et al., 2012). The PRR proteins transcriptionally repress *LHY/CCA1*, forming a feedback loop that restricts *LHY/CCA1* expression to the morning (Nakamichi *et al.*, 2010). The *PRR* with the latest expression peak, (*PRR1*, also known as *TOC1 – TIMING OF CAB EXPRESSION 1*) is transcriptionally repressed by LHY/CCA1 and so its expression is defined by its de-repression in the evening (Alabadi *et al.*, 2001). It was originally proposed that TOC1 indirectly promoted *LHY/CCA1* expression in the morning (Alabadi *et al.*, 2001), however it is now believed to act as a repressor to inhibit *LHY/CCA1* during the night (Pokhilko *et al.*, 2012). The EC forms another feedback loop that represses

PRR expression throughout the night to allow morning expression of *LHY/CCA1* by PRR de-repression (Huang & Nusinow, 2016)(Figure 4.1). The classification of LHY/CCA1, PRRs, and the EC as core clock components is justified based on the fact that mutation at any of these loci results in debilitated clock functions, mainly manifested by dramatic period alterations. The most severe clock defects are seen in the CCA1 over-expressor mutant (*CCA1-OX*) which is completely arrhythmic with respect to transcriptional circadian oscillations (Wang & Tobin, 1998).

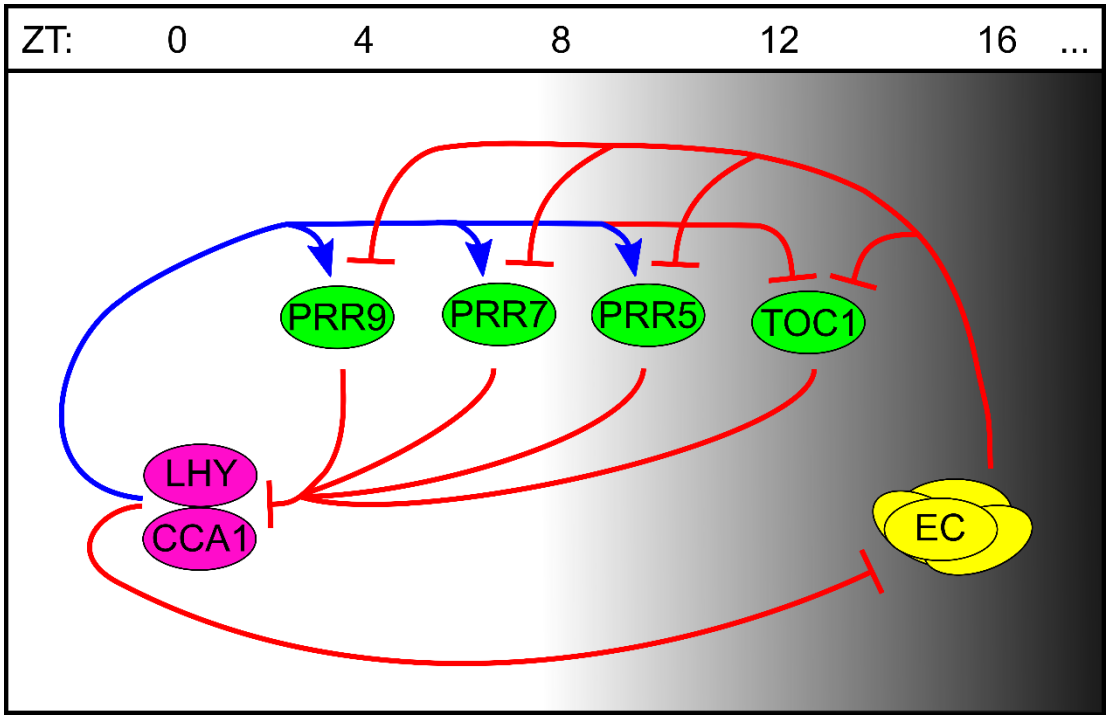

Figure 4.1 Simplified diagram of the plant transcriptional circadian oscillator, based mainly on the following published model (Pokhilko *et al.*, 2012). The approximate time of peak expression for each component is indicated by the ZT-hour scale, above. Blue/red lines indicate transcriptional induction/repression, respectively.

4.1.3 Functional relevance of the plant circadian clock

There are numerous examples which testify the importance of circadian rhythms in the lifecycle of plants. Firstly, it allows for temporal separation of processes to improve their efficiency or fidelity. For instance, cellular growth, division and DNA replication in plants occurs primarily during the night (Nozue *et al.*, 2007, Nusinow *et al.*, 2011). This is coincident with the catalysis of photosynthetic starch reserves which are stored during the day (Graf *et al.*, 2010). Hence peak growth occurs at the time when energy stores are most available. Furthermore, it has been suggested that restricting DNA replication to the night limits DNA damage from daytime mutagens such as solar UV irradiation and photosynthetic reactive oxygen species (Beator & Kloppstech, 1996, Pittendrigh, 1993). A second beneficial feature of the clock is that it allows plants to anticipate regularly changing external conditions. For instance, stomata opening and the expression of many photosynthesis related genes occur before dawn (Seo & Mas, 2015). This primes plants to begin photosynthesising as soon as photosynthetically active radiation becomes available. A third important role of the clock is that it enables plants to gauge seasonal changes. The best characterised example of this is the induction of flowering by changes in day length, by the so-called coincidence model (Bunning, 1936, Pittendrigh & Minis, 1964). Integration of signals from the endogenous oscillator and daylength affect the expression and stability of a central flowering regulator called CONSTANS (CO). Hence CO accumulates when the photoperiod is greater than a critical length. In plant species that flower in long days (long-day plants), CO promotes expression of flowering inducer genes. In contrast, CO acts as a suppressor of flowering inducer genes for short-day plants (Samach *et al.*, 2000, Samach & Gover, 2001).

The role of the circadian clock in responding to external abiotic cues and stresses has been relatively well characterised. Recently, evidence has begun to emerge that the clock may also play an important role in responses to biotic stresses such as pathogen attack. The first suggestions of this came from observations that many defence related genes exhibit diurnal and circadian rhythmic expression profiles (Molina *et al.*, 1997, Sauerbrunn & Schlaich, 2004, Weyman *et al.*, 2006). This should be interpreted with caution however, as whole-genome transcriptional time-course profiling in *Arabidopsis* has revealed that approximately 89% of all plant loci are expressed

rhythmically (Michael *et al.*, 2008), and approximately 35% of *Arabidopsis* genes are thought to be regulated by the circadian clock (Michael & McClung, 2003). Moreover, in *Drosophila melanogaster* it was observed that clusters of genes on the same chromosome have rhythmic expression profiles in a matching phase. This led to the speculation that, for some genes, circadian expression may have evolved to economise transcription at adjacent loci, implying that rhythmic expression of a locus does not necessarily signify an adaptive value of its rhythmic output *per se* (McDonald & Rosbash, 2001, Ueda *et al.*, 2002). However, mounting evidence suggests that the rhythmic expression of at least some of these pathogen-associated genes has evolved to provide temporal control over pathogen defence. For instance, *Arabidopsis* plants were found to be more resistant to the bacterial pathogen *Pseudomonas syringae* when inoculations were performed in the morning, compared to the evening (Griebel & Zeier, 2008). This response was, at least partially, explained by the observation that SA accumulated to higher levels during the day than the night, reflecting the requirement of light for its biosynthesis (Bechtold *et al.*, 2005, Zeier *et al*., 2004). In this study SA accumulation was found to oscillate in diurnal conditions (peaking at ZT0) but not under constant light indicating that it is under diurnal by not circadian regulation (Griebel & Zeier, 2008). However, recent studies have reported that SA biosynthesis does display circadian, as well as diurnal, rhythmicity (Goodspeed *et al.*, 2012, Goodspeed *et al.*, 2013), which is further corroborated by circadian regulation of genes involved in SA biosynthesis (Wang *et al.*, 2011, Wang *et al.*, 2014). Further to the diurnal differences in susceptibility to *P. syringae* (Griebel & Zeier, 2008), it was recently shown that *Arabidopsis* exhibits a circadian response to this same pathogen, such that plants are more resistant in subjective morning, compared to subjective night (Bhardwaj *et al.*, 2011). Importantly, the difference in susceptibility between subjective day/night was not observed in the arrhythmic *CCA-OX* mutant genotype, thus confirming that the response is regulated by the plant circadian clock. The authors further demonstrated that the increased resistance in the morning may be partly due to enhanced capacity for cell wall callose deposition (Bhardwaj *et al*., 2011). Hence a model is emerging which posits that plant resistance is enhanced in the morning compared to the evening, which is most likely driven by SA production, signalling, and associated immune responses. Care should be taken to not over-

generalise however, as most of these studies utilised the same pathogenic agent and have focussed mainly on SA-mediated defence responses. It is possible that circadian responses to other pathogens, involving different defence mechanisms, may exist, potentially exhibiting different characteristics. An example of this was recently demonstrated. The increased resistance to *P. syringae* at early ZT-hours under both diurnal and circadian conditions was confirmed, and the same trend was shown for resistance to the oomycete pathogen *Hyaloperonospora arabidopsidis*. However, when *P. syringae* was introduced to plants by spray inoculation rather than syringe inoculation, resistance was enhanced during the night and subjective night. This appears to contradict previous reports of enhanced pathogen resistance in the morning. The authors demonstrated that the enhanced night-time resistance seen in spray-inoculated plants is a result of stomatal regulation acting as a physical barrier to *P. syringae* invasion (Zhang *et al.*, 2013a). This phenomenon is masked when artificial inoculation methods such as syringe infiltration are used. In addition to highlighting the importance of taking a more holistic view of infection, to consider factors such as transmission as well as intracellular immunity, this study also raised an interesting concept – that different defence responses to the same pathogen may be timed to different, complementing phases of the day/night cycle to ensure a consistent and effective resistance response.

In the context of diurnal/circadian responses to viral pathogens, very little is currently known. An early report indicated that the necrotic lesion response to several viruses is enhanced during the day compared to the night, under natural diurnal conditions (Matthews, 1953a). This is in agreement with the fact that SA signalling (which is more active in the day) plays an important role in the development on necrotic lesions (Loebenstein, 2009). However, this early study (Matthews, 1953a) did not assess the role of the circadian clock directly, nor did it monitor accumulation of viral RNA but rather focussed only on a specific viral symptom (lesions). Hence, the roles of diurnal and circadian rhythms in plant antiviral responses are still largely unknown.

4.2 Chapter aims

This chapter aims to determine whether viral infections in plants exhibit diurnal and circadian rhythms and if so, how host factors and defence mechanisms may dictate and/or respond to this phenomenon.

4.3 Rhythmic transcription of antiviral RNA silencing genes

Analysis of the promoter regions of the *Arabidopsis DCL*, *RDR* and *AGO* paralogues involved in viral RNA silencing was carried out using an online promoter motif search tool (http://arabidopsis.med.ohio-state.edu/AtcisDB/). The original goal was to search for possible light- and temperature- regulatory motifs to corroborate the observations, presented in the previous chapter, that light and temperature shifts could modulate the efficacy of viral RNA silencing. While no heat responsive motifs were identified using this basic search, each of the loci tested contained multiple light responsive elements (LREs). These included the 'GATA', 'G-box', and 'SORLIP' motifs which have all been characterised as cis-regulatory sequences involved in light-mediated transcriptional regulation (Hudson & Quail, 2003, Kawoosa *et al.*, 2014). In addition to their roles in light-dependent transcriptional regulation in the context of growth and development, the GATA and G-box motifs have been implicated as important regulatory motifs for the entrainment of circadian-clock-regulated genes (Doherty & Kay, 2010). Interestingly, several of the vsiRNA-related loci analysed also contained well characterised circadian-clock-regulatory elements such as the CBS and EE motifs. These observations raised the intriguing possibility that many of the loci involved in RNA silencing of plant viruses may exhibit diurnal and/or circadian expression profiles. Table 4.1 lists the putative LREs and clock-regulatory elements which were bioinformatically detected for each of the vsiRNA-related loci.

Locus	Motif name	Sequence	Function
DCL2 & DCL4	GATA	agataa	Light regulation
	SORLIP1	agccac	Light regulation
	CBS	aaaaatct	Circadian clock regulation
RDR1	GATA	agataa	Light regulation
	SORLIP2	gggcc	Light regulation
RDR6	GATA	agataa	Light regulation
	G-box	cacgtg	Light regulation
AGO1 & AGO2	GATA	agataa	Light regulation
	SORLIP2	gggcc	Light regulation
	CBS	aaaaatct	Circadian clock regulation
	EE	aaaatatct	Circadian clock regulation
AGO7	GATA	agataa	Light regulation
	CBS	aaaaatct	Circadian clock regulation

Table 4.1 List of light- and circadian-clock- regulatory motifs detected, bioinformatically, in the promoters of vsiRNA-related loci in *Arabidopsis*.

To test whether the vsiRNA-related loci presented in Table 4.1 exhibit rhythmic expression profiles, a diurnal/circadian time-course experiment was conducted. For this, wild-type (Col-0) *Arabidopsis* plants were entrained for 4 weeks under photo-cycles consisting of a 12 hour light period and a 12 hour dark period (12:12 LD). After 4 weeks of entrainment under these conditions, plant tissue was sampled every 4 hours across one day/night cycle. In parallel, a set of plants which were entrained under the same 12:12 LD conditions for 4 weeks were shifted to constant light conditions (LL) and plant tissue was sampled every 4 hours for 2 day/subjective-night cycles, under LL conditions. qRT-PCR was then used to test for diurnal and circadian transcriptional regulation of the vsiRNA-related loci, which would be manifested by approximately 24 hour-period oscillations under LD and LL conditions, respectively. *DCL4*, *DCL2*, *RDR6,* and *AGO2* all exhibited rhythmic transcriptional activity under both diurnal (LD) and strictly circadian (LL) conditions (Figure 4.2A,B,D,E). The transcriptional profile of *AGO1* oscillated under LD but not LL conditions, indicating diurnal but not circadian rhythmicity at this locus (Figure 4.2G). In contrast, *RDR1* and *AGO7* did not exhibit obviously discernible transcriptional oscillations under either LD or LL conditions (Figure 4.2C,F).

The transcriptional profiles of *DCL4*, *DCL2*, *RDR6*, *AGO2*, and *AGO1* all oscillate with a period of ~24 hours, indicating *bona fide* diurnal/circadian rhythmicity. More interestingly, each of these rhythmic profiles oscillate with the same phase, specifically, rising over the night/subjective-night period with a peak of expression at ZT0, and falling over the day period (Figure 4.2A,B,D,E,G).

As for any qRT-PCR experiment, the reference gene used for ΔΔCp calculations to estimate the relative transcript abundance should be chosen with great care. This is particularly pertinent for circadian experiments due to the high proportion of genomic loci in plants that exhibit diurnal/circadian expression profiles (Michael & McClung, 2003, Michael *et al.*, 2008). Hence, normalising a target gene to a reference gene that oscillates in a diurnal/circadian manner could lead to erroneous results. *ACT7* was chosen as a reference gene for experiments presented herein, as it is commonly used for diurnal/circadian time-course qRT-PCR experiments (Edwards *et al.*, 2006, Flis *et al.*, 2015, Pokhilko *et al.*, 2012, Seaton *et al.*, 2015), due to its non-rhythmic expression over day/night time-series (Hong *et al.*, 2010). To further assess the validity of the qRT-PCR results, a qRT-PCR reaction for the core-oscillator gene *TOC1* was used as a positive control to confirm that the sampling conditions, or normalisation to *ACT7* expression did not bias the results for the various target loci. *TOC1* was chosen not only for its rhythmic expression profile, but also because its phase of oscillation (rising expression during the day and falling expression during the night) (Pokhilko *et al.*, 2012) is different to that observed for the vsiRNA-related genes shown in Figure 4.2. Observation of the predicted phase and period of *TOC1* circadian expression (Figure 4.2H), indicates that the experimental and analytical methods used accurately reflect the diurnal/circadian regulation of the target loci.

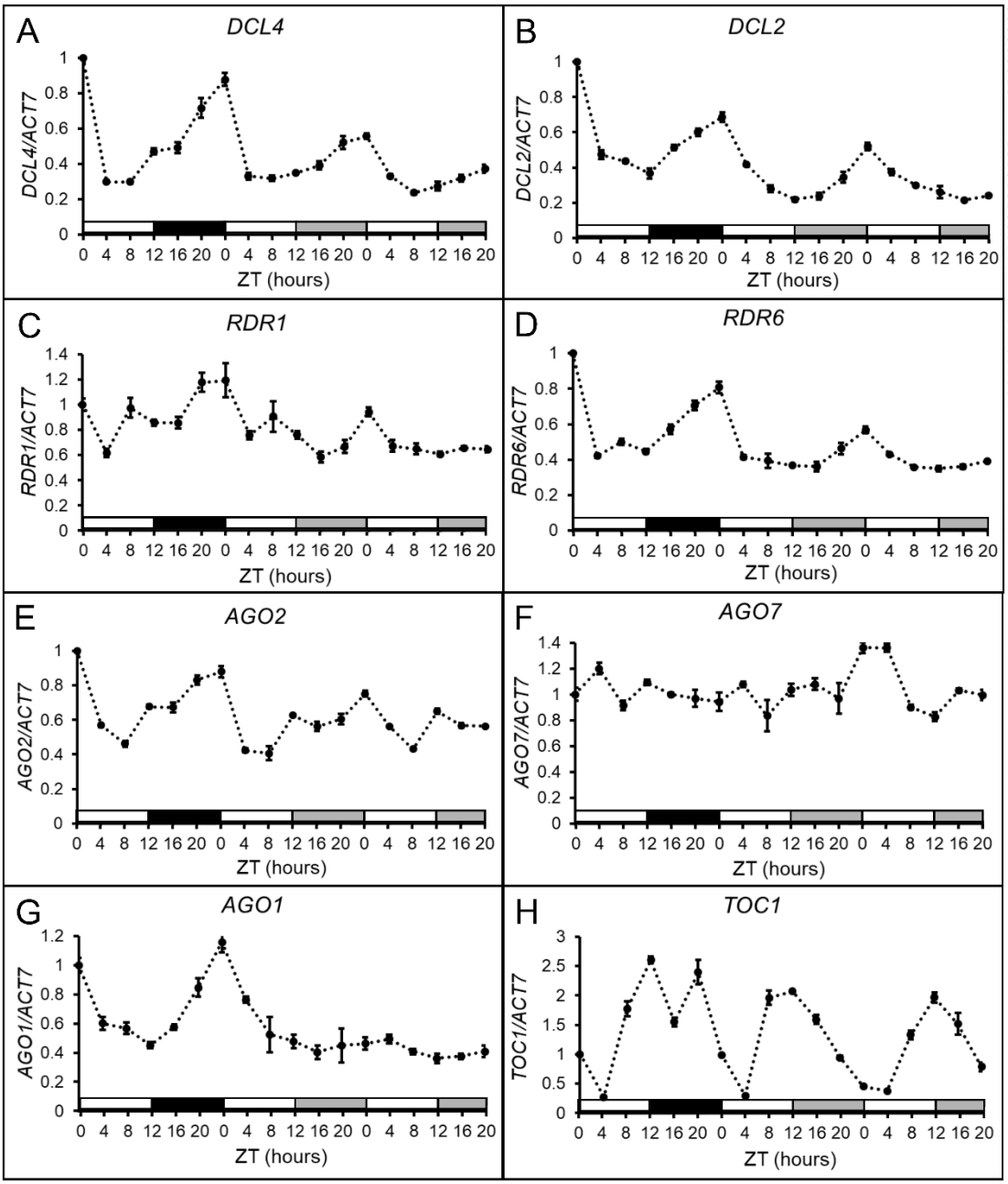

Figure 4.2 Transcriptional profiling of vsiRNA-related genes, listed in an order reflecting their functional hierarchies. *Arabidopsis* plants were entrained for 4 weeks under 12:12 light:dark cycles with 21ºC constant temperature. Whole rosettes were sampled every 4 hours over a 24-hour period in LD. Similar sampling was conducted over a 48-hour period for plants shifted to constant light (LL). Points represent the arithmetic mean of 3 plant samples ± SEM. *ACT7* was used as a reference gene for ΔΔCp calculations, and values are plotted relative to the mean value for ZT0 under diurnal conditions. White bars indicate the day period (lights on), black bars indicate the night period (lights off), grey bars indicate the subjective night period (lights on).

As a further measure to stringently and independently verify the results presented herein, published microarrays used to assess diurnal/circadian rhythmicity (Covington & Harmer, 2007, Harmer *et al.*, 2000, Mockler *et al.*, 2007) were analysed. Mining of these microarray data, validated most of the observed diurnal/circadian oscillations in the vsiRNA-related genes (Figure 4.3, Figure 4.4). The minor deviations between the qRT-PCR data (Figure 4.2) and published microarray data (Figure 4.3, Figure 4.4) are likely to be due to the differences in growth conditions and sampling. The qRT-PCR experiment sampled whole rosettes from mature, soil-grown plants whereas the microarrays sampled young (7 day old) seedlings grown on agar plates. On the whole, the microarray data supports the trends observed for the qRT-PCR time-course experiment.

Based on the qRT-PCR results for soil-grown *Arabidopsis* (Figure 4.2), it can be concluded that *DCL4*, *DCL2*, *RDR6*, and *AGO2* all exhibit ~24 hour period oscillations which persist both under diurnal (LD) and strictly circadian (LL) conditions. Similarly, *AGO1* transcripts oscillate with a ~24 hour period, though this oscillation appears to require external signals from photo-cycles to persist. Moreover, the synchronous phase of these oscillations imply that viral RNA silencing may be primed to be more active during the night compared to the day.

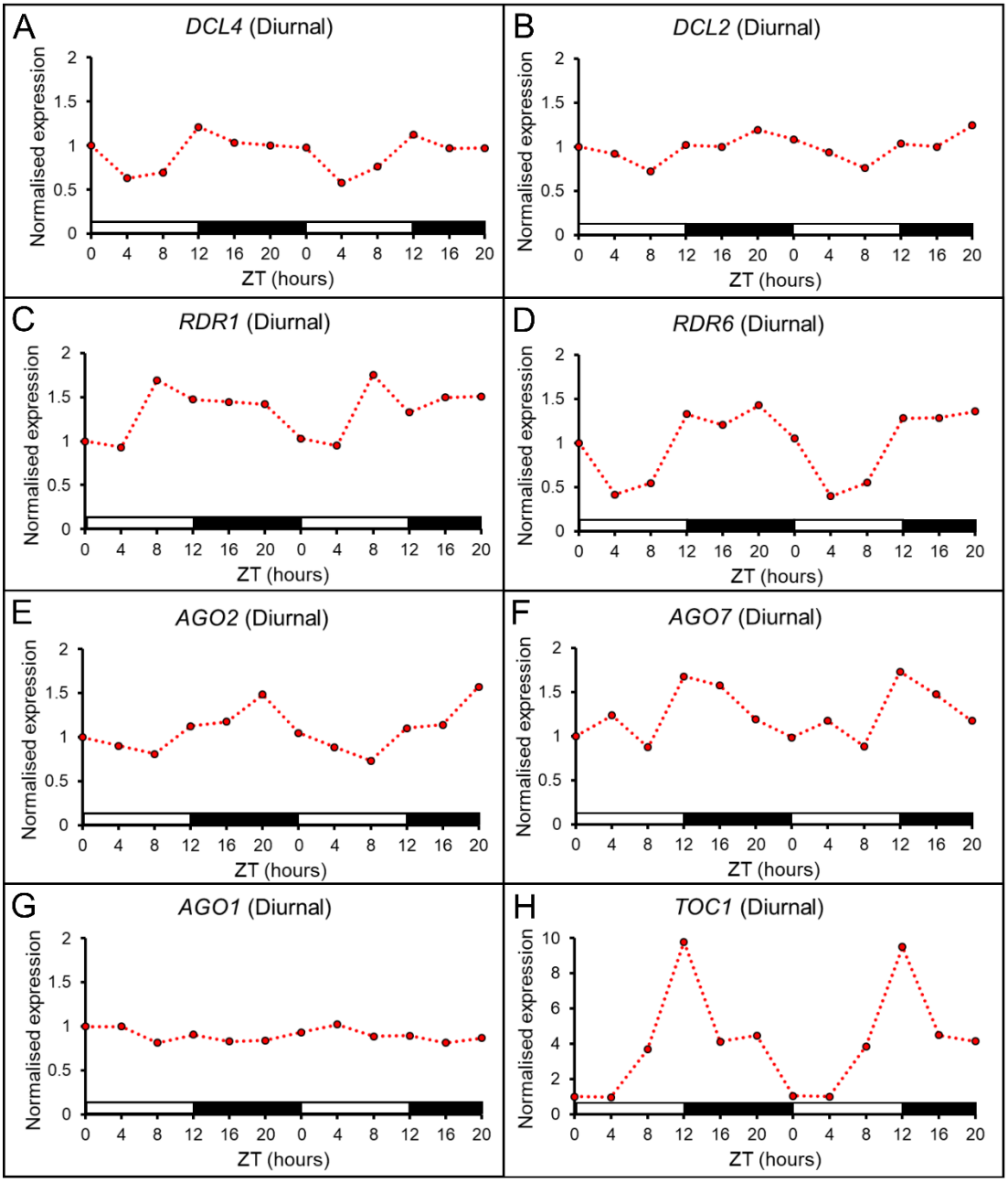

Figure 4.3 Relative expression of vsiRNA-related genes over a 48-hour time-course under diurnal (12:12 L:D) conditions. Microarray data was extracted from the following public database: http://diurnal.mocklerlab.org/ (Mockler *et al.*, 2007) Data points are plotted relative to the expression at ZT0 of the first day of sampling. White/black bars indicate day/night, respectively.

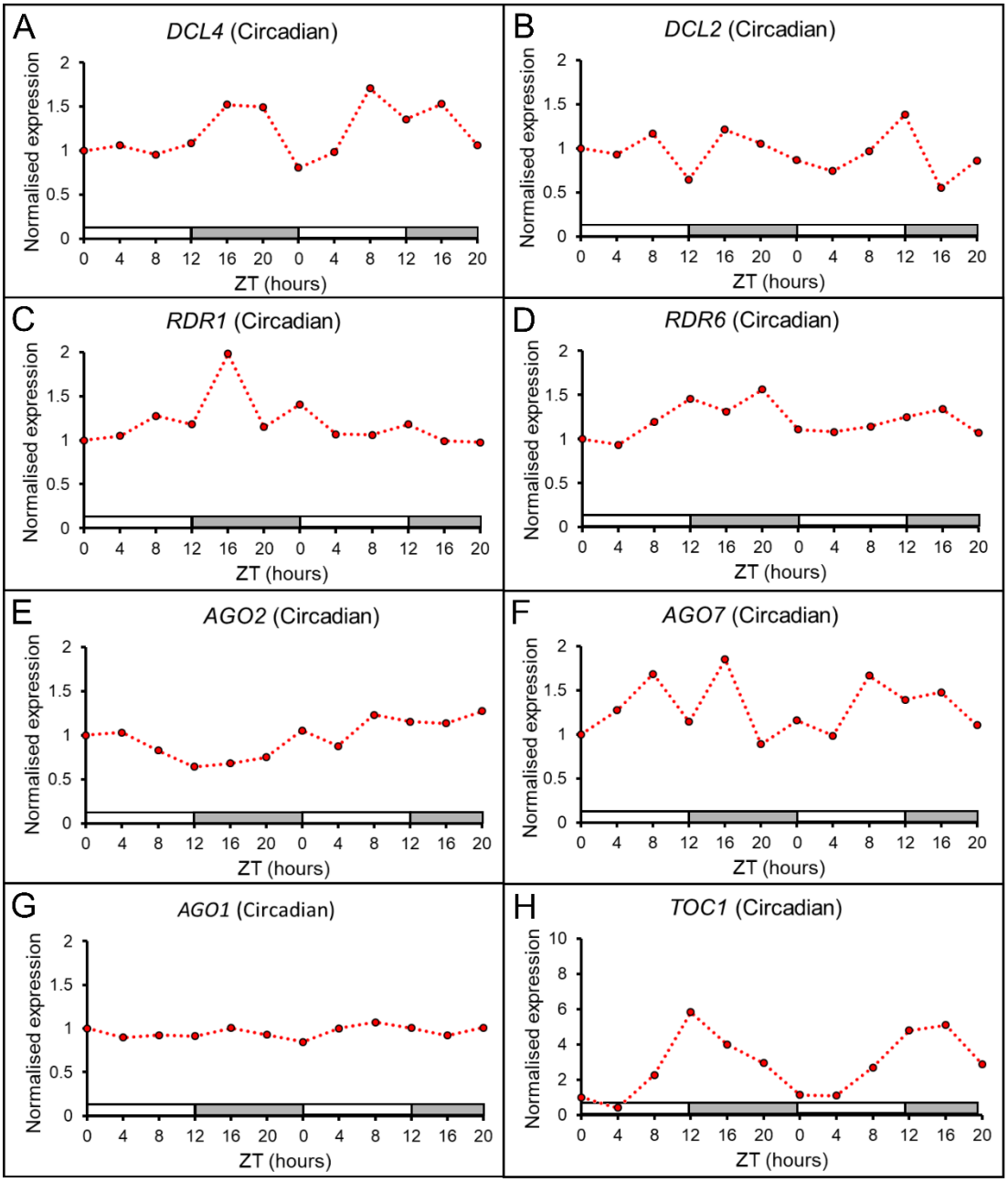

Figure 4.4 Relative expression of vsiRNA-related genes over a 48-hour time-course under constant light (12:12 L:L) conditions. Microarray data was extracted from the following public database: http://diurnal.mocklerlab.org/ (Mockler *et al.*, 2007) Data points are plotted relative to the expression at ZT0 of the first day of sampling. White/grey bars indicate day/subjective-night, respectively.

4.4 DCL4 protein abundance is rhythmic under circadian conditions

To further assess the rhythmic expression of vsiRNA-related genes, Western blotting was performed to test for oscillations in protein accumulation. DCL4 was chosen as a target for several reasons including; its rhythmic transcriptional expression profile (Figure 4.2A, Figure 4.3A Figure 4.4A), its role as the primary anti-TuMV DCL (Garcia-Ruiz *et al*., 2010), and the commercial availability of anti-DCL4 antibodies. The tissue sampled for the qRT-PCR time-course presented in Figure 4.2 was divided prior to RNA extraction to allow the same samples to be used for total protein extractions. A Western blot was performed for protein extracted from samples collected every 4 hours over 2 day/subjective-night cycles, to determine whether DCL4 protein abundance oscillates under LL conditions. Additionally, the ZT20 sample and ZT0, ZT4, and ZT8 samples from the end of the LD photo-cycle were included to assess the role of the dark-to-light transition in DCL4 protein accumulation. Concordant with previously published DCL4 Western blots (Azevedo *et al*., 2010, Schott *et al.*, 2015), two specific bands corresponding to the predicted molecular mass (~191kDa) for DCL4 were observed. It was recently discovered that two different DCL4 isoforms are expressed in *Arabidopsis* due to differential transcriptional initiation within the same *DCL4* open reading frame (ORF) (Pumplin *et al.*, 2016). The larger DCL4 isoform ($DCL4^{NLS}$) contains a 5'- nuclear localisation signal (NLS) and results from transcription of the full DCL4 ORF. The smaller DCL4 isoform ($DCL4^{\Delta}$) lacks the 5'-NLS and hence is predominantly localised to the cytoplasm. $DCL4^{\Delta}$ results from transcription initiating from an internal transcriptional start site within the DCL4 ORF. Based on these findings, it is reasonable to conclude that the slower-migrating (upper) anti-DCL4 signal on the Western blot corresponds to $DCL4^{NLS}$, while the faster-migrating (lower) anti-DCL4 signal represents $DCL4^{\Delta}$ (Figure 4.5). However, these assumptions are yet to be formally tested. The anti-DCL4 signal for both $DCL4^{NLS}$ and $DCL4^{\Delta}$ appears to oscillate over the day/subjective-night cycles under LL conditions, indicating that DCL4 protein abundance is a rhythmic output of the circadian clock. Furthermore, these oscillations exhibit a similar phase to that seen for the DCL4 transcript (Figure 4.2A), namely; an increase during subjective night and a decrease during the day. Interestingly, the peaks of $DCL4^{NLS}$ and $DCL4^{\Delta}$ protein abundance were not synchronous, occurring at ZT20 and ZT0, respectively.

This could possibly reflect temporal separation of the peak activity of the nuclear and cytoplasmic DCL4 isoforms, though further validation and the potential biological relevance of this observation remain to be further investigated. The decrease in DCL4 protein abundance is slightly less rapid than the decrease in DCL4 transcript abundance, over the circadian oscillations. This is likely to be partly due to the delay between transcript accumulation and translation, and possibly also due to a greater stability of the DCL4 protein isoforms, compared to the DCL4 transcript. It is possible that light-mediated degradation of the DCL4 protein isoforms may sharpen their expression profile to restrict their expression to the night period. This notion is supported by the fact that light-dependent proteosomal degradation is involved in maintaining oscillations of 'core clock' proteins (Baudry *et al.*, 2010). Evidence that this may apply to DCL4 regulation can be deduced from the protein expression patterns of the two DCL4 isoforms for the dark-to-light samples at the end of the LD cycle (Figure 4.5). The DCL4 protein abundance exhibits a rapid decrease during the transition from darkness to light, mirroring the similarly rapid decrease observed for DCL4 transcript abundance immediately after ZT0 under LD conditions (Figure 4.2A). Interestingly, the decrease in protein abundance at the dark-to-light transition was greater for $DCL4^{NLS}$ than for $DCL4^{\Delta}$ (Figure 4.2). It is noteworthy that the primers used for assessing *DCL4* expression by qRT-PCR were designed to amplify the 3' region of the *DCL4* transcript. Hence, the qRT-PCR results reflect the combined abundance of $DCL4^{NLS}$ and $DCL4^{\Delta}$ transcripts. Tentatively, these results suggest that, under LD cycles, the transition from darkness to light may help to sharpen the night-time peak of DCL4 protein accumulation.

In conclusion, DCL4 protein abundance (as well as transcriptional activity) appears to be temporally regulated under constant (LL) conditions, though determination of the exact peak and phase is difficult due to the short sampling period. Furthermore, light signals under LD cycles may be important for both entrainment of the transcriptional oscillation and period maintenance for the protein oscillation. Interestingly, the profiles of the nuclear- and cytoplasmic- localised DCL4 isoforms ($DCL4^{NLS}$ and $DCL4^{\Delta}$) may be slightly divergent, though this preliminary observation is subject to future investigation. Another worthwhile focus for future study would be to ascertain whether the other vsiRNA-related genes that exhibit transcriptional oscillations

(*DCL2*, *RDR6*, *AGO2*, and *AGO1*) also show rhythmic protein expression. A lack of reliable antibodies prevented assessment of these proteins in this study. The circadian expression of DCL4 protein however, supports the notion that antiviral RNA silencing may be primed for greater activity during the night/early morning compared to day-time/evening, at least when grown under 12:12 LD photoperiods.

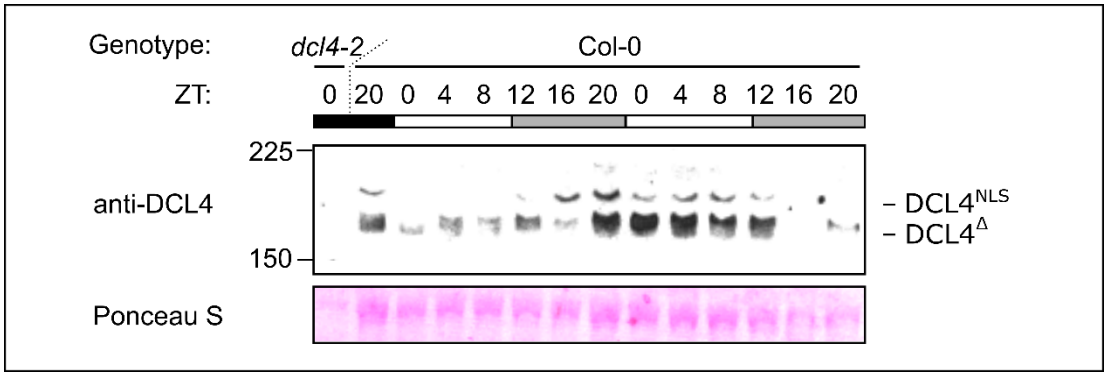

Figure 4.5 Western blots for semi-quantitative assessment of DCL4 protein accumulation. Plants were grown and sampled as described for Figure 4.2. A sample collected at ZT0 from a *dcl4-2* mutant was included as a negative control. The two DCL4 specific bands ($DCL4^{NLS}$ and $DCL4^{\Delta}$) migrated between the 150 and 225 kDa products of the molecular weight marker (GE Healthcare, Cat# RPN800E). A Ponceau S stain of the membrane is included to verify even sample loading. White bars indicate the day period (lights on), black bars indicate the night period (lights off), grey bars indicate the subjective night period (lights on).

4.5 Susceptibility to TuMV-GFP infection exhibits diurnal and circadian rhythmicity in *Arabidopsis*

The rhythmic expression of vsiRNA-related genes suggests that the susceptibility of plants to viruses may change across day/night cycles. Furthermore, the fact these oscillations in gene expression for vsiRNA-related loci appear to be regulated by the circadian clock in most cases, suggests that rhythms in virus susceptibility may also persist under constant growth conditions. To test these hypotheses, an experiment was conducted to assess the time-of-day effects on the susceptibility of *Arabidopsis* to TuMV-GFP infections. *Arabidopsis* plants were entrained for 4 weeks under 12:12 LD conditions and then separated into two groups, hereafter referred to as the 'LD group' and the 'LL group'. Both groups were rub-inoculated with TuMV-GFP at ZT0, ZT6, ZT12, and ZT18. For the 'LD group' these inoculations were performed under the same 12:12 LD conditions used for the 4-week entrainment period, and these conditions were maintained throughout the experiment. In contrast, plants in the 'LL group' were shifted to constant light conditions for 24 hours before the first inoculation time-point. The 'LL group' was henceforth kept under these LL conditions for the remainder of the experiment. The inoculated plants for the 'LD/LL groups' were kept under their respective LD/LL conditions for 5 days before sampling the inoculated leaves to quantify TuMV-GFP RNA by qRT-PCR. Importantly, the inoculated leaves were sampled at exactly the same ZT-hour that they were inoculated, to ensure that the virus incubation time was identical for all of the samples. Furthermore, as the time-of-day effects on TuMV-GFP resistance/susceptibility were presumed to be subtle, the inoculated leaves were thoroughly washed immediately prior to tissue sampling to limit the effects of residual sap on the surface of the inoculated leaves confounding the measurements of the intracellular viral RNA. In addition to wild type (Col-0) *Arabidopsis*, the *CCA1-OX* genotype was inoculated to assess the requirement of a functional circadian oscillator on any observed time-of-day effects on susceptibility to TuMV-GFP.

For the inoculations onto wild type plants, time-of-day effects on TuMV-GFP RNA accumulation were clearly discernible for both the 'LD' and 'LL' groups. Specifically, in the LD group TuMV-GFP RNA was highest when inoculations were performed at ZT0 and progressively lower at ZT6 and ZT12 (Figure 4.6A), and higher again at

ZT18. In the LL group, TuMV-GFP RNA was also highest for inoculations performed at ZT0, and lower at ZT6 and ZT12. In contrast to the LD group however, inoculations at ZT18 resulted in the lowest TuMV-GFP RNA (Figure 4.6A). This difference is possibly due to the fact that many outputs of circadian clocks show reduced amplitudes and altered periods under LL conditions (Thain *et al.*, 2002). This could be assessed in future experiments by extending the inoculation/sampling time-points to extend over at least 2 full day/subjective-night cycles. Nevertheless, the results for the 'LD' and 'LL' group inoculations in wild type *Arabidopsis* indicate that there are clear time-of-day differences in susceptibility to TuMV-GFP under both diurnal (LD) and strictly circadian (LL) conditions.

Differences in TuMV-GFP susceptibility were also observed in the 'LD group' of the *CCA1-OX* genotype, though the phase of the TuMV-GFP-susceptibility oscillation appears to have an approximately 6-hour phase-shift compared to that of the WT 'LD group' (Figure 4.6A,B). These data imply that the observed time-of-day effects on TuMV-GFP susceptibility in LD can persist in the absence of a functional transcriptional oscillator, but rely on the concerted action of the external light cues and the circadian clock for correct phasing. In contrast, no differences in TuMV-GFP RNA were observed for the different ZT-hour inoculations in the 'LL group' for the *CCA1-OX* genotype. Unfortunately, the ZT18 sample under these conditions could not be analysed due to an insufficient yield of extracted RNA. However, by comparing the ZT0, ZT6, and ZT12 time-points of the 'LL' groups in the WT and *CCA1-OX* genotypes it can be concluded that a functional circadian oscillator is required for the time-of-day effects on TuMV-GFP susceptibility in constant LL conditions (Figure 4.6A,B).

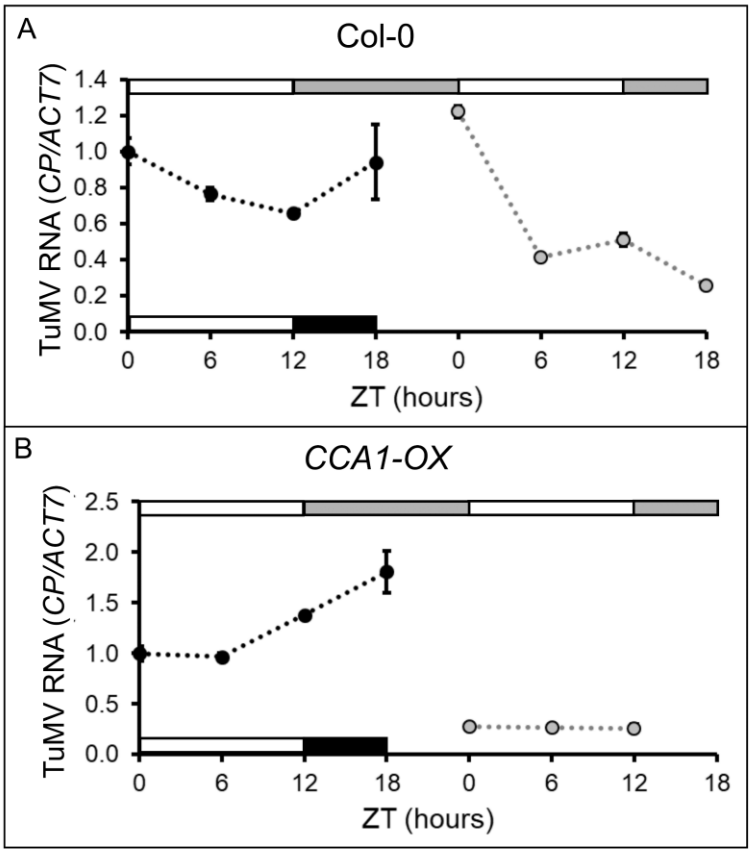

Figure 4.6 Quantification of TuMV-GFP RNA by qRT-PCR for inoculated leaves of **A** Wild type and **B** *CCA1-OX* genotypes. *Arabidopsis* plants were entrained for 4 weeks under 12:12 light:dark cycles with 21^{o}C constant temperature. Plants in the 'LD group' (black points/lines) were inoculated under light:dark photocycles. Plants in the 'LL group' (grey points/lines) were moved to constant light for 24 hours before inoculations. The ZT hour indicates the time of inoculation and tissue sampling. Samples were collected at 5dpi, pooling 16 inoculated leaves from 4 plants as one biological replicate. 3 biological replicates were sampled in this way. Hence, each point represents the arithmetic mean of 3 samples ± SEM. *ACT7* was used as a reference gene for ΔΔCp calculations, and values are plotted relative to the mean value for ZT0 under diurnal conditions. White bars indicate the day period (lights on), black bars indicate the night period (lights off), grey bars indicate the subjective night period (lights on). Shaded bars above and below the chart indicate the light regimes for the 'LL' and 'LD' groups, respectively.

4.6 The role of antiviral *DCL* genes in night-time suppression of TuMV-GFP infection

The finding that the greatest TuMV-GFP susceptibility occurred at ZT0 was unexpected as it coincides with the peak of expression for the vsiRNA-related genes that were found be under diurnal/circadian control. The initial assumption was that the peak in RNA silencing activity at, or shortly after, ZT0 would result in greater resistance (lower susceptibility) when inoculations were performed at this time point. The observation of the contrary (Figure 4.6A) could arise from a few non-mutually exclusive scenarios. First, as a null-hypothesis, it is conceivable that the rhythmic expression of vsiRNA-related genes could have no effect on the accumulation of TuMV-GFP RNA, and the observed differences in TuMV-GFP RNA for different ZT-hour inoculations could be completely dependent on un-related phenomena. Secondly, the rhythmic expression of vsiRNA-related genes could be the main determinant generating the TuMV-GFP-RNA oscillations but a delay between the artificial inoculation time and peak viral synthesis could result in the 'effective' inoculation ZT-hour lagging behind the ZT used for rub-inoculation. If this delay were to be approximately 12 hours, this could result in the naturally anti-correlated oscillations (of RNA silencing and viral susceptibility) becoming phase-shifted into apparent synchrony. Finally, the rhythmic expression of vsiRNA-related genes could be temporally and functionally relevant but may be a response to, rather than the determinant of, the observed oscillations in TuMV-GFP RNA. Potential candidates for additional factors that could generate oscillations in susceptibility include rhythmic expression of host factors required for viral synthesis, or loci involved in other antiviral defence pathways.

In order to test the above hypotheses, *dcl2-1dcl3-1dcl4-2* mutant plants were included (alongside the wild type and *CCA1-OX* genotypes) in the previously described experimental design for 'LD/LL group' inoculations at ZT0, ZT6, ZT12, and ZT18. If the temporal pattern of TuMV-GFP susceptibility were to be identical in the wild type and *dcl2-1dcl3-1dcl4-2* genotypes, then the null hypothesis (above) would be favoured. On the other-hand arrhythmicity, with respect to TuMV-GFP susceptibility, in the *dcl2-1dcl3-1dcl4-2* genotype would argue that these loci are the main determinants for generating oscillations in TuMV-GFP susceptibility. A more nuanced

difference between the TuMV-GFP-RNA profiles in the wild type and *dcl2-1dcl3-1dcl4-2* genotypes would argue that the oscillations in *DCL2*/*DCL3*/*DCL4* expression may be functionally relevant but do not constitute the main determinant of the rhythmicity of TuMV-GFP susceptibility.

For ease of interpretation, the TuMV-GFP qRT-PCR results for the *dcl2-1dcl3-1dcl4-2* genotype (Figure 4.7) are plotted superimposed to the previously presented results obtained for wild type genotype (Figure 4.6A). In the 'LD group' both wild and *dcl2-1dcl3-1dcl4-2* genotypes exhibited a similar temporal trend of TuMV-GFP susceptibility. Specifically, TuMV-GFP RNA was lower for day-time inoculations and higher for night-time inoculations (Figure 4.7). The persistence of time-of-day differences in TuMV-GFP susceptibility in the *dcl2-1dcl3-1dcl4-2* genotype indicates that these loci alone are not responsible for generating the rhythmic susceptibility phenotype.

While the temporal pattern of susceptibility in the 'LD group' was similar for the two genotypes, it was not identical, thereby refuting the null hypothesis (outlined above). For inoculations performed at ZT0 and ZT6 there was no difference in TuMV-GFP RNA for the two genotypes. In contrast, inoculations at ZT12 and ZT18 resulted in a greater than 2-fold increase in TuMV-GFP RNA in *dcl2-1dcl3-1dcl4-2* relative to wild type. These results imply that the DCL2/DCL3/DCL4 antiviral function is dispensable during the day but required for efficient night-time suppression of TuMV-GFP. This conclusion is further supported by the previous findings that expression of several vsiRNA-related genes oscillate with a peak in the night period (Figure 4.2, Figure 4.5).

A similar trend was observed for the TuMV-GFP RNA in wild type and *dcl2-1dcl3-1dcl4-2* in the 'LL group'. TuMV-GFP RNA were higher in *dcl2-1dcl3-1dcl4-2* than wild type for inoculations performed at ZT12 and ZT18 (Figure 4.7). In contrast to the 'LD group' however, the TuMV-RNA were not identical for the two genotypes for ZT0 and ZT6 inoculations. This could possibly be due to previously mentioned effects of period and phase alterations to circadian rhythms which can occur under LL conditions. Particularly, as LD cycles were previously suggested to be required to restrict the expression domain of DCL4 protein to the night time, it is possible that the functional output of vsiRNA-related loci may extend to parts of the day under LL

conditions. To better explain the different TuMV-GFP-susceptibility responses seen in the 'LD/LL groups' however, more extensive inoculation and sampling times across several subjective day/night LL cycles would be required.

In conclusion, the consistent increase in TuMV-GFP susceptibility in *dcl2-1dcl3-1dcl4-2* for inoculations performed immediately before, or during the night/subjective night period (ZT12 and ZT18 in LD and LL) indicates that DCL2/DCL3/DCL4 play an important role in the night-time suppression of TuMV-GFP under both diurnal and strictly circadian conditions. This is in agreement with the previous observation that vsiRNA-related genes exhibit a peak of expression during the night. The persistence of TuMV-GFP susceptibility oscillations in the *dcl2-1dcl3-1dcl4-2* genotype implies that factors other than DCL2/DCL3/DCL4 are required to generate the oscillatory pattern of TuMV-GFP susceptibility.

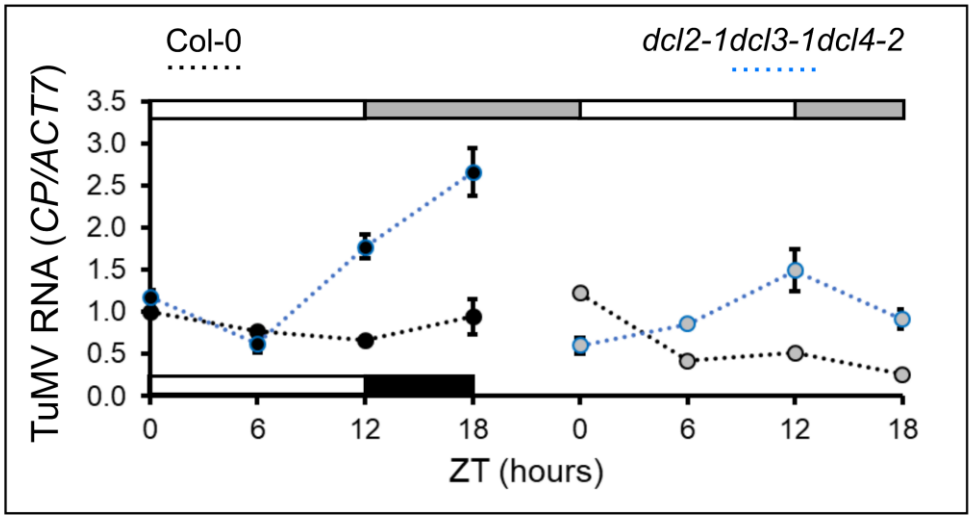

Figure 4.7 Quantification of TuMV-GFP RNA by qRT-PCR for inoculated leaves of wild type (black lines) and *dcl2-1dcl3-1dcl4-2* (blue lines) genotypes. *Arabidopsis* plants were entrained for 4 weeks under 12:12 light:dark cycles with 21°C constant temperature. Plants in the 'LD group' (black points) were inoculated under light:dark photocycles. Plants in the 'LL group' (grey points) were moved to constant light for 24 hours before inoculations. The ZT hour indicates the time of inoculation and tissue sampling. Samples were collected at 5dpi, pooling 16 inoculated leaves from 4 plants as one biological replicate. 3 biological replicates were sampled in this way. Hence, each point represents the arithmetic mean of 3 samples ± SEM. *ACT7* was used as a reference gene for ΔΔCp calculations, and values are plotted relative to the mean value for ZT0 under diurnal conditions in the wild type genetic background. White bars indicate the day period (lights on), black bars indicate the night period (lights off), grey bars indicate the subjective night period (lights on). Shaded bars above and below the chart indicate the light regimes for the 'LL' and 'LD' groups, respectively.

4.7 eIF(iso)4E is a candidate host factor which may generate the time-of-day rhythms in TuMV-GFP infection

eIF(iso)4E is an essential host factor that is strictly required for TuMV proliferation in *Arabidopsis* (Duprat *et al*., 2002, Lellis *et al.*, 2002, Sato *et al*., 2005) and is likely to be integral to several stages of the TuMV lifecycle including translation, replication, and intercellular movement (Gao *et al*., 2004, Leonard *et al*., 2000, Puustinen & Makinen, 2004, Wang *et al.*, 2000). Furthermore, in non-infected, differentiated leaf tissue the expression of *eIF(iso)4E* is much lower than the *eIF4E* paralogue (Rodriguez *et al.*, 1998). Based on these observations, it is conceivable that the availability of eIF(iso)4E may be a rate limiting factor for nascent TuMV infections. A rhythmic expression profile of *eIF(iso)4E*, with a matching or slightly earlier phase to the observed oscillations in TuMV-GFP susceptibility, would implicate this host factor as a likely candidate for generating the time-of-day effects on TuMV-GFP susceptibility.

To test whether *eIF(iso)4E* expression oscillates under diurnal and/or circadian conditions, the cDNA library from healthy, 4-week old *Arabidopsis* plants entrained under 12:12 LD cycles and sampled under LD and LL conditions (used for Figure 4.2) was used to profile the expression of *eIF(iso)4E* across day/night cycles. The transcriptional activity of *eIF(iso)4E* exhibits ~24-hour period oscillations under both LD and LL photocycles indicating diurnal and circadian rhythmicity at this locus (Figure 4.8A). Moreover, the oscillation rises over the night/subjective night period and falls during the day with a peak of expression at, or slightly before dawn/subjective dawn (ZT0). The fact that the phase of the *eIF(iso)4E* transcriptional oscillation (Figure 4.8A) is coincident to, or slightly in advance of, the oscillation in TuMV-GFP susceptibility (Figure 4.6A) supports the hypothesis that the diurnal/circadian expression of this essential host factor may result in time-of-day differences in TuMV-GFP susceptibility. Furthermore, the trend of eIF(iso)4E expression was supported by the previously mentioned microarray data (Figure 4.8B,C) (Mockler *et al*., 2007) Future experiments are required, however, to determine whether these *eIF(iso)4E* transcriptional oscillations are necessary or sufficient to generate TuMV-GFP susceptibility rhythms (see disucussion).

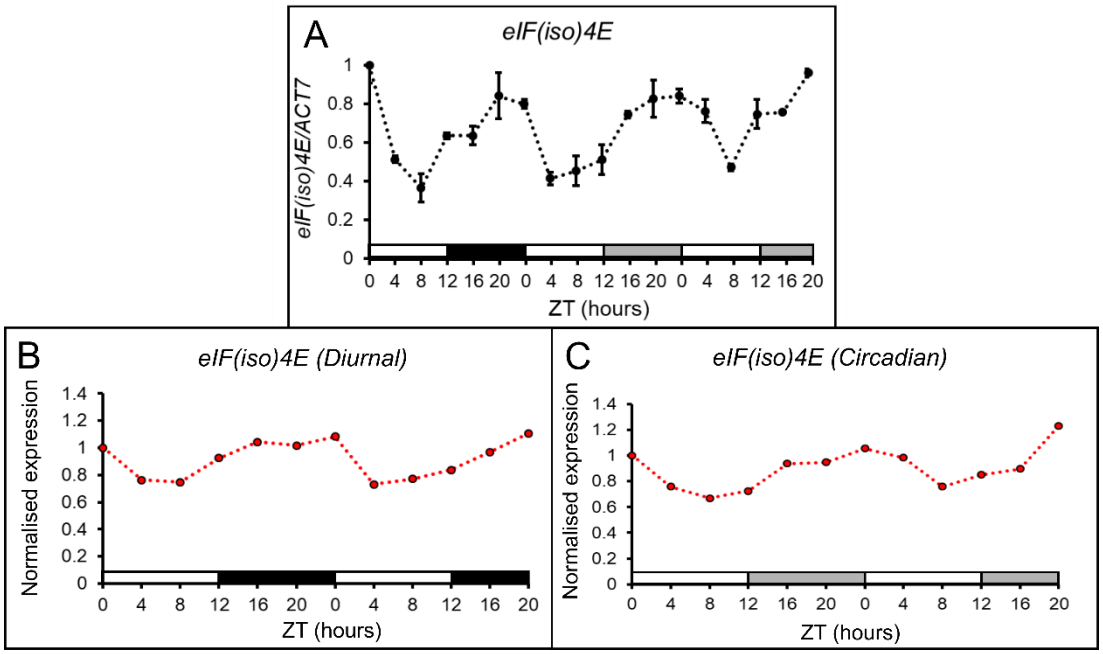

Figure 4.8 Transcriptional expression profile for *eIF(iso)4E*. **A** *Arabidopsis* plants were entrained for 4 weeks under 12:12 light:dark cycles with 21ºC constant temperature. Whole rosettes were sampled every 4 hours over a 24-hour period in LD. Similar sampling was conducted over a 48-hour period for plants shifted to constant light (LL). Points represent the arithmetic mean of 3 plant samples ± SEM. *ACT7* was used as a reference gene for ΔΔCp calculations, and values are plotted relative to the mean value for ZT0 under diurnal conditions. White bars indicate the day period (lights on), black bars indicate the night period (lights off), grey bars indicate the subjective night period (lights on). **B**,**C** Microarray data extracted from http://diurnal.mocklerlab.org/ (Mockler *et al.*, 2007), showing relative *eIF(iso)4E* expression across day/night (**B**) and day/subjective-night (**C**) cycles. Data are plotted relative to the value at ZT0 of the first day of sampling.

4.8 Chapter discussion

4.8.1 Virus susceptibility in plants varies over day/night cycles in a circadian-clock dependent manner

Results presented in this chapter show, for the first time, that susceptibility to a virus infection varies across day/night cycles in both diurnal conditions (LD) and under constant light (LL) (Figure 4.6A). This, along with the finding that these susceptibility oscillations under LL conditions are absent in the loss-of-function circadian clock mutant (*CCA1-OX*) (Figure 4.6B) implicate a function for the circadian clock in this response. However, experimental limitations restricted the sampling period in constant light conditions to less than 48 hours which makes it difficult to confidently assess circadian rhythmicity. Hence these data should be viewed as tentative results. Furthermore, where trends were observed under constant conditions they were more subtle than those observed under diurnal growth conditions, again necessitating further research to support the tentative observations under constant conditions. These findings extend our knowledge of plant-pathogen responses that are influenced by the circadian clock, beyond previous studies focussing on bacterial (Bhardwaj *et al.*, 2011, Griebel & Zeier, 2008, Zhang *et al*., 2013a) and oomycete (Zhang *et al*., 2013a) pathogens. Interestingly, these results also emulate recent findings of temporal and circadian influences on virus susceptibility in mammalian cells (Edgar *et al.*, 2016).

4.8.2 Antiviral RNA silencing operates primarily as a 'night-time' viral defence pathway

The finding that several vsiRNA-related genes may be expressed rhythmically and synchronously under the influence of the circadian clock and, most likely, environmental (light/dark) inputs (Figure 4.2, Figure 4.5) led to the hypothesis that rhythmic activity of viral RNA silencing may result in differential virus susceptibility across day/night cycles. However, the finding that the peak of TuMV-GFP susceptibility (Figure 4.6A) was coincident with the peak of vsiRNA-related gene expression suggested a more complex scenario exists than originally hypothesised. By comparing the temporal patterns of TuMV-GFP RNA accumulation between wild type (Col-0) and *dcl2-1dcl3-1dcl-4-2* genotypes, it was possible to deduce that DCL-dependent viral RNA silencing plays an important role limiting TuMV-GFP infections

during the night, but may not be necessary for efficient day-time defence (Figure 4.7). Consistent with this interpretation is the observation that vsiRNA-related genes, specifically *DCL4*, *DCL2*, *RDR6*, *AGO2*, and *AGO1*, exhibit peaks of expression at the end of the night period (Figure 4.2). Furthermore, these results are interesting in the context of previous findings that SA-mediated defence is more active during the day (Griebel & Zeier, 2008, Goodspeed *et al.*, 2012, Goodspeed *et al*., 2013). This suggests that SA-mediated defence constitutes the main day-time antiviral defence, while RNA silencing forms the predominant viral defence mechanism during the night. It would be interesting to test this hypothesis further by infecting *Arabidopsis* mutants deficient in SA production and/or perception with TuMV-GFP, or other viral pathogens. The model proposed herein would posit that no difference in viral susceptibility would be seen (relative to wild type plants) during the night, but mutants with defective SA signalling should be more susceptible during the day.

Plants may have evolved such a temporal separation of complementing defence pathways to enable a consistent defence across day/night cycles. This parallels previous findings of asynchronous physical and chemical defence responses shown to be active against *P. syringae* infections (Zhang *et al*., 2013a). Explanations for the adaptive value for timing SA-mediated defence to day-time often centre around the role of light in SA production and signal transduction (Bechtold *et al*., 2005, Griebel & Zeier, 2008, Zeier *et al*., 2004). In addition to this, it may be evolutionarily favourable to restrict SA-mediated defence pathways from being too active at night because of the so-called 'growth vs. defence' paradigm mentioned in the previous chapter discussion. As plant growth primarily occurs during the night (Nozue *et al*., 2007, Nusinow *et al*., 2011), and SA-associated defence responses are known to be restrictive to growth (Robert-Seilaniantz *et al*., 2011), the temporal regulation of SA-mediated defence may also be important limit the detrimental effects of pathogen defence responses on growth. RNA silencing may have evolved a night-time peak of activity to compensate for the reduced SA-mediated defence during this period. Though speculative, it is possible that viral RNA silencing may be less detrimental to growth as it provides a more specific defence response (in most cases) by producing vsiRNAs that directly target the viral genome for degradation. It is noteworthy that this timing of peak viral RNA silencing during the night is concordant with the results

presented in the previous chapter, whereby shifting plants to extremely low light intensities resulted in a greater activity of viral RNA silencing. Hence it seems likely that the evolution of nocturnal RNA silencing activity may partially explain the artificial 'skoto-therapy' phenomenon that can help to induce recovery from TuMV-GFP infections.

4.8.3 Rhythmic expression of host factors may be the underlying mechanism for time-of-day differences in virus susceptibility

The persistence of TuMV-GFP susceptibility oscillations in the *dcl2-1dcl3-1dcl4-2* mutant indicates that DCL2/DCL3/DCL4 functions are not necessary for generating the rhythmic pattern of TuMV-GFP susceptibility. However, it is possible that the residual anti-TuMV DCL activity of DCL1, along with the rhythmic expression of *RDR6* and *AGO1* could result in oscillatory activity of viral RNA silencing, even in the *dcl21-dcl3-1dcl4-2* mutant. Therefore, it is conceivable that the rhythmic activity of viral RNA silencing could be responsible for generating the oscillatory pattern of TuMV-GFP susceptibility. A more extensive genetic approach using mutant plants deficient in DCL2, DCL3, DCL4, RDR6, and AGO1 functions could be used to test this, though generating such a genotype would be extremely laborious. Additionally, experiments would be required to address the non-intuitive synchrony between peak activity of viral RNA silencing and the peak of TuMV-GFP susceptibility. This could be evaluated by quantifying the 'lag-phase' of local TuMV-GFP infections, namely the time delay between rub-inoculation and initiation of exponential replication of the virus. By correcting the ZT-hour to reflect the time of exponential replication of TuMV rather than time of rub-inoculation, it is possible that the peak of viral RNA silencing activity would correspond to the ZT-hour with lowest TuMV-GFP susceptibility. This would require a lag time of approximately 12 hours, which roughly corresponds to the lag time estimated for other plant RNA viruses (Yarwood, 1952a). While this somewhat complex scenario cannot be completely ruled out, it seems unlikely. As a simpler alternative hypothesis, the night-time peak of viral RNA silencing might not generate the susceptibility oscillations, but rather may be a compensatory response for the lowered SA-mediated defence during the night, as described above. The main determinant driving the susceptibility oscillations, therefore, remains an open question. The rhythmic profile of SA-mediated defence could be responsible for

driving the oscillations in TuMV-GFP susceptibility. This could be tested by comparing the temporal regulation of susceptibility between wild type and mutant plants defective in SA-signalling. An alternative hypothesis could posit that the susceptibility oscillations are not generated by any defence pathway but rather are determined by the rhythmic expression of host factors that are required for TuMV-GFP infections. The finding that *eIF(iso)4E,* which is critical for TuMV-GFP infections, exhibits rhythmic expression (Figure 4.8) offers support for this 'host-factor-hypothesis'. This hypothesis requires further investigation however, to confidently accept or reject it. An interesting experiment would be to complement a loss-of-function *eIF(iso)4E* mutant genotype with an eIF(iso)4E coding sequence, driven by a constitutive (non-rhythmic) promoter. A loss of TuMV-GFP susceptibility oscillations across day/night cycles in this genetic background would implicate rhythmic expression of *eIF(iso)4E* as the main determinant for the response. Of course, it is possible that the rhythmic expression of other host factors could also contribute to the response. Another topic worthy of future research would be to determine if similar oscillations in susceptibility occur for infections with different viruses. In addition to addressing how general this phenomenon is, comparison of different viruses with differing host factor requirements could help to assess and inform the 'host-factor-hypothesis'.

In sum, results presented herein clearly show that susceptibility to TuMV-GFP is dependent on the time of inoculation, which is at least partially dependent on the endogenous circadian oscillator. Circadian regulation of viral RNA silencing results in a peak of activity during the night/subjective-night period. The factors responsible for generating the rhythmic pattern of susceptibility are yet to be conclusively identified. The model favoured herein suggests that oscillatory expression of host factors such as *eIF(iso)4E* may drive the rhythms in TuMV-GFP susceptibility and SA-mediated- and vsiRNA-mediated- defences may exhibit alternating peaks of activity during the day and night, respectively, to ensure consistent suppression of TuMV-GFP. Many of the assumptions of this model, however, are subject to further investigation.

Chapter 5: Engineering potyvirus resistance in *Arabidopsis* using CRISPR/Cas9

5.1 Chapter introduction

5.1.1 Eukaryotic translation

In eukaryotes, translation of mRNAs is orchestrated by multi-component translation complexes composed of eukaryotic initiation factors (eIFs) which recruit ribosomes to the 5' UTR. eIF4E and eIF4G associate to form an eIF4F core complex. eIF4G acts as a scaffold protein which associates with the DEAD box RNA helicase eIF4A, and the polyA binding protein (PABP) which unwind and circularise the mRNA, respectively. eIF4E associates with the 5' m^7GpppN cap structure which is crucial for mRNA circularisation and anchoring of the translation complex to the 5' UTR. In higher plants, gene duplication has resulted in a second eIF4F complex called eIF(iso)4F which is composed of eIF(iso)4E and eIF(iso)4G (Browning, 1996) . While eIF4F and eIF(iso)4F are usually formed by cognate pairing of their respective subunits (eIF4E/eIF4G and eIF(iso)4E/eIF(iso)4G) (Bush *et al.*, 2009) functional redundancy exists such that single mutations in one complex can be compensated for by activity of the other (Duprat *et al*., 2002) (Figure 5.1A,B).

5.1.2 Viral usurpation of host translation factors

As previously outlined in *section 1.6.3*, the majority of recessive potyvirus resistance genes map to mutations in *eIF4E* or *eIF(iso)4E*. The molecular basis for this resistance was first elucidated by yeast-two-hybrid screens revealing protein-protein interactions between the potyviral VPg and eIF4E and/or eIF(iso)4E (Leonard *et al*., 2000, Schaad *et al*., 2000, Wittmann *et al*., 1997). Furthermore, mutations in the VPg coding region which abolish its interaction with eIF4E/eIF(iso)4E result in non-viable viruses, thus confirming the necessity of this interaction in the potyvirus lifecycle (Leonard *et al*., 2000). Interestingly, most natural eIF4E/eIF(iso)4E mutations which confer resistance to potyviruses harbour non-conservative amino acid substitutions within two neighbouring regions of the mature protein, proximal to the cap-binding pocket (Robaglia & Caranta, 2006). This strongly suggests that the *eIF4E/eIF(iso)4E*-based resistance is a result of loss of VPg-eIF4E/eIF(iso)4E interaction, rather than other

pleiotropic effects. While the most likely role for the VPg-eIF4E/eIF(iso)4E interaction is to enable translation of the potyviral polyprotein, there is evidence to suggest that eIF4E and eIF(iso)4E may also be involved in other viral processes including replication (Puustinen & Makinen, 2004, Wang *et al*., 2000) and movement (Gao *et al*., 2004). In contrast to the extensive functional redundancy between eIF4E and eIF(iso)4E for translating cellular mRNAs, some potyviruses have evolved to specifically interact with one of the two isoforms. It has been well documented that TuMV specifically utilises the eIF(iso)4E isoform, and consequently *eIF(iso)4E* loss-of-function mutants are resistant to TuMV infection (Duprat *et al*., 2002, Lellis *et al*., 2002, Sato *et al*., 2005) (Figure 5.1C,D). Hence the *Arabidopsis*-TuMV pathosystem can be used as a model to investigate methods to reverse engineer potyvirus resistance using biotechnological methods.

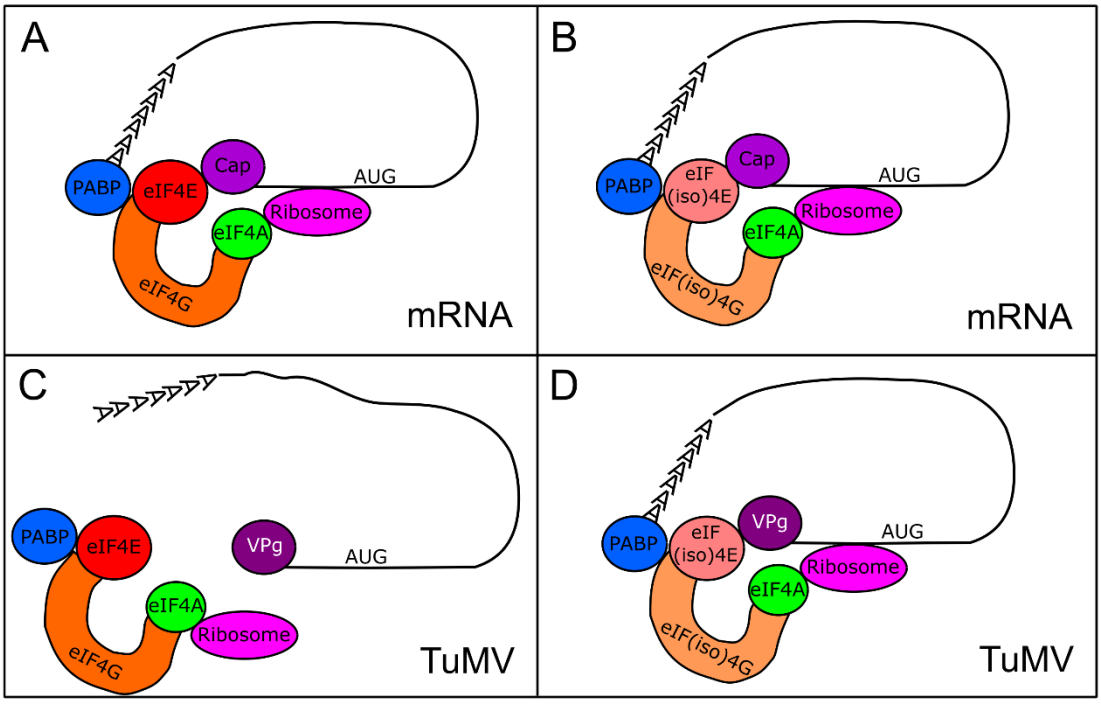

Figure 5.1 Translation of cellular mRNA and the TuMV genome by the eukaryotic translation complex, in *Arabidopsis*. Cellular mRNAs can be translated by complexes containing either eIF4E or eIF(iso)4E (**A**,**B**). The TuMV genome can only utilise complexes containing the eIF(iso)4E paralogue (**D**). If eIF4E is the only functional paralogue present (**C**) the TuMV VPg cannot interact with the eukaryotic translation complex, and hence cannot translate/replicate its genome. PABP = PolyA binding protein.

5.1.3 CRISPR/Cas9 as a tool for genome editing

The field of site specific genome editing was recently revolutionised by the discovery and characterisation of a programmable, RNA guided DNA endonuclease from *Streptococcus pyrogenes* called Cas9 which operates together with CRISPR (clustered regularly interspaced short palindromic repeats) loci in the bacterial genome as a form of defence against invading plasmids or DNA viruses, in bacteria (Jinek *et al.*, 2012). In its natural context, Cas9 associates with a so-called crRNA (CRISPR RNA) derived from the invading DNA, and a tracrRNA (trans-acting CRISPR RNA) which associate with the Cas9 protein. Cas9 is guided to the invading DNA by the crRNA via Watson-Crick base pairing, and it destroys the invading sequences by introducing double stranded breaks (DSBs) at a specific site. A breakthrough in the application of CRISPR/Cas9 came from the discovery that Cas9 can be programmed to introduce DSBs in eukaryotic cells guided by an artificial sgRNA (single guide RNA) which fuses components of the crRNA and tracrRNA (Jinek *et al*., 2012). The only constraint on the design of the sgRNA is that the 20nt region of complementarity between the sgRNA and the target DNA must be immediately upstream of an NGG sequence (where 'N' is any base) known as a PAM (proto-spacer adjacent motif). Introduction of sgRNA-programmed Cas9 into eukaryotic cells leads to Cas9 induced DSBs in the target DNA specifically 3nt upstream of the PAM (Gasiunas *et al.*, 2012, Jinek *et al.*, 2012). As these DSBs are primarily repaired by the error-prone non-homologous end joining (NHEJ) DNA repair pathway, Cas9 induced DSBs will often result in short insertions/deletions (indels) at the site of DNA cleavage. As such, CRISPR/Cas9 technology has opened up a facile means for introducing site specific mutations in eukaryotic genomes. Since its watershed publication, CRISPR/Cas9 technology has been adopted in various model organisms including zebrafish (Hwang *et al.*, 2013), mouse (Wang *et al.*, 2013), rat (Li *et al.*, 2013b), *Arabidopsis* (Feng *et al.*, 2013), and *N. benthamiana* (Nekrasov *et al.*, 2013). Importantly, CRISPR/Cas9 has recently been used to modify genomes of several crop plants (Brooks *et al.*, 2014, Cai *et al.*, 2015, Ito *et al.*, 2015, Jacobs *et al.*, 2015, Jiang *et al.*, 2013, Liang *et al.*, 2014, Shan *et al.*, 2013, Shan *et al.*, 2014, Zhou *et al.*, 2014, Zhu *et al.*, 2016), though most of these early studies were testing the efficacy of the CRISPR/Cas9 in crop species by targeting genes with obvious loss of function phenotypes, rather than functionally relevant loci.

5.2 Chapter aims

As loss of function mutations in components of the eIF4F translation complex have repeatedly been associated with stable resistance to several potyviruses (see chapter 1, table 1.3), the aim of this chapter was to generate virus resistant plants by novel mutation at the *eIF(iso)4E* locus in *Arabidopsis* using CRISPR/Cas9 technology. The rationale for inducing such mutations by CRISPR/Cas9 genome editing was to show-case the concept for generating virus resistance, which can be applied to important crops in the future. As CRISPR/Cas9 has been shown to be a viable technology for site specific genome editing in several plant species, this work could pave the way as a strategy for reverse engineering potyvirus resistance in a wide variety of crops. At the time of publication (Pyott *et al*., 2016) the research presented in this chapter provided one of the first examples of using CRISPR/Cas9 technology in plants to create an agronomically desirable trait – namely virus resistance.

5.3 Site specific mutation of *eIF(iso)4E* by transgenic expression of a sgRNA guided Cas9

The *Arabidopsis eIF(iso)4E* locus (At5G35620) was chosen as a target for CRISPR/Cas9 mutagenesis as mutations caused by ethyl methanesulfonate (EMS)(Lellis *et al*., 2002, Sato *et al*., 2005) and transposon insertion (Duprat *et al*., 2002) at this locus have previously been shown to result in complete resistance to several potyviruses including TuMV (Duprat *et al*., 2002), LMV (Duprat *et al*., 2002), and TEV (Lellis *et al*., 2002). The 5’ region of the ORF was targeted, as mutations here would increase the likelihood of creating non-functional proteins by causing a coding frameshift or early stop codons. Specifically, a sgRNA was designed to target bases +15 – +35 relative to the translation start site of the gene (Figure 5.2). This particular region was selected because it allowed the end of the target region of the sgRNA (immediately upstream of the PAM) to end in a guanine dinucleotide (GG) which has been reported to increase the efficacy of Cas9 induced mutations at the correct target site (Farboud & Meyer, 2015). Additionally, when this sgRNA was queried for possible off-targeting, using the online search tools CRISPR-Plant and CCTop (Stemmer *et al.*, 2015), no off-target loci were detected. For this, search criteria

were used which only qualified off-targets with fewer than 5 mismatching bases to the sgRNA and no more than 2 mismatching bases in the seed region (12 consecutive nucleotides upstream of the PAM), as these parameters have been experimentally validated as the minimum requirements for directing Cas9 to cleave DNA targets (Sternberg *et al.*, 2015, Hsu *et al.*, 2013). This sgRNA (Figure 5.2) was cloned into a binary vector (pDe-CAS9) (Fauser *et al.*, 2014) which allowed for the tandem expression of a plant codon optimised Cas9 and the sgRNA (driven by *PcUbi4-2* and *AtU6-26* promoters, respectively) along with the BASTA resistance gene. This construct was introduced into *Arabidopsis* (Col-0 accession) plants by *Agrobacterium* mediated transformation (floral dipping).

Transgenic T_1 seeds were BASTA selected, and 6 healthy looking plants were chosen to test for genome editing at the *eIF(iso)4E* locus. For this an assay was adopted which allows edited sequences to be detected on a gel after incubation with T7 endonuclease (hereafter T7). Briefly, primers spanning the target site (Figure 5.2) were used to generate an *eIF(iso)4E* amplicon, which was subsequently denatured and re-annealed before incubation with T7. Bulges caused by mismatching bases are recognised by T7 which results in cleavage of the re-annealed amplicons. Hence, samples with a mixture of wild type and mutant DNA will yield cleavage products which can be resolved on a gel, whereas homogenous samples of either wild type or fully mutated DNA will be resistant to T7 activity; yielding a full, un-cleaved target amplicon. A possible drawback of this assay is that samples with identical lesions in all DNA copies will be indistinguishable from wild type, and hence such samples would be missed in a screen for detecting mutations. However, it was reasoned that this would be a highly unlikely scenario in the T_1 generation as Cas9 expression could only have been present in one of the parental gametes and as such, genome editing at this stage would most likely result in chimeric plants containing a mixture of wild type and mutant sequences. The T7 assay was performed on *eIF(iso)4E* amplicons from 6 BASTA selected T_1 transformants alongside a wild type, non-transformed control and a T_1 plant transformed with a Cas9 vector lacking a sgRNA sequence (Figure 5.3). T7 cleavage products were absent from both negative controls but were clearly visible in candidate plants numbered 1,3, and 6 (Figure 5.3C). As the signal for the T7 cleavage product was strongest in sample number 1, this line was chosen for further analysis.

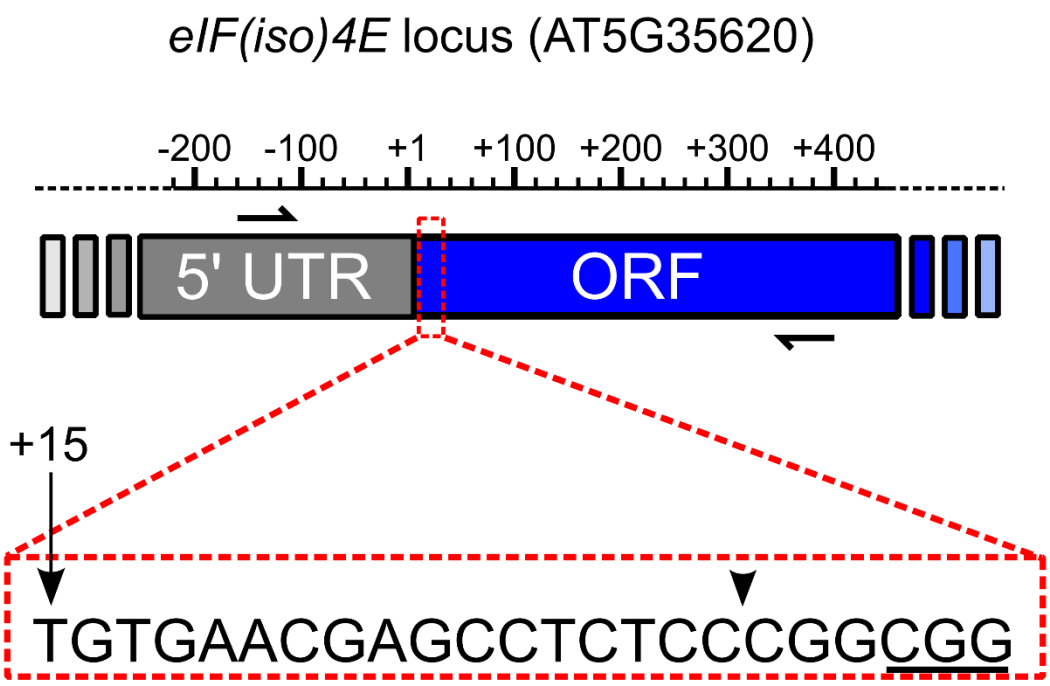

Figure 5.2 Schematic of the *eIF(iso)4E* locus targeted for editing by CRISPR/Cas9. Primers flanking the target site are shown by half arrows over the 5' untranslated region (UTR) and open reading frame (ORF) shaded grey and blue, respectively. Scale bar depicts the approximate positions, in base pairs, relative to the translation start site (+1). The enlarged area in red dashes shows the position and sequence of the sgRNA, with the protospacer adjacent motif (PAM) underlined. The position of the sgRNA relative to the +1 of the ORF is indicated by an arrow. The Cas9 cleavage site is indicated by an arrow head.

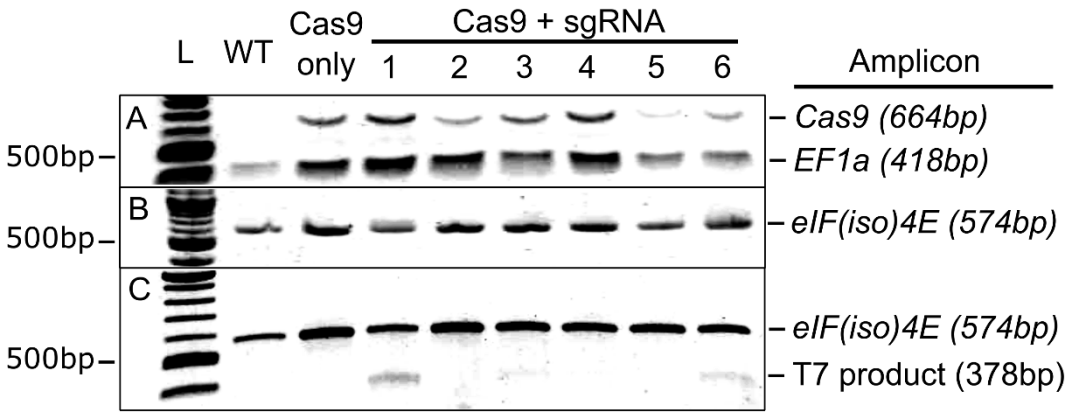

Figure 5.3 PCR/T7 endonuclease products for six independently tranformed plants (T_1 generation) containing a Cas9 transgene with a sgRNA targetting the eIF(iso)4E locus. Transformation with a *Cas9* transgene with no sgRNA (Cas9 only) and a non-transformed, wild type (WT) plant were used for controls. L denotes a 100bp DNA ladder. **A:** 1% agarose gel showing multiplex PCR products confirming the presence/absence of the *Cas9* transgene, using the constitutively expressed house-keeping gene *EF1a* as a loading control. **B:** 1 % agarose gel showing PCR amplicons spanning the putative mutation site at the *eIF(iso)4E* locus. **C:** 2% agarose gel showing *eIF(iso)4E* cleavage products after a self-annealing reaction of the PCR amplicon, and subsequent digestion with T7 endonuclease. Presence of the 3'-end cleavage product (378bp) in samples 1, 3, and 6 is indicative of Cas9-induced mutation at the *eIF(iso)4E* locus. The cleavage product corresponding to the 5'end (196bp) is not visible on the gel.

5.4 Segregation of the induced mutation from the transgene in the T_2 generation

T_2 seeds were produced by allowing the selected T_1 plant (line number 1) to self-pollinate. It was reasoned that it would be possible to segregate stable, uniform *eIF(iso)4E* mutations from the transgene at this generation. To establish the number of integration sites for the transgene in this line, approximately 200 of these T_2 seedlings from the T_1 line number 1 were sprayed with BASTA. The BASTA resistant:susceptible ratio was approximately 3:1, indicative of a single integration site of the transgene. In parallel, with a separate batch of segregating T_2 seedlings, a multiplex PCR screen was employed to identify T_2 plants which had lost the

Cas9/sgRNA expressing transgene (Figure 5.4). 144 plants were tested in this way and 54 non-transgenic plants were recovered. The general strategy to segregate the transgene from the induced mutation is outlined in Figure 5.5.

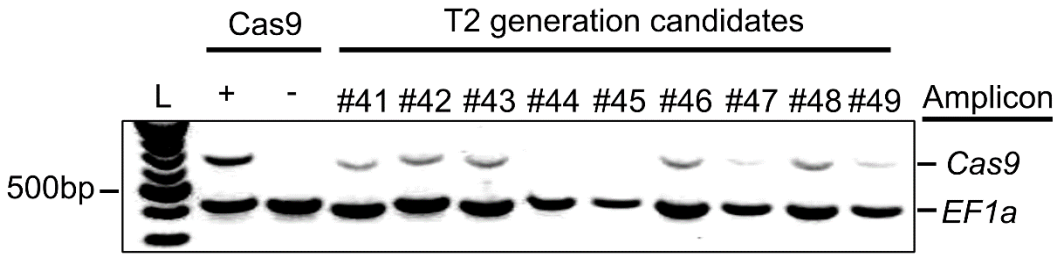

Figure 5.4 Representative 1% agarose gel for the selection of T_2 candidates lacking the *Cas9* transgene. Multiplex PCR was used to confirm the presence/absence of the *Cas9* transgene, using the constitutively expressed house-keeping gene *EF1a* as a loading control. L denotes a 100bp DNA ladder. A Cas9 transformant (T_1 generation) and a non-transformed wild type plant were used as positive and negative controls for *Cas9* amplification, respectively. Samples #41 - #49 are a representative selection of T_2 progeny from T_1 plant number 1 (as shown in Figure 5.4) Candidates #44 and #45 represent two out of a total of 54 candidates lacking the *Cas9* transgene, which were selected by this method.

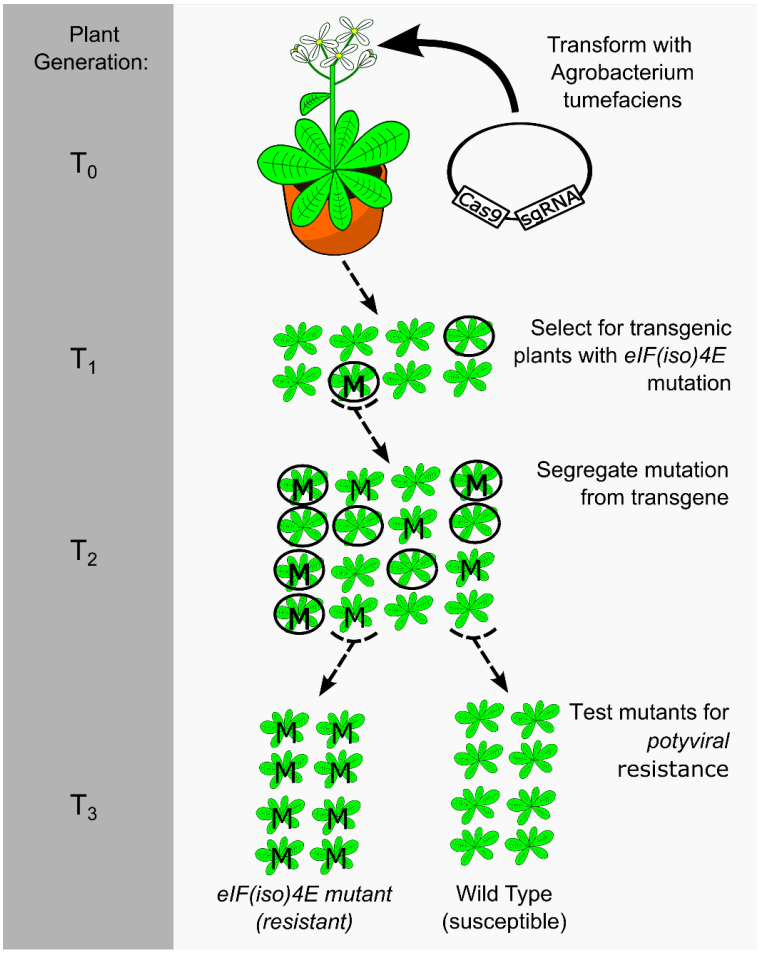

Figure 5.5 Schematic of project work-flow. *Arabidopsis* Col-0 plants were transformed by floral dipping with the *Cas9/sgRNA* recombinant binary vector containing a BASTA resistance gene (pDe-Ubi:SpCas9-U6:sgRNA_AteIF(iso)4E). After self-pollination, the seeds were collected and germinated in soil. Plants carrying the transgenic construct (black circle) were selected in the T_1 generation by spraying with BASTA. The transgenic plants were then tested for *eIF(iso)4E* mutation (M) to identify plants with an active CRISPR/Cas9 nuclease using a T7 assay. One transgenic T_1 plant with clear signs of *eIF(iso)4E* editing was used to produce the T_2 generation. PCR was used to identify T_2 generation plants which had lost the transgene by Mendelian segregation. The non-transgenic T_2 plants were then screened for *eIF(iso)4*E mutations by Sanger sequencing (please note: the mutations depicted as 'M' in the diagram are not identical, as the mutations in the T_1 generation occurred in somatic cells so were not heritable, and different mutations were recovered in the T_2 generation due to independent editing events in the germline of T_1. For simplicity 'M' was used to depict all CRISPR/Cas9 induced mutations). Non-transgenic T_2 plants, which were homozygous for either the mutated or wild type *eIF(iso)4e* alleles, were used to produce T_3 populations, which were then tested for viral resistance.

Next, this population of transgene-free plants was screened for *eIF(iso)4E* mutations. As homozygous mutants could potentially be identified at this stage, the T7 assay was inappropriate for this analysis as such mutants would be undetected. Instead, mutations were screened by directly sequencing *eIF(iso)4E* amplicons for each of the 54 lines by Sanger sequencing. An online software, CRISP-ID (Dehairs *et al*., 2016) was used to parse the mixed sequence base calling which occurs downstream of heterozygous or biallelic (where both alleles carry different mutations) editing. Alignment of the parsed sequence files enabled the different mutation types to be scored (Figure 5.6A). Of the 54 non-transgenic plants, 32 (59%) harboured mutations in *eIF(iso)4E* of which 13 (24%) were heterozygous, 14 (26%) were biallelic, and 5 (9%) were homozygous (Figure 5.6B). Interestingly the majority of all mutations were single nucleotide indels, with the exception of one heterozygous mutant which had a 57bp deletion, spanning the predicted Cas9 cleavage site (Figure 5.7). This over-representation of single nucleotide indels is consistent with previous reports for Cas9 induced editing in plants (Nekrasov *et al*., 2013). Importantly, all of the indels aligned perfectly with the expected Cas9 cleavage site, 3bp upstream of the PAM (Figure 5.6A). Furthermore, mutations of interest were confirmed by resequencing the *eIF(iso)4E* amplicons from the opposite direction to rule out the unlikely possibility of false positives due to sequencing errors. Of the 5 homozygous mutants, 2 (#65 and #66) were genetically identical, and hence only one of these was selected for further analysis. Each of the homozygous mutants have early stop codons in their predicted amino acid sequences (Figure 5.8), which would most likely create complete functional knock-outs by severe truncation of the eIF(iso)4E protein. 4 of the homozygous mutants (named #44, #65, #68 and #98) were self-pollinated to produce T_3 populations for each of the different *eIF(iso)4E* point mutations. A non-transgenic T_2 plant with homozygous wild type *eIF(iso)4E* alleles (#105) was also selected to produce a wild type T_3 population.

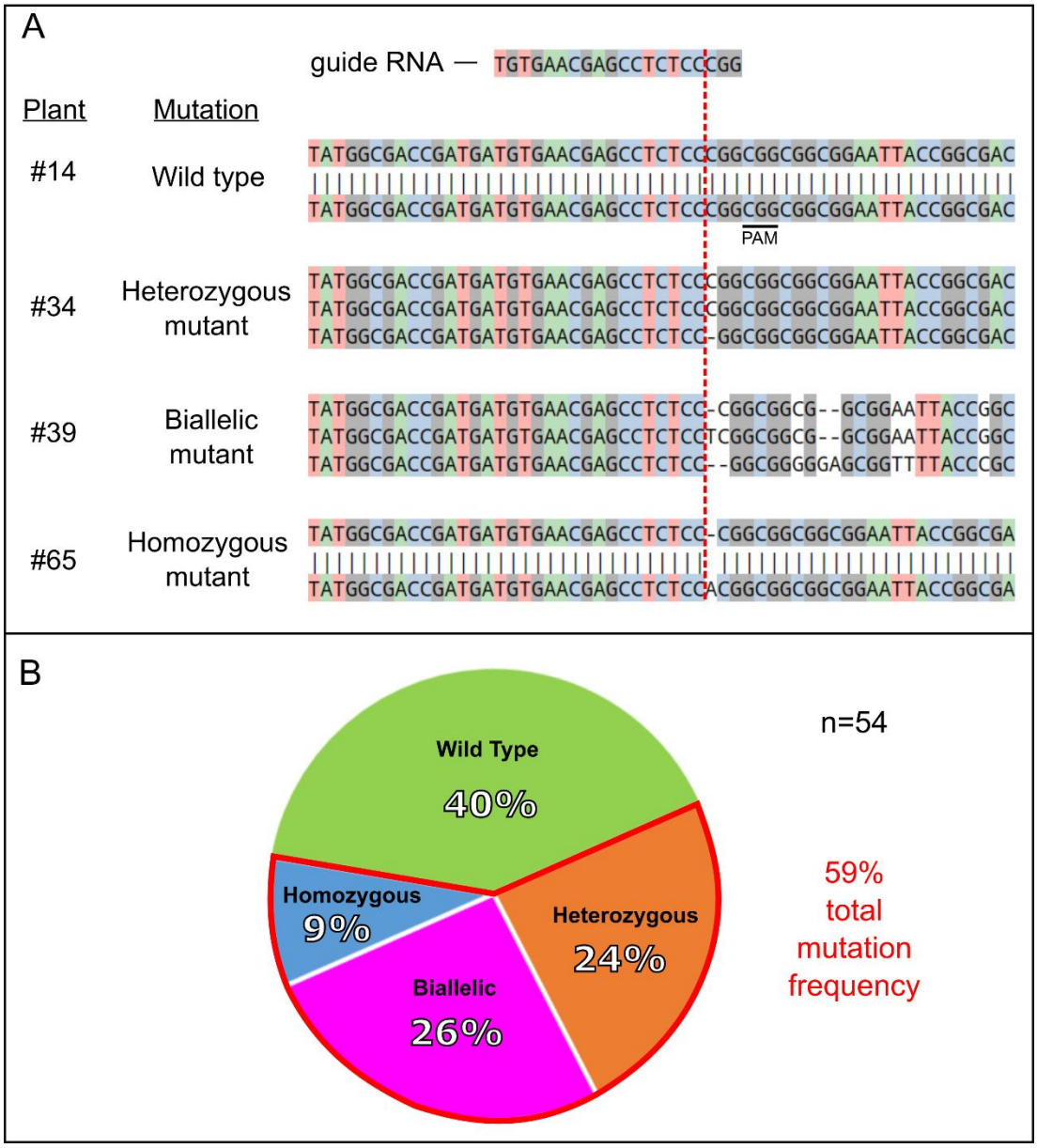

Figure 5.6 Summary of *eIF(iso)4E* mutations in the T_2 generation. **A** Sequence alignments using CRISP-ID online software. Guide RNA sequence is indicated, with PAM underlined. Red dashed line indicates the predicted Cas9 cleavage site. A representative alignment for each mutation type is shown. **B** Pie chart summary of the *eIF(iso)4E* mutations recovered from 54 T_2 plants.

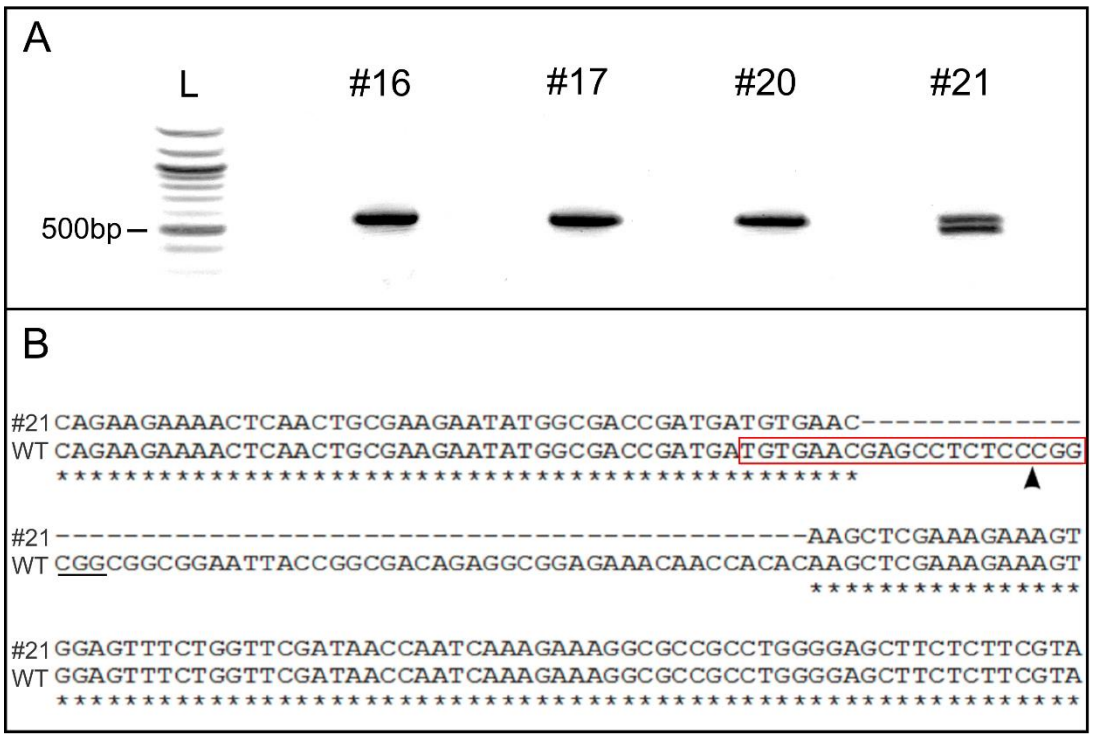

Figure 5.7 Sequence alignment for plant #21 showing a 57bp deletion in one of the two alleles. **A:** PCR spanning the predicted mutation site at *eIF(iso)4E*. A shorter amplification product can be resolved on a 1% agarose gel for plant #21. L denotes a 100bp DNA ladder. **B:** Alignment of the mutant allele to the Wild Type (WT) sequence. The lower band was gel isolated and cloned into a vector (pGEMT-EASY®) for sequence analysis. Amplicons from plants #16, #17, and #20 are all homozygous wild type. Red box indicates the target sequence of the guide RNA, with the PAM sequence underlined.

A

Sample	Nucleotide Sequence	Mutation
WT	TGTGAACGAGCCTCTCC-CGGCGG	
#44	TGTGAACGAGCCTCTCC--GGCGG	Deletion (-C)
#65	TGTGAACGAGCCTCTCCACGGCGG	Insertion (+A)
#68	TGTGAACGAGCCTCTCCTCGGCGG	Insertion (+T)
#98	TGTGAACGAGCCTCTCCCCGGCGG	Insertion (+C)

B

Sample	Amino Acid Sequence
WT	1 MATDDVNEPLPAAAELPATEAEKQPHK.LERKWSFWFDNQSKKGAAWGASLRKAYTFDTV
#44	1 MATDDVNEPLRRRRNYRRQRRRNNHFSSKESGVSGSITNQRKA..PPGELLFVKPILSTP
#65	1 MATDDVNEPLHGGGGITGDRGGETTTQ.ARKKVEFLVR*~~~~~~~~~~~~~~~~~~~~
#68	1 MATDDVNEPLLGGGGITGDRGGETTTQ.ARKKVEFLVR*~~~~~~~~~~~~~~~~~~~~
#98	1 MATDDVNEPLPGGGGITGDRGGETTTQ.ARKKVEFLVR*~~~~~~~~~~~~~~~~~~~~
WT	60 EDFWGLHETIFQTSKLTANAEIHLFKAGVEPKWEDPECANGGKWTWVVTANRKEALDKGW
#44	59 SKIFGDCTRLY.....FRLAN*~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
#65	38 ~~~~~~
#68	38 ~~~~~~
#98	38 ~~~~~~
WT	120 LETLMALIGEQFDEADEICGVVASVRPQSKQDKLSLWTRTKSNEAVLMGIGKKWKEILDV
#44	75 ~~~~~~
#65	38 ~~~~~~
#68	38 ~~~~~~
#98	38 ~~~~~~
WT	180 TDKITFNNHDDSRRSRFTV*
#44	75 ~~~~~~~~~~~~~~~~~~~
#65	38 ~~~~~~~~~~~~~~~~~~~
#68	38 ~~~~~~~~~~~~~~~~~~~
#98	38 ~~~~~~~~~~~~~~~~~~~

Figure 5.8 Summary of CRISPR induced *eIF(iso)4E* homozygous mutations. **A:** DNA sequence alignments for four of the homozygous *eIF(iso)4E* mutants (#44, #65, #68, #98) identified in the T_2 generation, along with a Wild Type (WT) control. Lines #65, #68, and #98 exhibit single nucleotide insertions while line #44 has a single nucleotide deletion. **B:** Predicted amino acid sequence alignments for the four homozygous mutants and the wild type consensus. Each of the mutant alleles code for severely truncated and disrupted proteins.

5.5 The induced mutation in *eIF(iso)4E* confers complete resistance to TuMV

The 4 separate T_3 populations (#44, #65, #68, #98) harbouring each of the induced *eIF(iso)4E* mutations were grown alongside the T_3 wild type control and a previously published transposon insertion *eIF(iso)4E* mutant, which is known to be resistant to TuMV infection (Duprat *et al*., 2002). After 4 weeks of growth, 40 plants from each of the 6 different genotypes (#44, #65, #68, #98, #105, and the transposon insertion mutants) were rub-inoculated with TuMV-GFP. At 7 and 14 dpi, TuMV infection was assessed by monitoring the expression of GFP in inoculated and systemic leaves (Figure 5.9). GFP expression (indicative of TuMV-GFP infection) was clearly visible at 14dpi in 37/40 (92.5%) of the wild type plants but not in any of the CRISPR-induced mutants (#44, #65, #68, and #98), nor the previously reported transposon insertion mutant (Figure 5.9A). To ascertain that the lack of GFP signal in the *eIF(iso)4E* mutants was due to complete TuMV-GFP resistance, and not just lower viral RNA, inoculated and systemic leaves for each genotype were analysed by RT-PCR amplifying the coat protein coding region of TuMV-GFP. TuMV-specific amplicons were clearly detected in both inoculated and systemic leaves from 3 wild type plants, but no TuMV-specific amplicons were detected in any samples from any of the *eIF(iso)4E* mutants (Figure 5.9B). To gain a more quantitative insight into the viral RNA in the inoculated plants, qRT-PCR was performed for the inoculated and systemic leaves at 7dpi (Figure 5.9C). A serial dilution of purified TuMV-GFP RNA was included in the qRT-PCR to construct a standard curve (Figure 5.10) from which the absolute RNA (in pg) of TuMV-GFP in the 7dpi samples could be interpolated. In agreement with the RT-PCR results (Figure 5.9B) TuMV-GFP was detected in all of the wild type samples but not in systemic leaves of the *eIF(iso)4E* mutants. Interestingly, a very low concentration of TuMV-GFP RNA (1-10pg) was detected in the inoculated leaves of the *eIF(iso)4E* mutants (Figure 5.9C, Figure 5.10A). However, it is very likely that this TuMV-GFP signal is due to residual infectious sap from the rub-inoculation, though it is also possible that a low level of TuMV replication occurs in the inoculated leaves of *eIF(iso)4E* mutants.

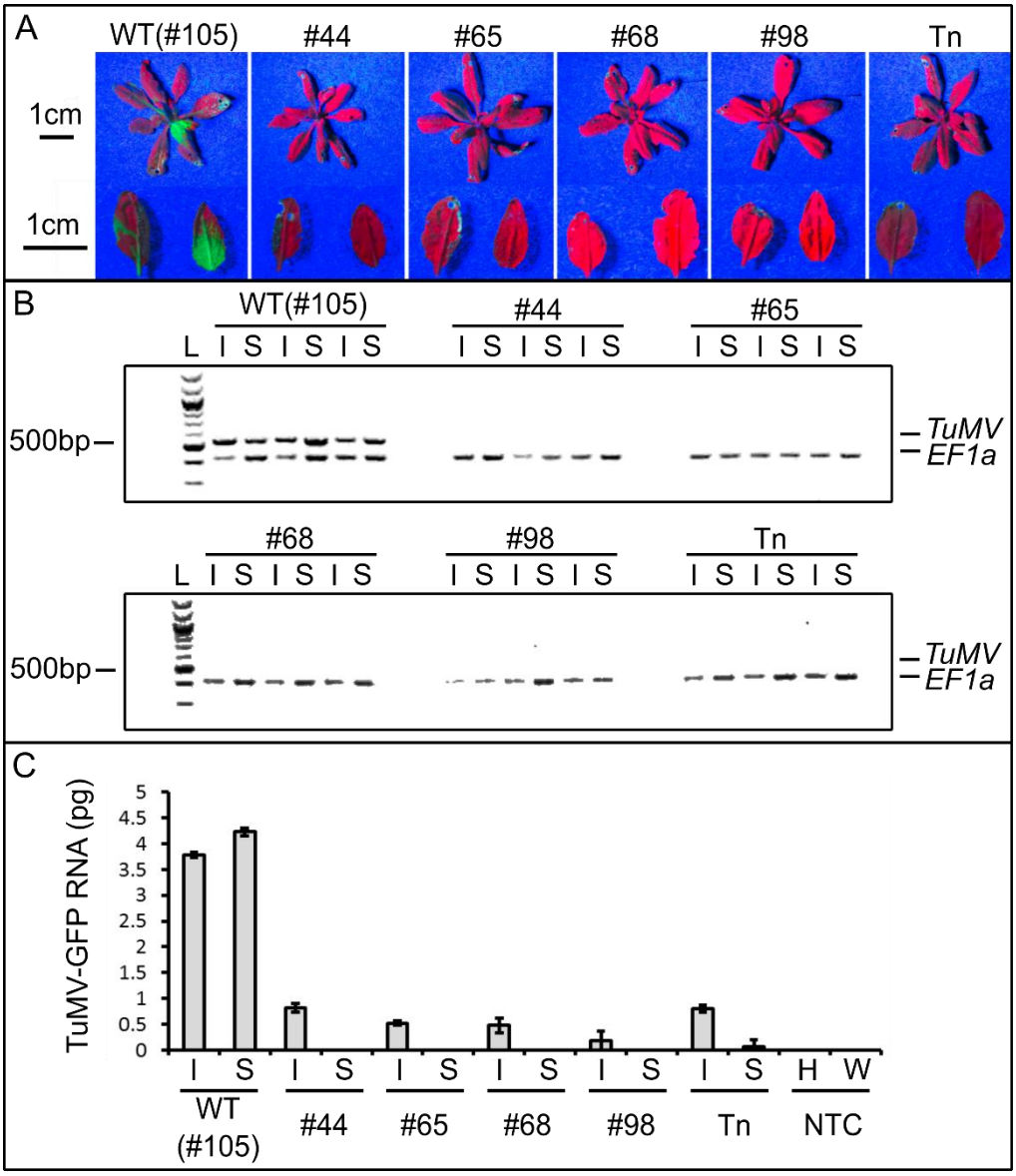

Figure 5.9 A: Representative photos of TuMV-GFP infected plants, imaged under UV light at 7 dpi. A transposon induced *eIF(iso)4E* mutant (Tn) was used as a resistant control. An enlarged image of inoculated (left) and systemic (right) leaves are shown below each rosette. **B:** RT-PCR to detect the presence of TuMV-GFP in inoculated (I) and systemic (S) leaves for each genotype. Three separate plants were analysed per genotype. The first two lanes for each genotype correspond to the leaves imaged in panel **A**. Amplicons of the TuMV coat protein region (537bp) and the housekeeping gene *EF1a* (418bp) were PCR amplified separately from the same cDNA, then mixed and run together on a 2% agarose gel. L denotes a 100bp DNA ladder. **C:** qRT-PCR to detect the mean absolute viral RNA (in pg) for the samples shown in panels A and B. qRT-PCR reactions were performed with cDNA from a healthy plant (H) and water (W) as negative controls (NTC). Error bars show SEM of 3 biological replicates.

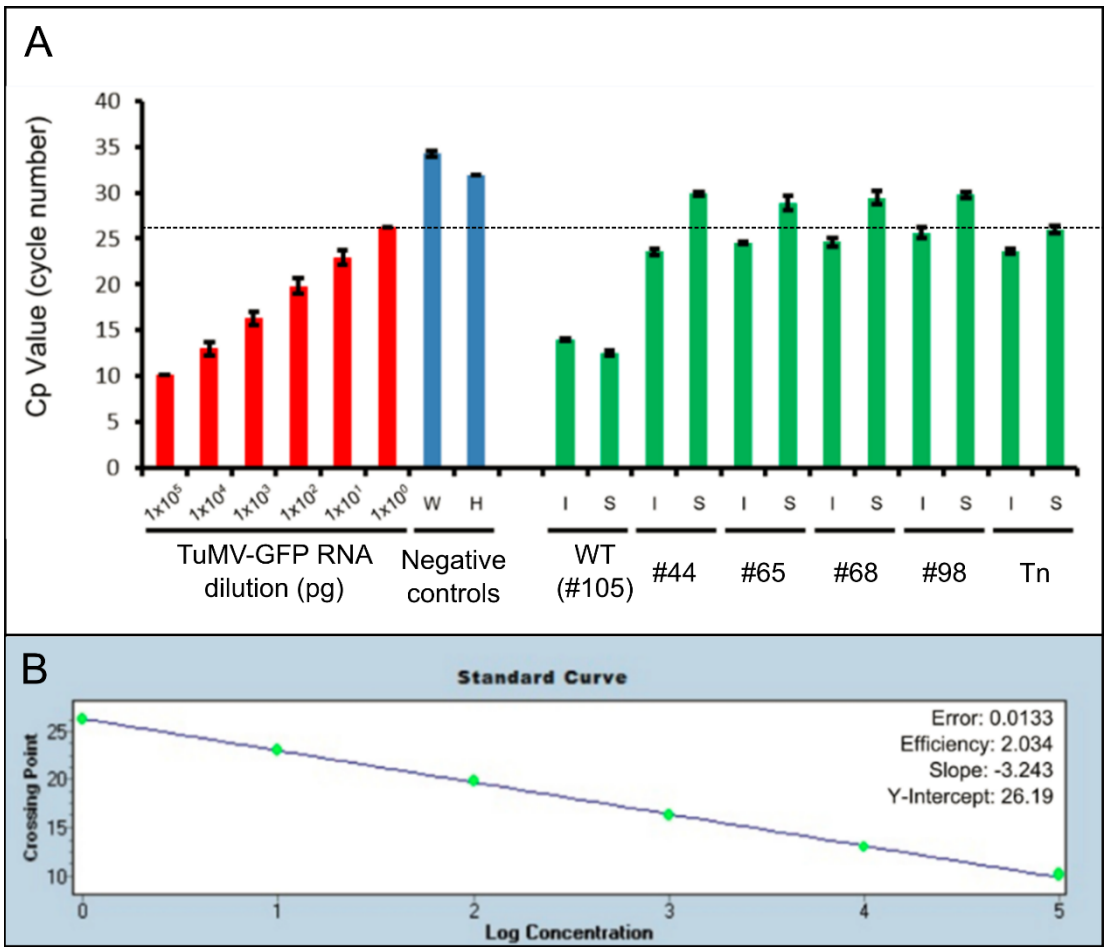

Figure 5.10 A: Cp values for qRT-PCR quantification of TuMV-GFP in inoculated samples, as described in Figure 5.8. Red columns show the mean Cp values for cDNA made from a serial dilution of gel purified viral RNA, which was used to calibrate the absolute quantification of viral RNA shown in Figure 5.8. Blue columns show the mean Cp values for negative controls, using water (W) and cDNA from a healthy, non-inoculated plant (H). Green columns show the mean Cp values for inoculated (I) and systemic (S) leaves taken from TuMV-GFP inoculated plants. Error bars for red and blue columns depict the SEM of technical replicates. Error bars for green columns depict the SEM of 3 biological replicates. Dashed line indicates the detection threshold for 1pg of TuMV-GFP RNA. **B:** Standard curve used to interpolate the absolute concentration of TuMV-GFP from measured Cp values, shown in Figure 5.8. The curve was plotted using LightCyler480 software, using the Cp values for the dilution series shown by the red columns in panel **A**.

To stringently test the TuMV resistance of the *eIF(iso)4E* mutants, 20 systemic leaves for each genotype were collected at 28dpi and pooled to make sap for back-inoculating *N. benthamiana* plants, which are highly susceptible to TuMV and hence would reveal the presence of very low TuMV-GFP titres present in the sap. 5 days after the back-inoculations, TuMV-GFP infection was clearly visible in both local and systemic leaves inoculated with sap taken from wild type *Arabidopsis* plants. In contrast, TuMV-GFP was not detected at all in any of the back-inoculations using sap prepared from the *eIF(iso)4E* mutants (Figure 5.11). This indicates that the ablation of the eIF(iso)4E protein by the induced single nucleotide mutations in lines #44, #65, #68 and #98 render the plants completely resistant to TuMV-GFP in a similar way to the previously published transposon mutant (Duprat *et al*., 2002).

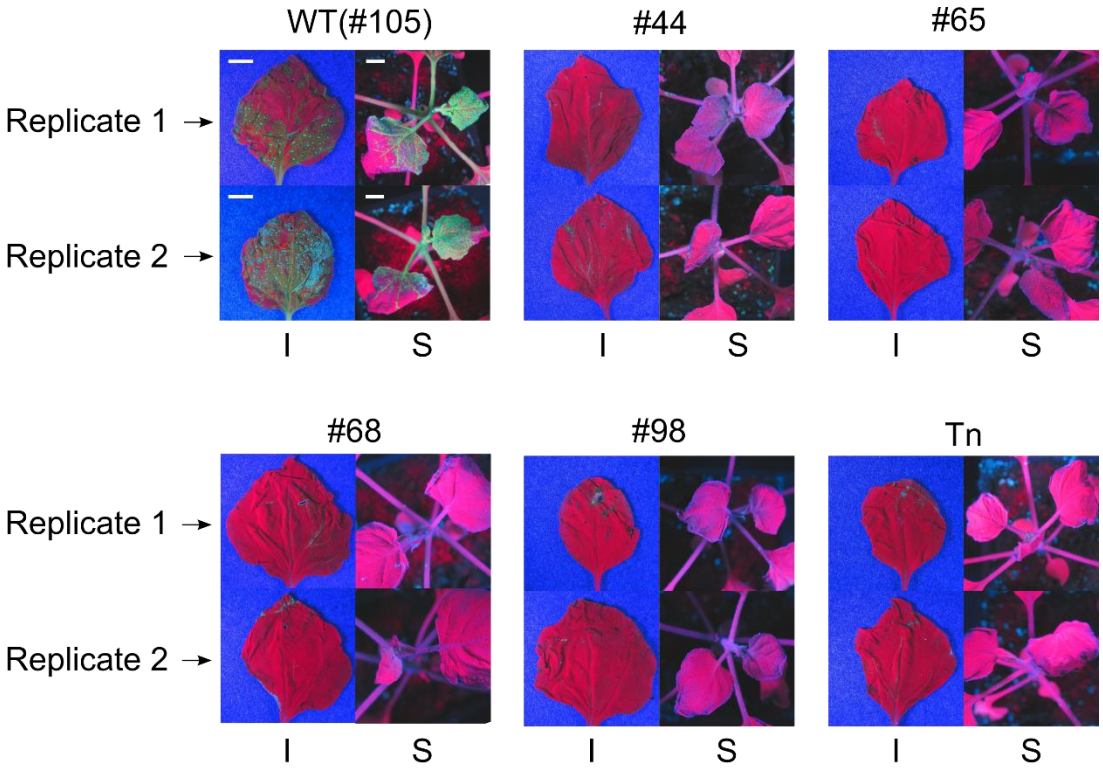

Figure 5.11 Back-inoculations of *N. benthamiana* plants using sap from TuMV-GFP inoculated *Arabidopsis*. Sap was prepared by pooling 20 systemic leaves from TuMV-GFP inoculated *Arabidopsis* plants as shown in Supplemental Figure 3. Labels above each quadrant refer to the genotype [Wild Type (WT), #44, #65, #68, #98 and Transposon(Tn)] of the inoculated *Arabidopsis* used to make sap. Each quadrant shows an inoculated leaf (I) and systemic tissue (S) for two replicate plants, imaged under UV light. A 1cm scale-bar is included in the first quadrant.

5.6 The *eIF(iso)4E* mutants show no growth defects compared to wild type plants, when grown under standard growth conditions

The *eIF(iso)4E* mutants looked indistinguishable from the wild type plants as they grew. To more accurately assess the growth vigour of the *eIF(iso)4E* mutants in comparison to wild type plants, measurements were made for the dry mass and flowering times for non-infected populations of each of the 4 homozygous *eIF(iso)4E* mutant lines (#44, #65, #68, #98) and a wild type line in the T_3 generation. For both the dry weight and flowering time experiments, seeds of each genotype were grown in a randomised block design. 10 plants from each genotype were grown together in a tray, with 6 replicate trays (resulting in 60 plants per genotype, spread across 6 different trays). The relative position of the different genotypes within any tray was randomised. These measures were taken to avoid systematic biases of the growth environment affecting one genotype more than another and hence confounding the analysis. After 4 weeks of growth, just prior to the onset of flowering, 30 plants from each set were randomly sampled to estimate the dry weights of each genotype. The remaining plants were scored for flowering emergence by counting the number of days from germination to the first appearance of a floral bolt. From these experiments, it can be concluded that there are no statistically significant differences (assessed by one-way ANOVA) in the growth of *eIF(iso)4E* mutants and wild type plants, with respect to total dry mass ($F_{4,70} = 1.372$, $p=0.252$) or flowering time ($F_{4,119} = 1.597$, $p=0.180$) (Figure 5.12). From this, it is reasonable to surmise that this strategy of site specific disruption of *eIF(iso)4E* will be useful for generating virus resistant crops without concomitant constraints on plant growth.

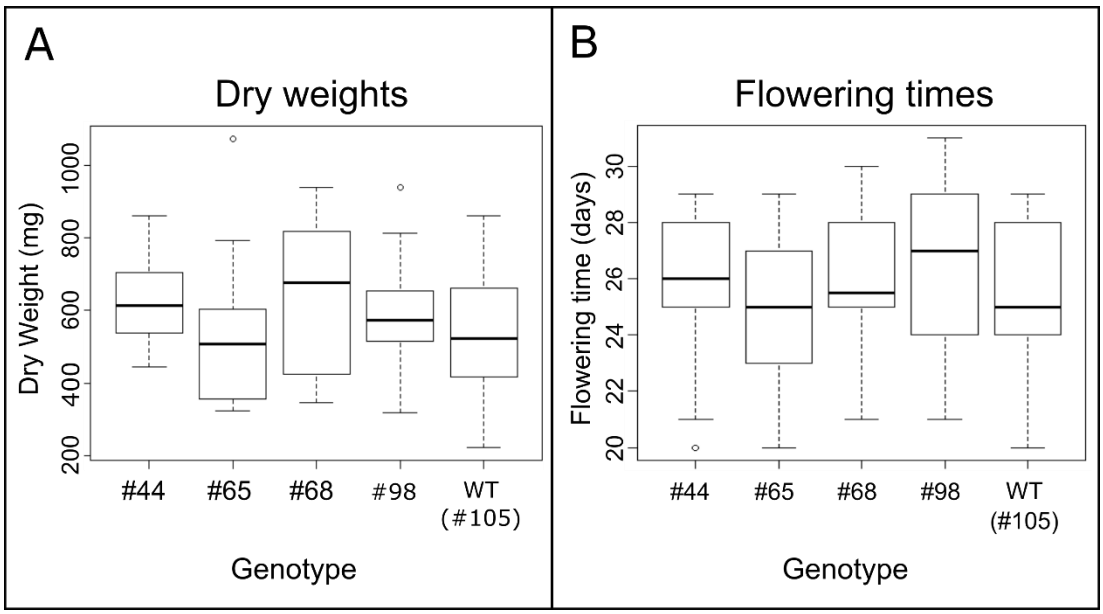

Figure 5.12 Box plots of dry weights (**A**) and flowering times (**B**) for the CRISPR/Cas9 edited *eIF(iso)4E* mutants (lines #44, #65, #68 and #98) alongside Wild Type (WT) plant .

5.7 Chapter discussion

5.7.1 Introducing R genes into crops by plant breeding

Arguably, utilising genetic resistance in crops is the most sustainable approach for controlling virus infections – other methods such as pesticides to control insect vectors or manual inspection and removal of infected plants are costly, laborious, and often ineffective. There are several approaches for introducing *R genes* into crops. Firstly, classical breeding can be used to identify and introgress *R genes* from related species to the crop of interest. Often wild relatives of the domesticated crop are used for these breeding programmes and successive back-crosses are required to segregate out additional non-beneficial (or even deleterious) alleles from the hybrid. Hence breeding programmes can be very costly and laborious, and take many years to achieve the desired elite cultivar. Breeding programmes can be additionally limited by constraints such as hybrid incompatibility or outbreeding depression.

5.7.2 Introducing R genes into crops using biotechnological approaches

Biotechnological approaches can also be adopted to generate *R genes* in crops. A popular platform for introducing and selecting beneficial mutations in crops is TILLING (targeting induced local lesions in genomes). This relies on non-site-specific mutation of the genome of interest, usually by EMS treatment (McCallum *et al.*, 2000). The induced mutations can be subsequently identified and screened for positive attributes, such as virus resistance (Piron *et al.*, 2010). This strategy overcomes some of the pitfalls of classical breeding as it generates novel genetic diversity within the species of interest. However, this approach is also costly, time consuming, and slow to produce elite cultivars, particularly for species with polyploid genomes. Furthermore, off-target mutations due to the random nature of the mutagenesis can result in unintended deleterious effects in the new cultivars. A second biotechnological approach to generate viral resistance is the stable integration of transgenes into crops. This transgenic approach has been utilised in a variety of ways including; overexpression of dominant *R genes* identified in other species (Oldroyd & Staskawicz, 1998), overexpression of non-coding sequences to direct RNA silencing of specific viruses (Asad *et al.*, 2003, Bonfim *et al.*, 2007, Hameed *et al.*, 2017, Pooggin *et al.*, 2003, Scorza *et al.*, 2001, Yang *et al.*, 2004b), overexpression of null alleles to sequester functional host susceptibility factors in a dominant negative manner (Cavatorta *et al.*, 2011), and stable expression of a *Cas9/sgRNA* cassette to target DNA viruses directly for CRISPR/Cas9 mediated cleavage (Ali *et al.*, 2015). While these methods can be effective for introducing viral resistance into crops, transformation of plants with transgenes can be costly and laborious and limited to only certain species or varieties with well-established transformation protocols. Moreover, many governments are opposed to the use of transgenic products which seriously undermines the application of these technologies for improving crop yields in the near future. More recently, site-specific genome editing technologies have offered new ways to improve crops. One attractive feature of these technologies is that, once the desired genome alterations have been made, the transgenes can be crossed out from the improved variety, thus circumventing public and political concerns around the use of persistent transgenes in crops. Early genome editing technologies include TALENs (transcription activator-like effector nucleases) and ZFNs (zinc

finger nucleases) (Gaj *et al.*, 2013). Both TALENS and ZFNs combine DNA nucleases (such as FokI) with a DNA binding protein to induce DNA DSBs at a specific site. A major limitation of these technologies is that tailoring the DNA binding proteins to target a sequence of interest can be costly and time-consuming. Furthermore, it is only possible to engineer DNA binding proteins for certain DNA target sequences. The advent of CRISPR technology has revolutionised the field of genome editing. Crucially, the fact that the Cas9 nuclease is guided by RNA rather than protein overcomes the major limitations of TALEN and ZFN technologies. RNA-based guiding is cheaper and easier to engineer and the range of possible target sequences is greatly expanded, requiring only the commonly occurring NGG PAM sequence.

5.7.3 The potential for CRISPR/Cas9 technology for generating virus resistance in crops

The results presented in this chapter demonstrate the utility of CRISPR/Cas9 technology for generating novel genetic resistance to TuMV in *Arabidopsis* by deletion of a host factor (eIF(iso)4E) which is strictly required for viral survival. It is anticipated that this work will pave the way for a similar strategy to be adopted in important crop species and thus provide an alternative, novel strategy for the introduction of *R genes*. Indeed, since the publication of this chapter (Pyott *et al*., 2016) another publication emerged utilising the same strategy in cucumber (*Cucumis sativus*) to generate resistance to several viruses within the family *Potyviridae* (Chandrasekaran *et al.*, 2016). Credence for this approach is given by the fact that many natural sources of potyvirus resistance rely on the same principle: loss of function mutations in host eIFs (Duan *et al.*, 2012, Kanyuka *et al.*, 2005, Gao *et al.*, 2004, Naderpour *et al.*, 2010, Nieto *et al.*, 2006, Nicaise *et al*., 2003, Ruffel *et al*., 2002, Stein *et al*., 2005). Therefore, it would be hard to justify objections to commercial application of such a strategy, as the final genome-edited product is essentially no different to varieties carrying mutant alleles arisen from 'natural' methods of mutagenesis. It is noteworthy that the engineered viral resistance reported here is the result of a single nucleotide point mutation arisen from the plant's own natural DNA damage repair mechanism, namely NHEJ. Moreover, while the approach in this study was to use transgenic delivery of the *CRISPR/Cas9* cassette, it was shown that it is feasible to segregate out the transgene from the induced mutation at the target

eIF(iso)4E locus at an early stage to produce stable, heritable point mutations without a persistent transgene. Furthermore, alternative strategies could be employed to avoid the use of transgenic delivery of the CRISPR/Cas9 complex. For instance, biolistic delivery of purified Cas9-sgRNA ribonucleoprotein (RNP) complexes has been shown to be a viable strategy for transgene-free genome editing in various plant species (Svitashev *et al.*, 2016, Woo *et al.*, 2015). Encouragingly, the US Department of Agriculture (USDA) has recently approved the use of CRISPR/Cas9-edited mushrooms and maize (reviewed in (Globus & Qimron, 2017)). However, it remains to be seen whether other governing bodies will choose to regulate CRISPR-modified crops under similar criteria.

The fact that CRISPR/Cas9 induces sequence specific mutations, and that there were no predicted off-targets in this study, means that it can be concluded with a high degree of certainty that the genetic differences between the wild type and mutant plants are solely at the *eIF(iso)4E* locus. This results in a better system for investigating the effects of the mutation on plant growth, as previous studies using EMS (Lellis *et al.*, 2002, Sato *et al.*, 2005) or transposon insertion (Duprat *et al*., 2002) mutagenesis are likely to be confounded by multiple off-target genomic mutations. The results presented herein indicate that there are no significant differences in the growth and development of the *eIF(iso)4E* mutants compared to wild type plants. It is still possible that under certain growth environments, particularly stress conditions, the *eIF(iso)4E* mutants may grow differently to wild type plants, though this is entirely speculative and requires further investigation. Recently, a comparison of *eIF4E* resistance alleles between cultivated and wild pepper varieties revealed that these loci were under positive selection in cultivated but not wild scenarios, suggesting a fitness cost to the mutation under natural environments (Poulicard *et al.*, 2016). However, these effects may be stronger for mutations within *eIF4E* compared to *eIF(iso)4E,* based on the fact that growth defects are frequently associated with the former but not the latter (Duprat *et al.*, 2002, Martínez-Silva *et al.*, 2012, Lellis *et al*., 2002, Sato *et al*., 2005).

The durability of this engineered resistance also remains to be tested. It is assumed that recessive resistance arising from the loss of a host factor required by the virus will be more durable than dominant *R genes,* due to lower selective pressures on the virus to evolve counter defence strategies (de Ronde *et al.*, 2014). However, resistance

breaking has been previously reported for recessive *eIF(iso)4E* resistance to TuMV, possibly due to VPg polymorphisms acting via an eIF(iso)4E independent pathway (Gallois *et al.*, 2010). It remains to be seen whether this resistance breaking will also occur in the CRISPR/Cas9 induced *eIF(iso)4E* mutants in this study. However, the fact that resistance breaking was reported when a similar approach was adopted in cucumber (Chandrasekaran *et al.*, 2016) suggests a high propensity of resistance breaking in complete loss-of-function *eIF4E/eIF(iso)4E* mutants. Interestingly, it has recently been demonstrated that complete knock-out (KO) *eIF4E* mutations exhibit a narrower resistance spectrum compared to natural mutations, which usually maintain translational activity but result in weaker VPg-eIF4E/eIF(iso)4E interactions *(Piron et al.*, 2010, Ruffel *et al.*, 2005). This suggests that site directed mutagenesis to create *R genes* at *eIF4E/eIF(iso)4E* loci which resemble natural mutations, may be a superior method to complete loss-of-function mutations.

5.7.4 Technological outlook: the potential to copy natural mutations by genomic editing

In addition to supporting a greater durability and range of resistance, subtle mutations (affecting VPg binding, with little effect on host translation) may be more appropriate than KO mutations at *eIF4E/eIF(iso)4E* loci for several reasons (reviewed in (Bastet *et al.*, 2017)). Firstly, retention of ability to translate host mRNA should circumvent most of the potential fitness trade-offs associated with the mutations at single *eIF4E/eIF(iso)4E* loci. Furthermore, subtle mutations can support 'stacking' of *R genes* at both *eIF4E* and *eIF(iso)4E* loci. Such stacking is not possible for KO mutations as loss of function at both *eIF4E* and *eIF(iso)4E* results in severe developmental defects and the two loci are not concurrently heritable (Callot & Gallois, 2014, Gauffier *et al.*, 2016, Patrick *et al.*, 2014). The ability to stack *R-genes* at these two loci is necessary to generate resistance to some viruses which can utilise both isoforms (Hwang *et al.*, 2009, Jenner *et al.*, 2010), and will also be useful to generate resistance to multiple virus species which have evolved specificity for complementary isoforms. There already exists a wealth of information about the amino acid changes at eIF4E/eIF(iso)4E which are likely to be responsible for resistance to certain viruses (Lebaron *et al.*, 2016). Moreover, the fact that overexpression of such alleles can result in potyvirus resistance despite the presence of the native susceptible

alleles (Cavatorta *et al.*, 2011, Kim *et al.*, 2014) indicates that copying these mutations by genome editing at native loci is likely to be a successful strategy. Clearly, our limitation to apply this knowledge of natural variation at *eIF4E/eIF(iso)4E* loci to generate potyvirus resistance lies with the technological implementation. While it is relatively simple to generate site specific KO mutations using CRISPR/Cas9 methods, site directly mutagenesis altering only certain amino acids is more challenging. Recent methods have been developed to this end. Firstly CRISPR/Cas9 has been used to knock-in (KI) sequences by homology directed repair (HDR) by supplying an exogenous repair template to insert into the targeted DSB produced by Cas9 (Baltes *et al.*, 2014, Li *et al.*, 2013a, Sauer *et al.*, 2016, Shi *et al.*, 2017, Svitashev *et al.*, 2015, Wang *et al.*, 2017a, Zhao *et al.*, 2016). However, the frequency of KI events is low because in plants the NHEJ pathway is more active than HDR pathways (Puchta, 2005). An alternative approach has been explored using a recombinant Cas9 protein fused to a cytidine deaminase. Targeting the cytidine deaminase to specific loci by CRISPR/Cas9 can result in tailored de-amination cytosine residues (base-editing) which has the eventual effect of a cytosine to thymine (C→T) substitution. While base editing has been shown to be efficient (Komor *et al.*, 2016), the types of mutations it can induce are limited so it may not be the most appropriate method for engineering potyvirus *R genes*. Technological developments which may further aid the creation of potyvirus *R genes* are explored in the subsequent chapter.

Chapter 6: Exploring new CRISPR methods

6.1 Chapter introduction

6.1.1 Recombinant forms of Cas9 expand its biotechnological repertoire

Most CRISPR/Cas9 studies to date have utilised Cas9 to introduce DSBs at target loci resulting in random mutagenic indels, as outlined in the previous chapter. However, several modifications to this canonical approach have begun to emerge. Cas9 possesses two nucleolytic domains, named RuvC and HNH. Upon recognition of a target DNA sequence the HNH domain cleaves the DNA strand which is complementary to the guide RNA sequence, and the RuvC domain cleaves the opposite strand *(Sternberg et al*., 2015). Both cleavage events occur exactly 3bp upstream of the PAM, resulting in a blunt DSB. Mutating either of the nucleolytic domains, leaving only one functional, converts Cas9 into a 'nickase' which cleaves (or 'nicks') only one DNA strand at the target locus. Paired Cas9 nickases can thus generate DSBs if they are targeted to complementary strands at the same locus (Ran *et al.*, 2013). The main advantage of using a paired nickase approach is that it greatly improves the specificity of the gene editing by reducing the likelihood of off-target mutations. While each Cas9 nickase individually has the potential to cleave non-target loci, the fidelity of DNA repair for single strand breaks (SSBs) is much higher than for DSBs. As the likelihood of both 20nt guide sequences occurring in close proximity at non-target loci is generally very low, paired nickases greatly improve the specificity of Cas9-induced mutations. A further advantage of using Cas9 nickases is that they can be designed to generate staggered DSBs by careful design of the two guide RNAs. The importance of staggered DSBs is that they are likely to be more conducive for homology directed repair (HDR) to facilitate targeted KI of DNA sequences. This is further discussed below. So far, Cas9 nickases have been used most widely in mammalian cell systems (Kurihara *et al.*, 2017, Ran *et al*., 2013, Shen *et al.*, 2014), though there are a few reports emerging for plants (Mikami *et al.*, 2016, Schiml *et al.*, 2014). This is possibly because the frequency of off target mutations has been shown to be much lower in plant cells

compared to animal cancer cell lines with impaired repair pathways (Pan *et al.*, 2016, Shan *et al*., 2013, Sun *et al.*, 2015, Xie & Yang, 2013, Zhang *et al.*, 2014).

Further to the creation of Cas9 nickases by ablation of one nuclease domain, both nuclease domains can be mutated to generate a catalytically inactive, or dead, Cas9 (dCas9). Although unable to cleave DNA, the dCas9 protein retains its ability to bind sgRNAs and DNA. By fusing transcriptional activators or repressors to dCas9, this modified CRISPR/Cas9 system can be used to manipulate gene expression by targeting specific endogenous promoters (Chavez *et al.*, 2015, Gilbert *et al.*, 2014, Mali *et al.*, 2013). A further application of dCas9 has recently emerged in the form of the aforementioned base-editing method. Fusion of dCas9 to a cytidine deaminase allows for targeted genome editing without the need for DSBs, hence resulting in more deterministic (less random) mutations (Komor *et al*., 2016). Addition of fluorophores to dCas9 has also enabled confocal imaging of precise DNA sequences (Anton *et al.*, 2016), though these methods have not been widely used so far.

6.1.2 Natural diversity of Cas9 orthologues

The most frequently used Cas9 orthologue is derived from *Streptococcus pyrogenes* (SpCas9). However, its widespread usage is most likely a result of convention rather than scientific rationale, as it was one of the first orthologues to be appropriated for biotechnological purposes and hence has a wealth of resources. Several additional Cas9 orthologues have begun to be used for biotechnological purposes. Interestingly, several of the identified Cas9 variants require PAM sequences which deviate from the 'NGG' of SpCas9 (Table 6.1). Hence, alternative sources of Cas9 can be used if the 'NGG' PAM of SpCas9 is restrictive to experimental design. Furthermore, there are reports that, in certain organisms, SpCas9 may not be as active as other orthologues (Esvelt *et al.*, 2013, Steinert *et al.*, 2015). Despite this, use of SpCas9 remains prevalent though this may change as the characterisation and resources improve for the alternative Cas9 variants.

Cas9 variant	PAM	References
Streptococcus pyrogenes	NGG	(Mojica *et al.*, 2009)
Streptococcus agalactiae	NGG	(Mojica *et al*., 2009)
Streptococcus mutans	NGG	(van der Ploeg, 2009)
Streptococcus thermophilus	NGGNG/NNAGAAW	(Horvath *et al.*, 2008)
Listeria monocytogenes	NGG	(Mojica *et al*., 2009)
Neisseria meningitidis	NNNNGATT	(Esvelt *et al*., 2013, Zhang *et al.*, 2013b)
Camplyobacter jejuni	NNNNACA	(Fonfara *et al.*, 2014)
Francisella novidida	NG	(Fonfara *et al*., 2014)
Treponema denticola	NAAAAN	(Esvelt *et al*., 2013)

Table 6.1 Summary of Cas9 orthologues from different bacterial species, and associated PAM sequences (N=A,T,C, or G ; W=A or T)

6.1.3 Cpf1: a novel Cas orthologue with distinct features to Cas9

The CRISPR-associated (Cas) nucleases have been classified functionally which, at the broadest level, divides them into two classes. Class I Cas enzymes require hetero-oligomers of Cas proteins to target and cleave DNA, whereas for class II Cas enzymes these processes are mediated by a single paralogue (Makarova *et al.*, 2015). Biotechnological exploitation of CRISPR/Cas systems has focussed on class II due to the comparative simplicity of this class. The two classes are further divided into types. Class I comprises types I, III, and IV, while class II contains types II and V. Cas9 is a class II, type II enzyme and has been extensively used for genome editing purposes. Recently, a novel Cas orthologue was characterised which defines the emergent type V grouping. This orthologue was named Cpf1 (CRISPR from *Prevotella* and *Francisella* 1)(Zetsche *et al*., 2015). Cpf1 has several features which are distinct from Cas9 endonuleases (Figure 6.1).

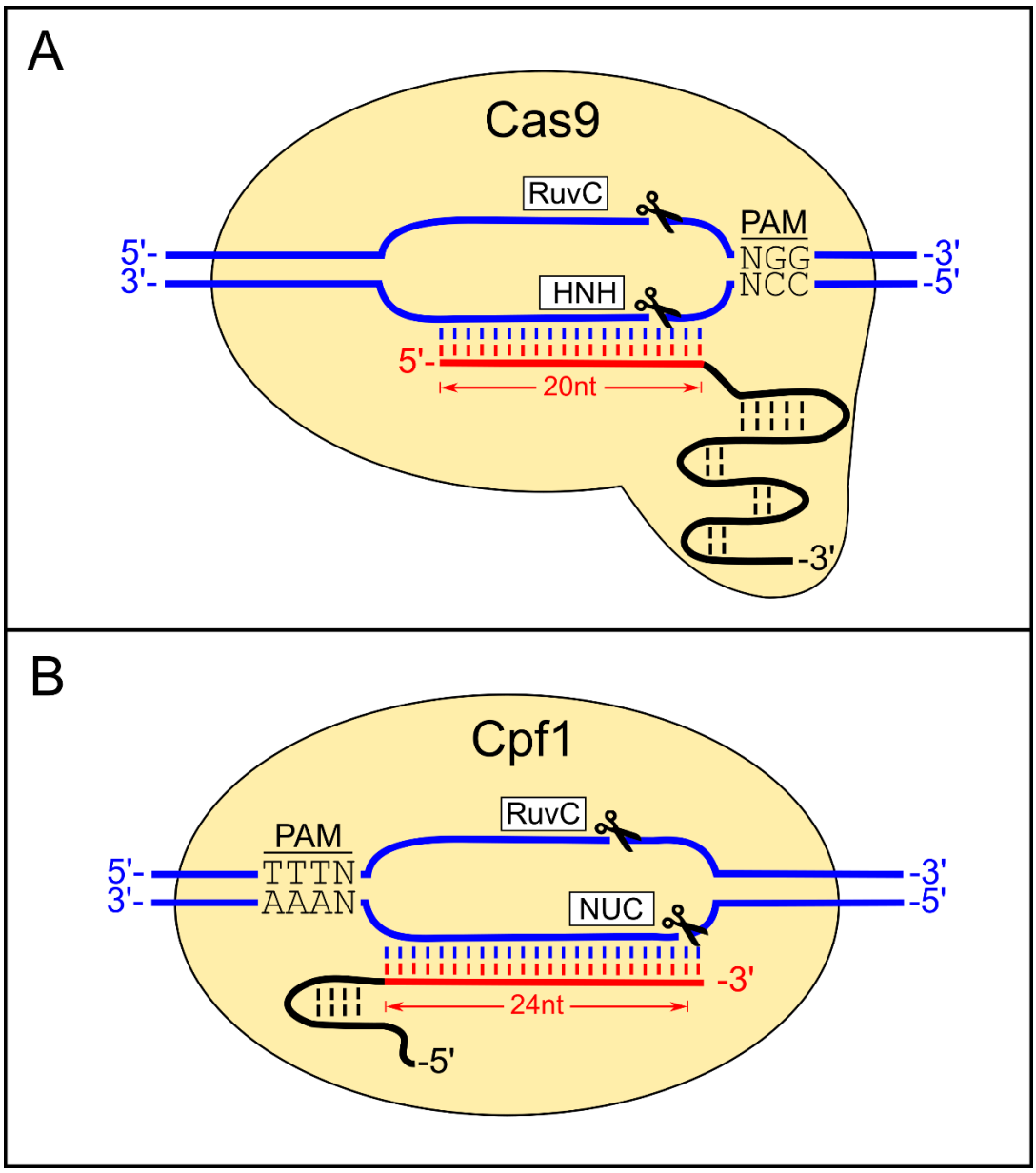

Figure 6.1 Schematic of Cas9 and Cpf1 induced DSBs in DNA. The guide sequence of the sgRNA/crRNA is shown in red, and the structural sequences of the sgRNA/crRNA is shown in black. DNA is represented by blue lines.

Cpf1 contains a single RuvC nuclease domain but does not share the HNH domain of Cas9. This led to initial speculation that Cpf1 may operate as a homo-dimer to cleave DNA (Zetsche *et al.*, 2015). However, experimental evidence suggested that it acts as a monomer (Dong *et al.*, 2016, Fonfara *et al.*, 2016, Yamano *et al.*, 2016). Recently a second putative nuclease domain, called NUC, was described for Cpf1 (Yamano *et al.*,

2016). The current model suggests that Cpf1 monomers generate DSBs at target DNA by the combined action of the NUC and RuvC domains which cleave the target and non-target DNA strands, respectively (Yamano *et al*., 2016). Interestingly, unlike the blunt DSBs induced by Cas9, Cpf1 produces DSBs with approximately 4-5bp cohesive overhangs (Zetsche *et al*., 2015). In contrast to the 3', predominantly 'G'-rich PAM of Cas9; Cpf1 uses a 5', 'T'-rich PAM (Zetsche *et al*., 2015).

The guide RNAs of Cas9 and Cpf1 systems are also dramatically different. In nature, Cas9 is guided to target DNA by the concerted action of a crRNA (CRISPR RNA) and a tracrRNA (trans-acting CRISPR RNA). The crRNA contains a 5' structural motif which interacts with the tracrRNA, and a 3'-20nt guide sequence (or spacer) with homology to the target DNA. The tracrRNA forms multiple hairpin loops which are necessary for complexing the crRNA/tracrRNA with Cas9. As previously mentioned, the crRNA and tracrRNA can be fused for biotechnological purposes to form a roughly 100nt sgRNA (Jinek *et al*., 2012). In contrast to this relatively complex guiding system of Cas9, Cpf1 naturally utilises a single crRNA with no requirement for a tracrRNA. The Cpf1 crRNA is approximately 60nt (prior to crRNA maturation) and contains a 5'-36nt direct repeat and a 3'-24nt guide sequence (or spacer). This direct repeat can be divided into two domains, a 5' repeat leader and a 3' stem-loop region (Zetsche *et al*., 2015). The stem-loop region (proximal to the 23nt guide sequence) interacts with the Cpf1 protein to allow the formation of crRNA/Cpf1/target-DNA tertiary complexes, and the repeat leader directs the sequential cleavage of individual crRNAs from long CRISPR array transcripts (Zetsche *et al*., 2015). Cleavage at flanking repeat leaders releases a so-called pre-crRNA which contains a full length direct repeat. The 'pre-crRNA' is then further processed to remove the 5' repeat leader to yield a 'mature crRNA' consisting of the 5' stem loop and the 3' guide sequence (Fonfara *et al*., 2016).

A further important difference between the guide RNAs of Cas9 and Cpf1 relate to their processing from long CRISPR array transcripts. For Cas9, the processing of crRNA and tracrRNA requires several enzymes including other Cas paralogues and a RNAseIII enzyme (Deltcheva *et al.*, 2011). In contrast, Cpf1 possesses RNAse activity allowing it to process its own crRNAs from transcribed CRISPR arrays without the need for additional factors (Fonfara *et al*., 2016).

The divergent features of Cpf1 offer exciting potential for biotechnological exploitation which in several ways may be complementary or ever superior to Cas9. Firstly, the 'T'-rich PAM favours targeting of genomes or genomic regions with a low 'GC' content. Secondly, the staggered DSBs generated may facilitate KI events by HDR (discussed below). Thirdly, the smaller coding sequence of Cpf1 nuclease and simpler crRNA alleviate the detrimental genetic load when attempting to deliver these components to cells by viral vectors. Finally, the self-processing of crRNAs by Cpf1 can greatly simplify multiplex editing (where multiple loci are simultaneously targeted), as multiple guides can be delivered as a polycistronic sequence resembling a natural CRISPR array (Wang *et al.*, 2017b, Zetsche *et al.*, 2017).

6.1.4 Using CRISPR systems for targeted gene replacement

The ability to introduce exogenous DNA or rearrange endogenous DNA at targeted loci in an inducible manner has be an important goal for plant biotechnology for many years. While this would represent a powerful tool for modifying plant genomes to enable fundamental research in plant biology, arguably the greatest impetus is the potential to improve methods for crop breeding. Research towards this goal has largely focussed on exploring ways to introduce DNA sequences by inducing HDR. HDR is a DNA repair pathway whereby DSBs are repaired by intact DNA with extensive homology to the sequences either side of the DSB. In a natural context, the sister chromatid is the most frequent template for repair. In HDR, the 5' ends of the DNA either side of the DSB are degraded in a process called resection, resulting in 3' overhangs. These 3' overhangs invade the DNA duplex of the repair template, forming a so-called D-loop. The DNA repair can occur by one of two mechanisms known as double strand break repair (DSBR) and synthesis dependent stand annealing (SDSA). Precise details of these two models of repair can be found in the following reviews (Pfeiffer *et al.*, 2000, Rodgers & McVey, 2016).

It has been estimated that only 0.01% of DSBs are repaired by HDR in higher eukaryotes (Puchta, 2005), as NHEJ (which involves direct fusion of broken DNA ends) is the predominant repair pathway. Hence, the low propensity for HDR is a major hurdle to its application for targeted DNA modifications. Nevertheless, targeted KI of exogenous sequences to endogenous loci has been achieved several times in plants,

albeit at low frequencies. Site specific induction of DSBs is the first requirement for targeted KI events. Early studies utilised ZFNs and meganucleases to this end, resulting in gene replacement in tobacco (Cai *et al.*, 2009), maize (Shukla *et al.*, 2009), and cotton (D'Halluin *et al.*, 2013). However, in addition to the low rate of KI events, the cost and labour required to produce ZFNs and meganucleases is greatly prohibitive to their use in crop breeding and hampers research progress. CRISPR/Cas technology has revolutionised the field of targeted gene replacement. This is reflected by the number and frequency of reports using CRISPR/Cas technology to induce KI in a variety of plant species (Table 6.2). While CRISPR/Cas technology greatly enables the induction of targeted DSBs, KI frequencies remain low due to the inherently low preponderance of HDR mediated repair. Nonetheless, the reduced cost and labour conferred by CRISPR/Cas approaches has accelerated research to determine factors which may promote HDR. For instance, increasing the copy number of exogenous repair templates (Baltes *et al.*, 2014, Sun *et al.*, 2016, Wang *et al.*, 2017a), and using templates with long homology arms flanking the induced DSB have been shown to increase KI efficiencies (Sauer *et al.*, 2016). Additionally, research in animal systems has indicated that suppression of factors required for NHEJ repair can result in an increase in HDR mediated repair, resulting in higher frequencies of targeted KI events (Chu *et al.*, 2015, Robert *et al.*, 2015, Zhu *et al.*, 2015).

An alternative, putative repair mechanism has been described which is conceptually similar to HDR but is mechanistically more similar to NHEJ. So-called microhomology-mediated end joining (MMEJ) results when compatible cohesive ends exist around a DSB (McVey & Lee, 2008). MMEJ can potentially be used to KI sequences by flanking exogenous DNA with overhangs that are complimentary to a staggered DSB. The MMEJ repair mechanism does not rely on extensive homology between the repair template (or donor DNA) and the DSB flanking sequences, and does not involve strand invasion. Rather, microhomology between the staggered DSB and donor DNA bring the exogenous sequence in close proximity to the DSB whereupon NHEJ-style ligation can result in targeted KI. The NHEJ, MMEJ, and HDR DNA repair pathways are summarised in Figure 6.2. The potential for CRISPR/Cpf1 to generate staggered DSBs at target loci offers tantalising prospects for efficient gene replacement by MMEJ. Such methods however, have not yet been explored.

DBS inducer	Delivery method	Repair substrate	KI frequency	References
ZFN	Agrobacterium	Integrated transgene	10%*	(Cai *et al*., 2009)
Meganuclease	Biolistic bombardment	Integrated transgene	2%	(D'Halluin *et al*., 2013)
Cas9	PEG	ssODN	2.4%	(Sauer *et al*., 2016)
Cas9	PEG	dsDNA	9%	(Li *et al*., 2013a)
Cas9	Agrobacterium	Donor DNA plasmid	4%	(Svitashev *et al*., 2015)
Cas9	Agrobacterium	Donor DNA plasmid	0.8%	(Zhao *et al*., 2016)
Cas9	Biolistic bombardment	Donor DNA plasmid	1%	(Shi *et al*., 2017)
Cas9	Agrobacterium	DNA replicon	N/A	(Baltes *et al*., 2014)
Cas9	Agrobacterium	DNA replicon	19.4%*	(Wang *et al*., 2017a)

Table 6.2 Summary of HDR mediated gene KI in plants. PEG = polyethylene glycol, ssODN = single stranded oligodeoxynucleotide, dsDNA = double stranded DNA. *indicates KI frequencies which are over-optimistic due to selection bias in the scoring method.

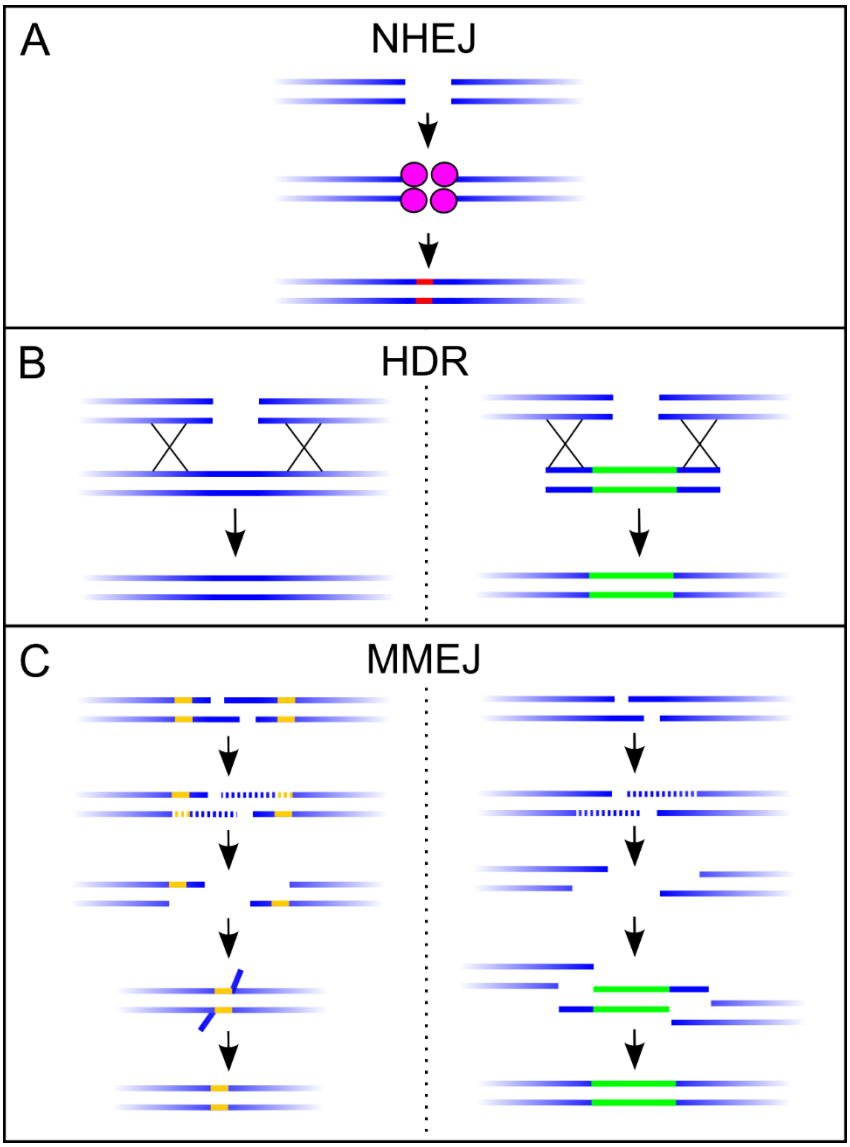

Figure 6.2 Endogenous DNA DSB repair pathways and their biotechnological exploitation. **A** The main DSB repair pathway, NHEJ, requires binding of several host factors (pink) to the cleaved DNA ends which results in ligation of the DSB. This process is error prone and often introduces short mutations (red) around the DSB. **B (Left)** The natural HDR pathway most frequently involves a non-damaged sister chromatid acting as a repair template. **(Right)** HDR can be utilised for targeted KI of exogenous sequences (green) by delivering donor DNA with extensive homology arms flanking the DSB. **C (Left)** In a natural context MMEJ frequently results in short deletions. 5' to 3' resection around the DSB reveals regions of 4-30nt microhomology (yellow) on opposite DNA strands which directs the DSB repair in a similar way to NHEJ. **(Right)** MMEJ can potentially be exploited to KI exogenous sequences (green) by supplying a donor DNA with staggered microhomology domains which are complementary to a staggered DSB.

6.2 Chapter aims

This chapter aims to explore several features of the recently described CRISPR-associated endonuclease, Cpf1. At the time this work was started, Cpf1 had not been utilised for genome editing in plant species. Therefore, the first aim was to determine which Cpf1 orthologue would be most appropriate to use for plant genomes. Secondly, a screen was devised to compare the activity of Cpf1 to the well characterised SpCas9. After exploring these basic features of Cpf1, the possibility of using MMEJ-mediated repair mechanisms as a novel method for targeted DNA KI into plant genomes was explored. An ultimate goal of this research, beyond the scope of this chapter, is to develop technologies that could enable the creation of novel *R genes* in crops that closely resemble naturally occurring mutations, as outlined in the previous chapter discussion.

6.3 Testing Cpf1 for *in planta* genome editing

6.3.1 Selecting a suitable Cpf1 orthologue for in planta genome editing

In the first report describing the novel Cas nuclease, Cpf1, it was demonstrated that two out of several Cpf1 orthologues could induce genome editing in mammalian cells (Zetsche *et al*., 2015). Hence these orthologues, from *Acidaminococcus sp.* (AsCpf1) and *Lachnospiraceae bacterium* (LbCpf1), were chosen to test their *in planta* activity, based on the assumption that editing requirements in eukaryotic cells would be similar. AsCpf1 and LbCpf1 coding sequences were cloned into a binary vector (pK7FWG2) for plant transformation, flanked by 35S promoter and terminator sequences. crRNAs were designed to target the transgenic GFP locus present in the stably transformed *N. benthamiana* plant line, called 16c. The same 23nt guide sequence was used for both Cpf1 orthologues but the direct repeats were designed based on the published sequences for the As/LbCpf1 direct repeats, which are subtly different (Zetsche *et al*., 2015). These crRNAs were cloned downstream of a *U6* promoter into separate pK7FWG2 vectors. The AsCpf1 and LbCpf1 constructs, along with their cognate crRNAs, were then introduced to 4 week old *N. benthamiana* leaves by transient *Agrobacterium* infiltration. A previously tested GFP-targetting sgRNA was used along

with an SpCas9 expression vector as a positive control for the T7 assay. A T7 cleavage product was clearly visible for the SpCas9 + sgRNA control and also for LbCpf1 + crRNA sample, but not for the AsCpf1 + crRNA sample (Figure 6.3). As the guide sequence for the two Cpf1 crRNAs was identical these results suggest that the LbCpf1 orthologue is superior to the AsCpf1 orthologue for targeted mutagenesis of plant genomes. In support of this conclusion, several groups have recently produced similar results when comparing the *in planta* mutagenesis frequency of these two Cpf1 orthologues (Hu *et al.*, 2017, Kim *et al.*, 2017a, Tang *et al.*, 2017).

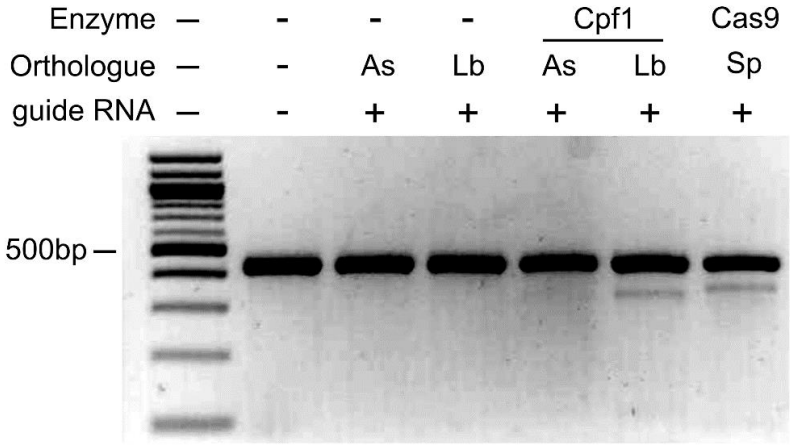

Figure 6.3 T7 assay to assess the *in planta* activity of AsCpf1 and LbCpf1. Constructs expressing Cas9/Cpf1 orthologues along with their respective crRNAs were introduced to *N. benthamiana* leaves by syringe infiltration with *Agrobacterium tumefaciens*. A non-infiltrated sample (lane 1) and samples infiltrated with only As/Lb crRNAs (lanes 2/3) were included as negative controls. crRNAs were designed to target the GFP transgene which was stably integrated into the *N. benthamiana* genome. A previously verified sgRNA targeting the GFP sequence, along with Cas9, was used as a positive control for the T7 assay. Samples were collected 3 days after infiltration. Infiltration and T7 assay performed by Aron Ferenczi.

6.3.2 Design of a phenotypic screen to compare the efficiencies of SpCas9 and LbCpf1 for inducing heritable mutations

After establishing that LbCpf1 is the most suitable orthologue for plant applications, the next basic feature of this novel nuclease to assess was whether it could produce heritable mutations at a similar frequency to the commonly used SpCas9. To this end,

a phenotypic screen was designed. Most of the phenotypic screens used to test the editing efficiency of Cas9 in plant genomes have relied on targeting genetically dominant genes which have easily discernible loss-of-function phenotypes (Feng *et al*., 2013, Feng *et al.*, 2014, Li *et al*., 2013a). A limitation of using genetically dominant loci to score mutation frequencies is that it results in an underestimation of editing because mutations in just one allele cannot be phenotypically detected, due to the presence of the dominant wild type allele. Targeting recessive loci, with easily discernible loss-of-function phenotypes, on the other hand would allow all heritable mutations to be detected, as a loss-of-function mutation at one or both alleles would result in a scorable phenotype. However, it would still not be possible to score how many of these genotypes are mono-allelic, bi-allelic, or homozygous mutations as all three would result in the same phenotype. While most genes behave either dominantly of recessively, some are classified out-with this dichotomy and are described as exhibiting semi-dominant behaviour. At semi-dominant loci homozygous mutant, heterozygous mutant, and wild type genotypes can each be distinguished by a unique phenotype, which can arise by either partial gain or partial loss of function. When a semi-dominant trait arises by partial loss of function it is described as haploinsufficient (Meinke, 2013). Thus, targeting haploinsufficient loci for CRISPR-induced mutations could offer a means to phenotypically score total mutations frequencies and simultaneously distinguish heterozygous mutants from homozygous/biallelic mutations. To identify a haploinsufficient locus that could be appropriate for such a screen, a survey of loci with documented haploinsufficient phenotypes was conducted. The candidate loci from this survey (Table 6.3) were given a rank based on their appropriateness for phenotypic screening. The lowest rank (4) comprised loci in which mutations resulted in embryo or fertility defects, as these would obviously impede scoring heritable mutations. The next rank (3) contained loci which exhibit haploinsufficient phenotypes that require tissue sampling and chemical or microscropic analysis to accurately score them. Ranks 1 and 2 represented the haploinsufficient loci with phenotypes that could be easily quantified without the need for tissue sampling. Rank 1 was defined by the presence of a phenotype that can be scored during early stages of development and can be screened by high-throughput methods (the flowering phenotypes of the Rank 2 candidates do not meet either of

these criteria). This ranking method resulted in one locus, *ARF19*, emerging as the most appropriate target for designing a haploinsufficiency screen to compare the editing capabilities of SpCas9 and LbCpf1.

6.3.3 Design of ARF19-targeting sgRNAs and crRNAs for CRISPR/Cas9/Cpf1-induced mutations

To accurately compare the mutagenic capabilities of Cas9 and Cpf1, sgRNAs/crRNAs were designed for both enzymes containing an identical, or nearly identical, guide RNA sequence. This measure was taken to rule out, or at least reduce the effect of differences in the guide sequence confounding the comparison of the activity of the two enzymes. Because Cpf1 and Cas9 utilise guide RNAs with 5'(TTTN) and 3'(NNG) PAM sequences, respectively, guide RNAs were designed by scanning the *ARF19* locus for a TTTN(N20-23)NNG sequence. Two target regions, in the second and seventh exons, were identified by this method (Figure 6.4). These guide RNA sequences were cloned as mature crRNAs/sgRNAs to guide Cpf1/Cas9, respectively, to the two regions of the *ARF19* locus. Hereafter, 'arf19_1' and 'arf19_2' will be used to refer to the *ARF19* guides/targets in the second and seventh exon, respectively. Each sgRNA/crRNA expression cassette was cloned downstream of a *U6* promoter into a pK7FWG2 vector containing either Cas9 or Cpf1, flanked by *UBIQUITIN* promoter/terminator sequences. Additionally, an *OLEOSIN:RFP* cassette was included in the T-DNA region of the vectors to enable selection and counter-selection of the transgene at the T_1 and T_2 generations, respectively (Figure 6.5A). The *OLEOSIN:RFP* cassette was previously developed as a technology, called FAST(Fluorescence Accumulating Seed Technology)-red, to facilitate the selection of transgenic seeds (Shimada *et al.*, 2010). Expression of RED FLUORESCENT PROTEIN (RFP) under the embryo-specific and highly active *OLEOSIN* promoter results in accumulation of RFP in the embryo which can be visualised through the translucent *Arabidopsis* seed coat. This results in a 'red seed' phenotype for genotypes containing the *OLEOSIN:RFP* transgene (Figure 6.5B/D).

Locus	Full gene name	Heterozygous phenotype	Homozygous phenotype	Rank
ARF19	Auxin Response Factor 19	partial auxin sensitivity	auxin insensitivity	1
MAF1	MADS Affecting Flowering 1	slightly early flowering	extreme early flowering	2
CO	Constans	slightly late flowering	extreme late flowering	2
GIS	Glabrous Inflorescence Stems	slightly reduced trichome number	severely reduced trichome number	3
PHOT2	Phototropin 2	Slightly reduced chloroplast light acclimation	Severely reduced chloroplast light acclimation	3
NYE1	Non-Yellowing 1	slightly reduced chlorophyll	severely reduced chlorophyll	3
LUT2	Lutein Deficient	slightly reduced β-carotene content	severely reduced β-carotene content	3
NPQ4	Non-Photochemical Quenching 4	slightly reduced photochemical quenching	severely reduced photochemical queching	3
ESM	Epithio-Specifier Modifier	slightly altered glucosinolate hydrolysis	severely altered glucosinolate hydrolysis	3
IRT1	Iron Transport	early flowering in short days	non-fertile	4
SOZ1	Sensitive to Ozone 1	hypersensitive to ozone	embryo lethal	4
AML1	*Arabidopsis* Minute-Like 1	small plants	embryo lethal	4
TIC110	Translocon at Inner Envelope Membrane of Chloroplasts 110	pale yellow-green plants	embryo lethal	4

Table 6.3 Summary of semi-dominant mutations resulting in haploinsufficient phenotypes.

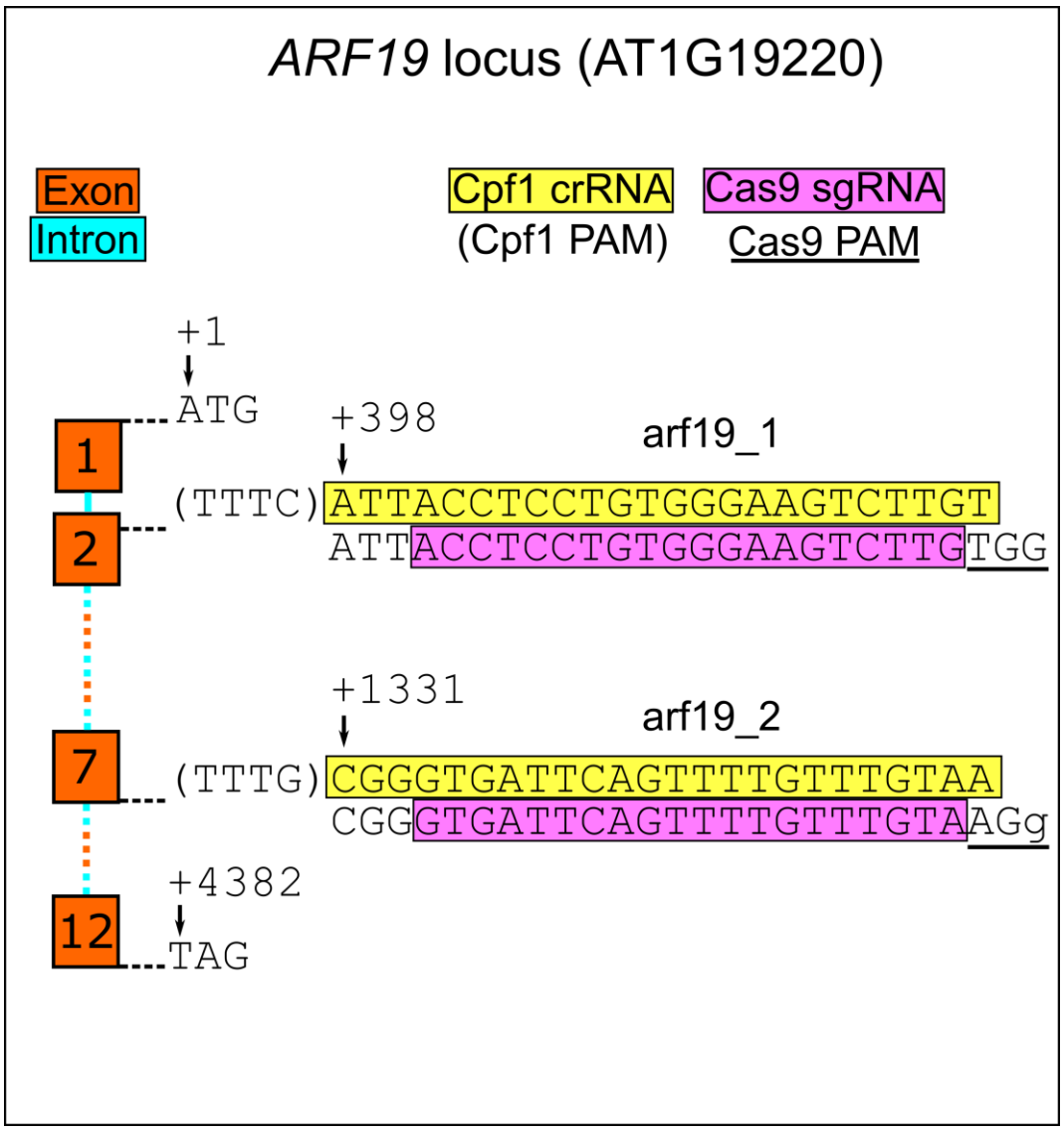

Figure 6.4 Design of overlapping Cas9 and Cpf1 guide RNAs to target the *ARF19* locus in *Arabidopsis*. Cpf1 crRNA sequences are shown in yellow, Cas9 sgRNAs are shown in pink. The Cpf1 PAM is enclosed in brackets and the Cas9 PAM is underlined. Exonic and intronic sequences are indicated orange boxes and blue lines, respectively. Sequence positions relative to the start of translation (+1) are indicated.

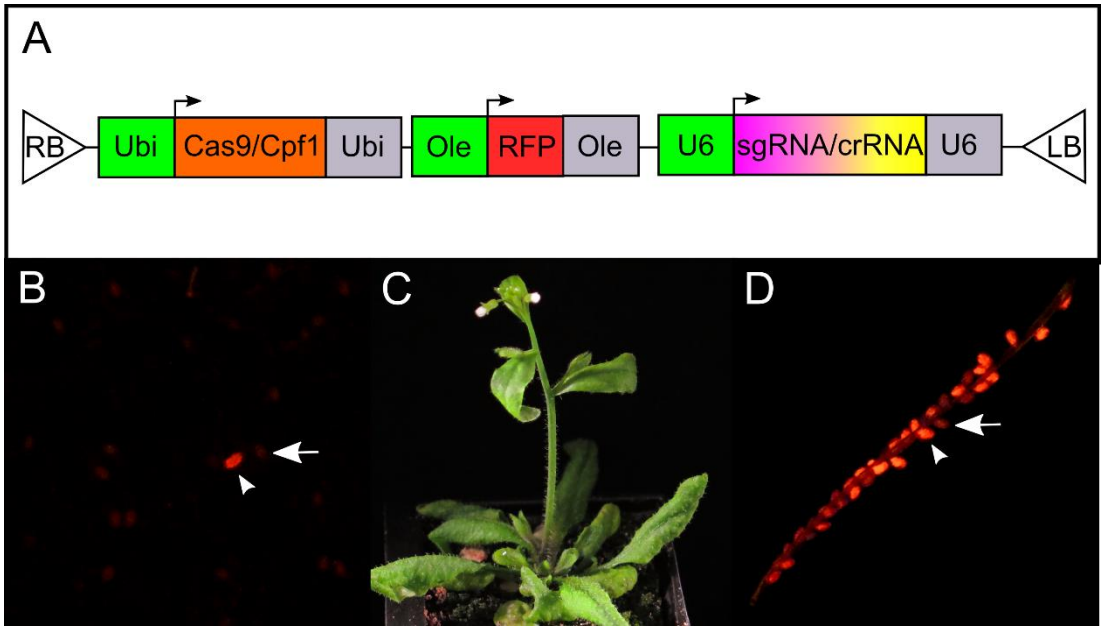

Figure 6.5 Cas9 and Cpf1 vectors using FAST-red technology for selection of transgenic T_1 seeds and counter-selection to identify non-transgenic T_2 seeds. **A** T-DNA map of the Cas9 and Cpf1 expression vectors. LB/RB denote the left/right border sequences that demarcate the T-DNA. Promoter/terminator sequences are shown in green/grey, respectively. *Ubi* = *UBIQUITIN*, *Ole* = *OLEOSIN*. Black arrows indicate the transcriptional start sites. **B** Selection of transgenic T_1 seed. **C** Transgenic T_1 plant. **D** Representative silique from a T_1 plant showing roughly 3:1 segregation of the fluorescent: non-fluorescent phenotype conferred by the dominant FAST-red transgene. White arrowheads indicate red fluorescing seeds, white arrows indicate non-fluorescing seeds.

6.3.4 Testing the arf19 haploinsufficient phenotype

Mutations at the *ARF19* locus have previously been reported to result in an insensitivity to exogenous auxin, in a haploinsufficient manner (Li *et al.*, 2006, Okushima *et al.*, 2005). Dark-grown seedlings display what is known as skotomorphogenic growth which is manifested by a dramatic elongation of the hypocotyl (Josse & Halliday, 2008). If wild type plants are grown in the presence of the synthetic auxin analogue, 2,4-dichlorophenoxyacetic acid (2,4-D), the skotomorphogenic hypocotyl extension is inhibited resulting in short hypocotyls. *arf19* mutants are insensitive to application of exogenous 2,4-D such that homozygous mutants display long hypocotyls when grown in the dark, even in the presence of 2,4-

D. Heterozygous *arf19* mutants exhibit hypocotyl lengths which are an intermediate between the short-wild type and long-homozygous-mutant phenotypes. While this haploinsufficent phenotype has previously been described (Li *et al*., 2006, Okushima *et al*., 2005), it was necessary to determine whether the hypocotyl assay would be accurate and sensitive enough to use as a marker to diagnose CRISPR-induced *ARF19* mutations. To this end, the previously described *arf19-1* homozygous mutant (Okushima *et al*., 2005) was obtained for testing the hypocotyl assay. When grown under the previously published conditions, the mean hypocotyl length of a population of homozygous *arf19-1* mutants was 1.85X greater than a wild type population. However, box plots of the hypocotyl lengths revealed significant overlap between the normal distribution curves of these two populations (Figure 6.6A). This indicates that the hypocotyl assay would not be appropriate, under these conditions, for a diagnostic screen for *ARF19* mutations. It was reasoned that most of the variation in hypocotyl length for both the *arf19-1* and wild type populations was due to variation in germination rates within each population, despite attempts to synchronise germination by seed stratification followed by a 5-hour light pulse. In an attempt to further synchronise germination to reduce the within-population hypocotyl length variation, deviations to the published assay were made by extending the light pulse period and using different intensities of light (Figure 6.6B-D). The lowest within-population variation was achieved by growing plants for 8 days under low light intensity. However, the between population variation was also significantly reduced under these conditions such that the two normal distributions still had considerable overlap (Figure 6.6B). In a further attempt to reduce the within-population variation the experiment was repeated, imaging the seedlings over a daily time-course to allow relative hypocotyl lengths to be calculated. The rationale behind this was that variation in germination time would have less influence on the relative hypocotyl growth rate, compared to the absolute hypocotyl lengths. However, there was still significant overlap in the normal distributions of relative hypocotyl growth rate for the *arf19-1* and wild type populations (Figure 6.7). Moreover, each of the hypocotyl assays were performed with homozygous *arf19-1* mutants. As the phenotype of the heterozygous mutants is an intermediate of the wild type and homozygous *arf19* mutant phenotypes, the assay conditions would need to be significantly improved to be able to

phenotypically discern the three genotypes. Taken together, these data imply that conditions could not be optimised to allow the *arf19* mutant phenotype to be used to diagnose CRISPR-induced mutations at the *ARF19* locus.

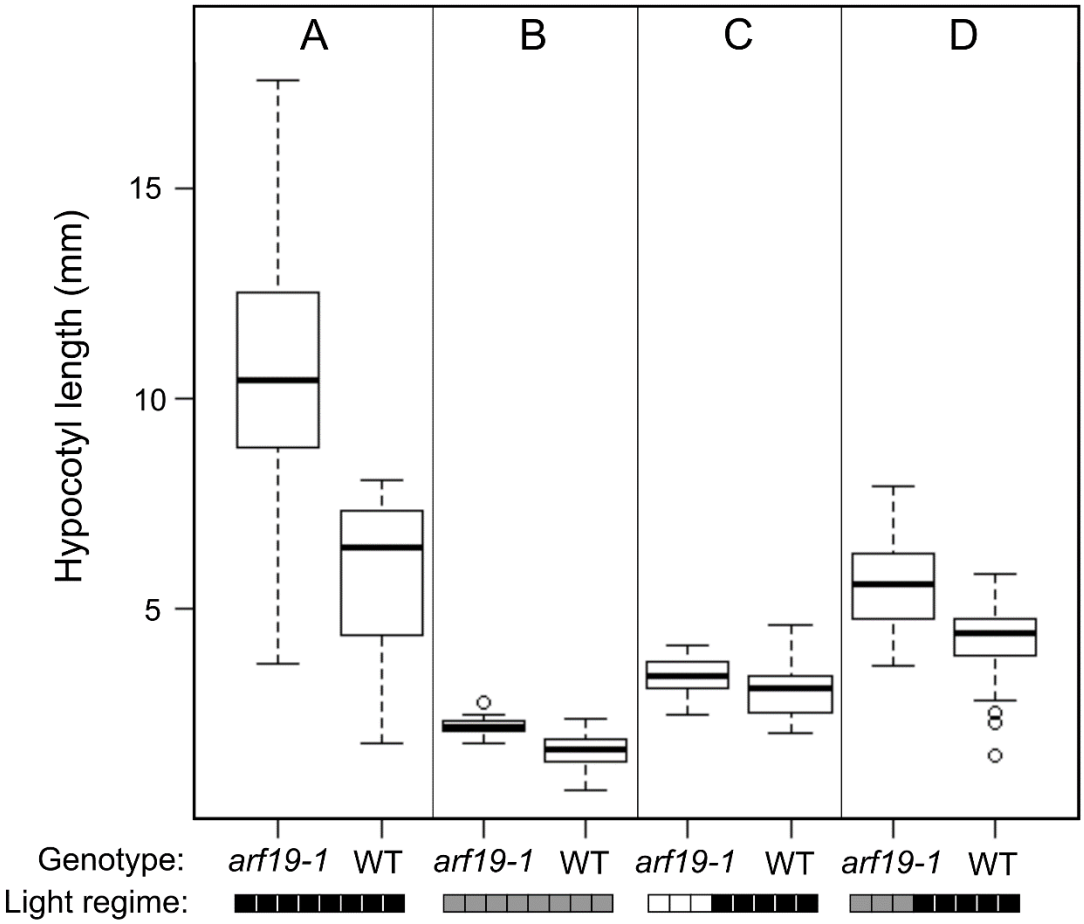

Figure 6.6 Quantitative assessment of the *arf19* mutant hypocotyl phenotype. Box plots indicate the median hypocotyl lengths for Wild Type (WT) and homozygous *arf19-1* seedlings, grown on agar plates containing 1µM 2,4-D. Error bars represent the range in hypocotyl lengths for approximately 30 seedlings. Seeds were stratified for 3 days at 4ºC in darkness and given a 5 hour light pulse of 200 µE white light to induce germination. Replicate plates were shifted to different light regimes, as indicated by shaded boxes. Black = darkness, grey = 50 µE white light, white = 200 µE white light. Each shaded box represents 24 hours in 16:8 light:dark photo-cycles. Hypocotyl lengths were measured after 8 days in the different light regimes.

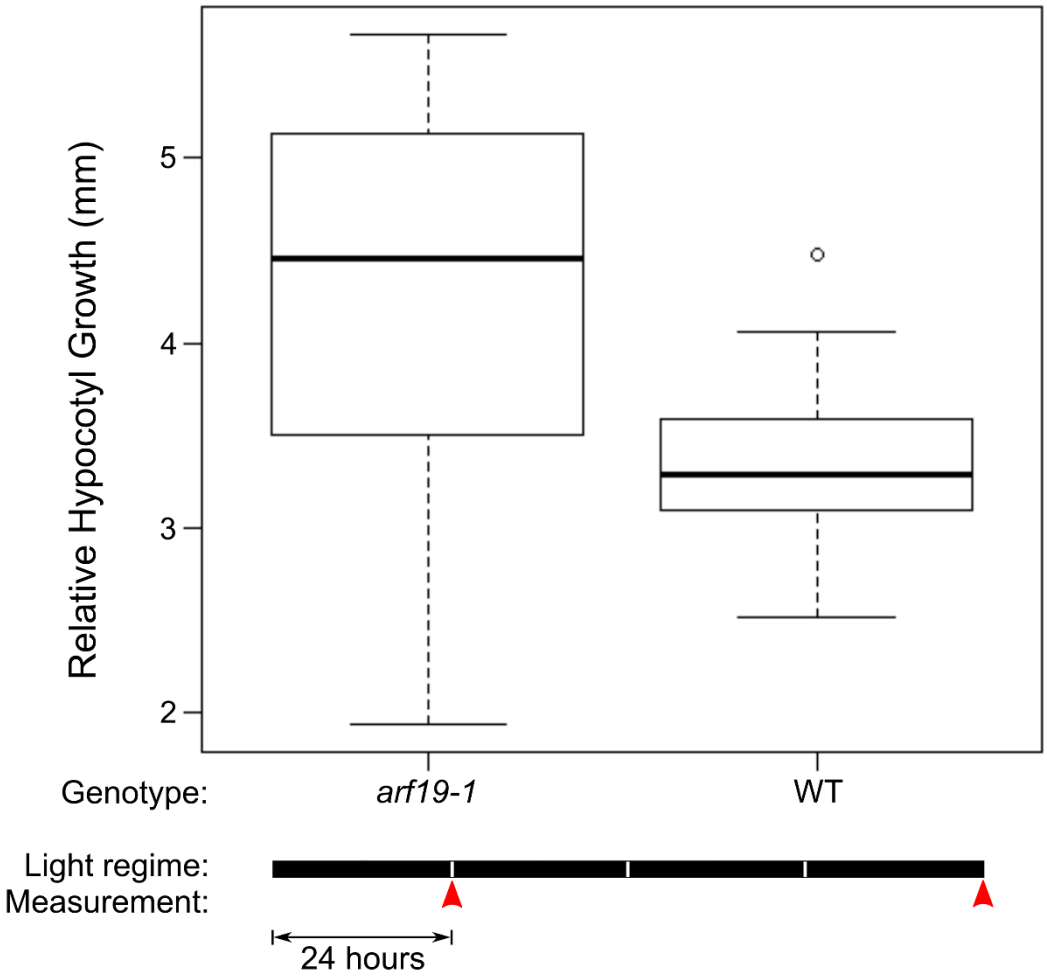

Figure 6.7 Quantitative assessment of the *arf19* mutant hypocotyl phenotype. Box plots indicate the median relative hypocotyl lengths for Wild Type (WT) and homozygous *arf19-1* seedlings, grown on agar plates containing 1µM 2,4-D. Error bars represent the range in hypocotyl lengths for approximately 30 seedlings. Seeds were stratified for 3 days at 4ºC in darkness and given a 5 hour light pulse of 200 µE white light to induce germination. Seedlings were grown in complete darkness for 4 days, taking hypocotyl measurements at the end of day 1 and day 4 (red arrow-heads). The difference in length for each seedling was taken between these two time points to estimate the relative hypocotyl growth.

6.3.5 Scoring CRISPR/Cas9/Cpf1-induced mutations in the T_1 generation

As the conditions for the phenotypic *arf19* mutant screen could not be satisfactorily optimised, a direct genotyping approach, using Sanger sequencing, was adopted. T_1 seeds for the four transgenic lines (SpCas9+sgRNAarf19_1, SpCas9+sgRNAarf19_2, LbCpf1+crRNAarf19_1, LbCpf1+crRNAarf19_2) were selected using the red seed fluorescence phenotype (Figure 6.5B). DNA was extracted from 35 seedlings from each of the transgenic lines and PCR amplicons of the two *ARF19* target loci were used to detect CRISPR/Cas9/Cpf1-induced mutations by Sanger sequencing. Strangely, no mutations were recovered for either of the SpCas9+arf19_1 or SpCas9+arf19_2 transgenic lines, nor the LbCpf1+arf19_2 line. Sequencing the sgRNA/crRNA-expressing regions from a randomly chosen sample from each of the four transgenic lines confirmed that lack of recovered mutations was not a result of mis-labelling seed batches or DNA samples. The only transgenic line which resulted in mutations at *ARF19* was LbCpf1+arf19_1. This line harboured mutations at the target locus at a frequency of 71.4% (25/35). Importantly, these mutations aligned to the predicted Cpf1 cleavage site indicating that they are *bona fide* CRISPR/Cpf1-induced mutations (Figure 6.8A) The frequencies of heterozygous, biallelic, and homozygous mutations were 2.8% (1/35), 28.6% (10/35), and 40% (14/35), respectively (Figure 6.8B). Such a high incidence of homozygous mutations in the T_1 generation is unprecedented, and is much higher than the 13% homozygous mutation frequency in T_1 seedlings previously reported for a SpCas9 editing system driven by an egg-cell specific promoter (Wang *et al.*, 2015). More intriguing still, is the fact that all of the homozygous mutants were genetically identical, harbouring an insertion of single thymine (T) nucleotide in between the two predicted Cpf1 cleavage sites. This indicates that the mutation induced by the arf19_1 crRNA was non-random. By careful analysis of this target region, a DNA repair model was proposed to explain this non-random 'T-insertion' mutation (Figure 6.9). Cpf1-induced cleavage results in a staggered cut at the *ARF19_1* target locus, leaving 5' overhangs on both DNA strands (Figure 6.9B). Resection of these overhangs could result in a 'TT' microhomology at the overhang termini (Figure 6.9C). Correct base pairing of this 'TT' microhomology would result in non-mutagenic repair and restoration of the wild type allele (Figure 6.9D-F, left). However, if the single terminal 'T' nucleotide was to pair with the

terminal ‘A’ on the complementary strand, this microhomology-directed repair would result in a single ‘T’ insertion proximal to the ‘TT’ microhomology domain (Figure 6.9D-F, right). It would be interesting to test this model by designing Cpf1 guides to target loci with microhomology domains within the two Cpf1 cleavage sites. If a similarly high frequency of non-random homozygous mutations were to be obtained in these cases, it would provide support for this model of microhomology-directed mutation. Proof of this concept could be of potential value, both to help design guides with higher mutagenic efficiencies and also to enable deterministic (non-random) mutations.

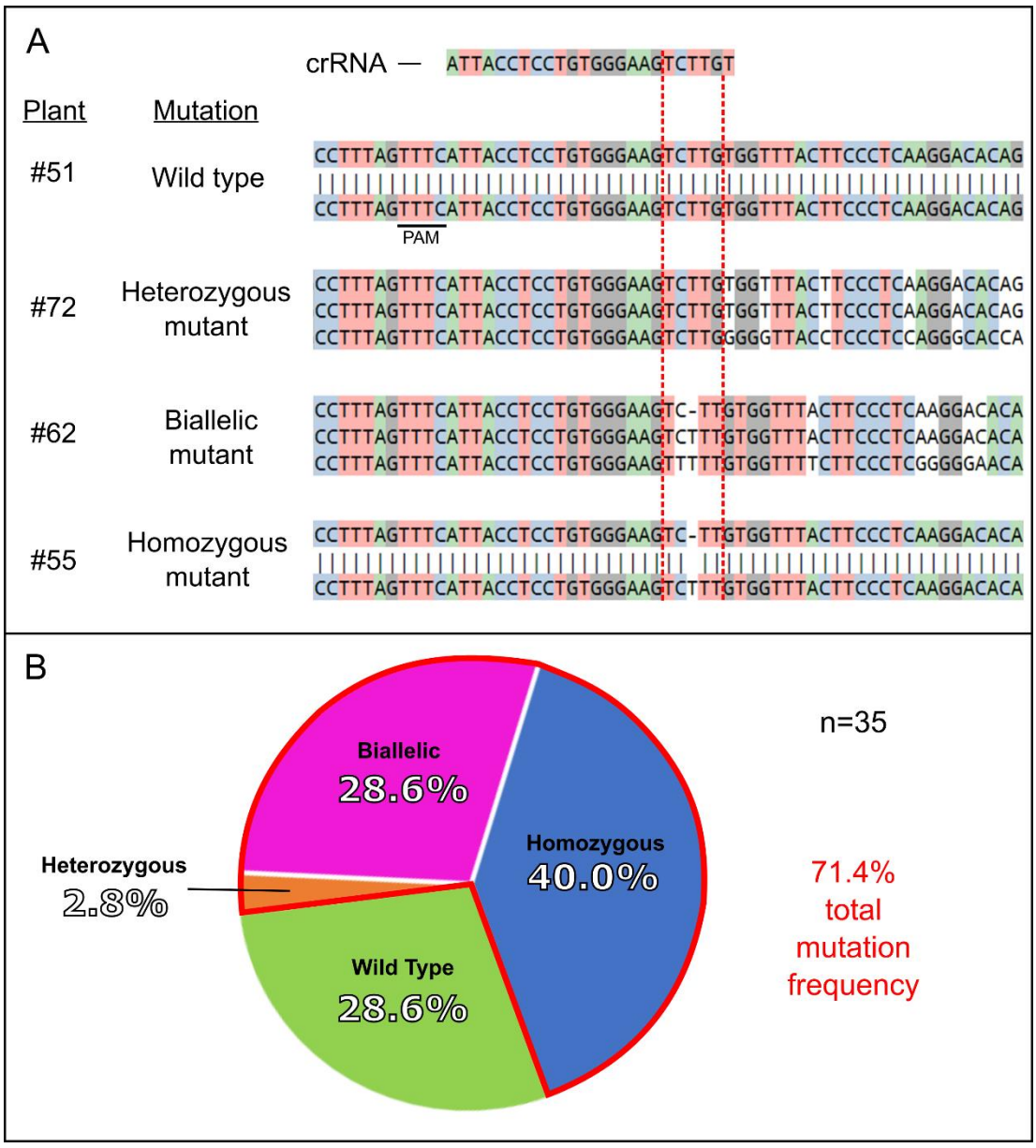

Figure 6.8 Summary of Cpf1-induced mutations at the *ARF19* locus in the T_1 generation. **A** Sequence alignments using CRISP-ID online software. The guide sequence of the crRNA is indicated and the PAM sequence is underlined. Red dashed lines indicate the predicted Cpf1 cleavage sites. A representative alignment for each mutation type is shown. **B** Pie chart summary of the *arf19* mutations recovered from 35 T_1 plants.

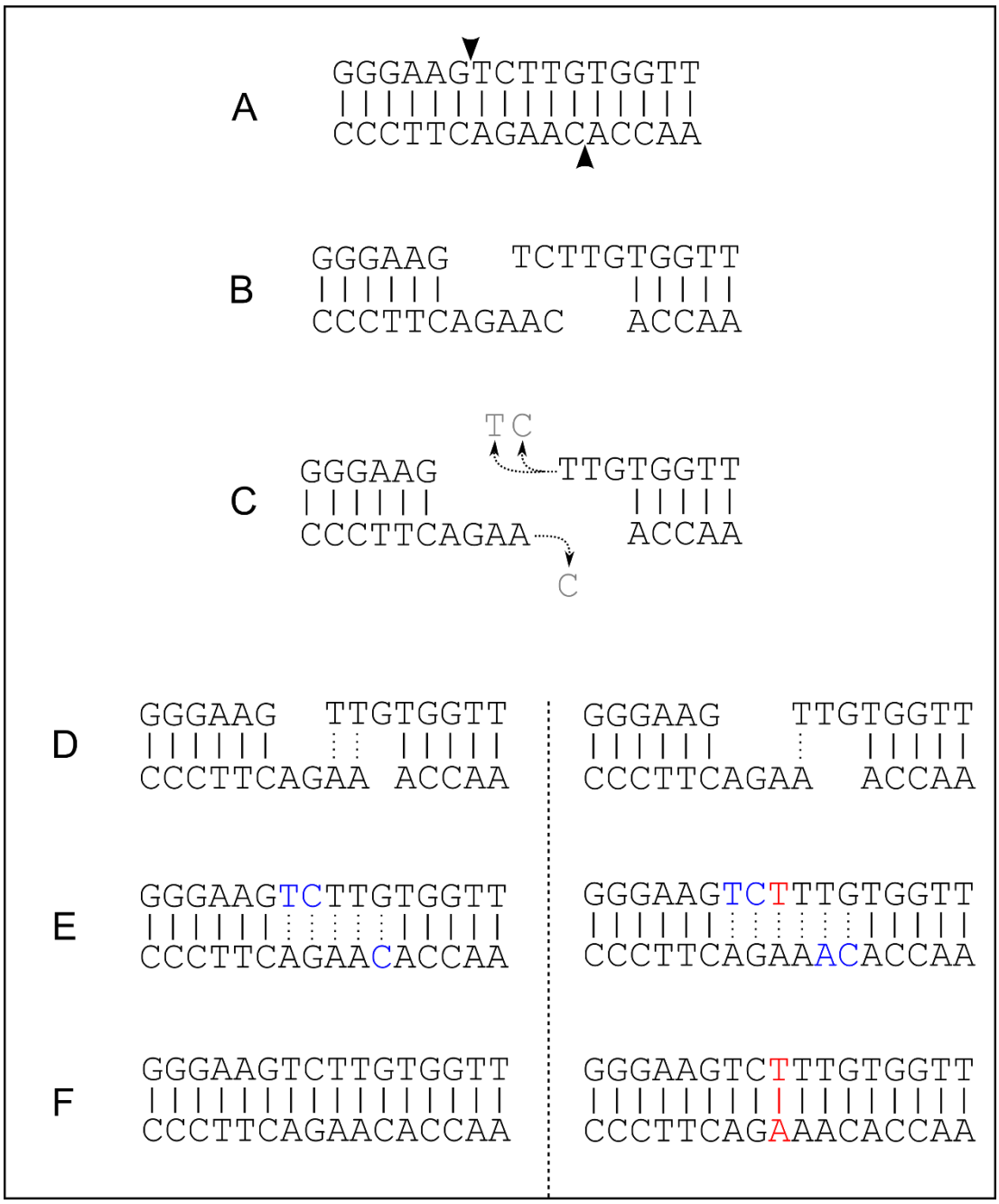

Figure 6.9 Model explaining the frequent homozygous 'T insertion' mutation. **A** Cpf1 cleavage sites at the target locus are shown by black arrowheads. **B** Staggered Cpf1-induced DSBs results DNA ends with compatible 5' overhangs. **C** Resection of the 5' overhangs. **D,E,F(left)** Correct base pairing results in DNA repair restoring the wild type allele. **D,E,F(right)** Base paring of the terminal nucleotides results in a slippage-repair scenario, introducing an additional A/T base pair in between the two Cpf1 cleavage sites. Nucleotides restoring the wild type sequence are shown in blue, extraneous nucleotides (resulting in an insertion mutation) are shown in red.

To confirm that the homozygous 'T-insertion' mutations (hereafter referred to as *arf19-T*) identified in the T_1 generation were stable and heritable, one of the homozygous mutant T_1 seedlings was used to produce T_2 seed. Non-red-fluorescent seeds were selected to generate a non-transgenic population of T_2 plants. Segregation of the transgene at this stage ensures that additional CRISPR/Cpf1-induced mutations cannot occur in the T_1 germ cells or T_2 somatic cells, which would confound the genetic analysis of the T_2 population. *ARF19_1* target amplicons for 8 randomly selected non-transgenic T_2 homozygous *arf19-T* mutant progeny were sequenced. All 8 sequences harboured the same 'T-insertion' mutation as the parental plant, indicating that this homozygous mutation is stable and heritable. In parallel, 30 of these non-transgenic T_2 *arf19-T* mutants, were planted on 2,4-D agar plates, alongside a non-transgenic T_2 population from a wild type T_1 plant. After 7 days of growth in constant darkness, the mean hypocotyl lengths for the *arf19-T* mutant and wild type populations were 8.43mm and 3.99mm, respectively, reflecting a more than 2-fold statistically significant difference ($t_{59} = 19.253$, $p = 2.2e10^{-16}$). This 2,4-D insensitive phenotype of the *arf19-T* mutant population provides further evidence that the CRISPR/Cpf1-induced mutations were heritable and resulted in ARF19 loss of function. Interestingly, plotting the hypocotyl lengths as boxplots revealed that the two populations had non-overlapping normal distributions (Figure 6.10). Further experiments are required however, to determine whether the mutation at this locus could be reliably used for 'haploinsufficiency screening' to test the efficiency of CRISPR/Cpf1 editing technologies.

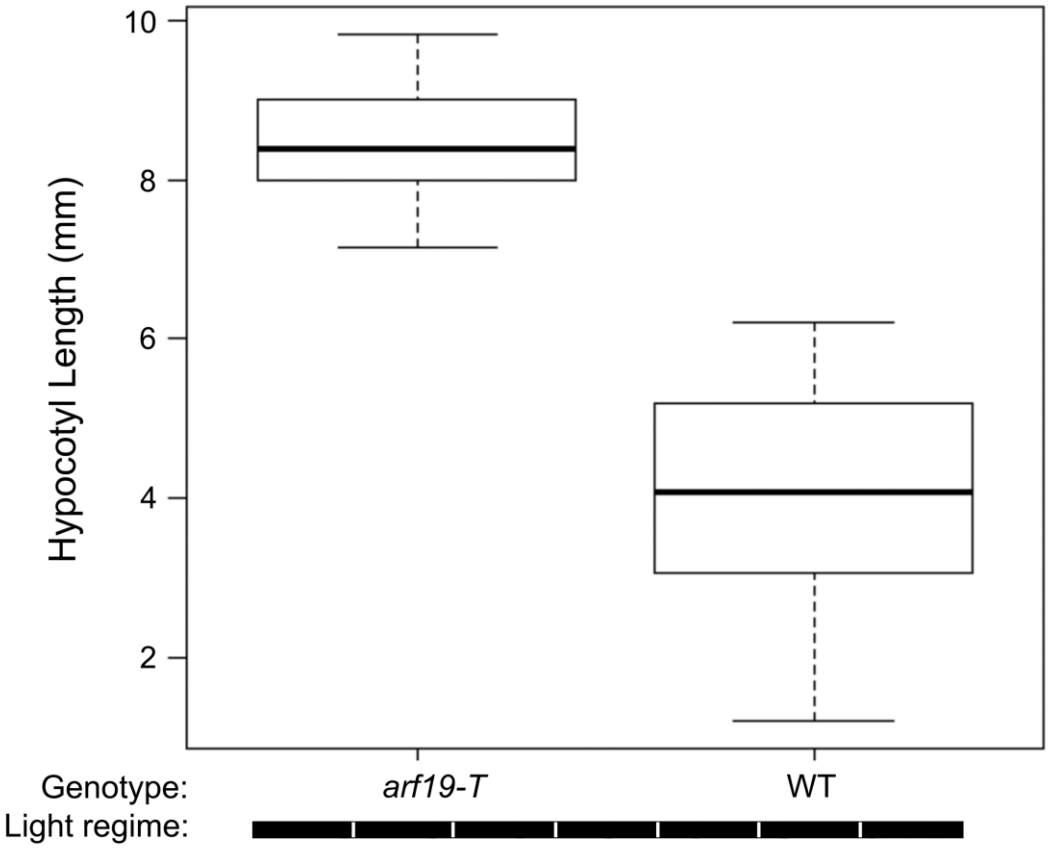

Figure 6.10 Quantitative assessment of the *arf19-T* mutant hypocotyl phenotype. Box plots indicate the median hypocotyl lengths for Wild Type (WT) and homozygous *arf19-T* mutant seedlings in the T_2 generation, grown on agar plates containing 1µM 2,4-D. Error bars represent the range in hypocotyl lengths for approximately 30 seedlings. Seeds were stratified for 3 days at 4ºC in darkness and given a 5-hour light pulse of 200 µE white light to induce germination. Seedlings were kept in complete darkness for 7 days before measuring hypocotyl lengths.

6.3.6 Assessing the determinants for efficient Cpf1 crRNAs

It was unexpected to observe such a dramatic difference in mutagenic efficiency between the two Cpf1 crRNAs targeting the *ARF19* locus (71.4% and 0% editing efficiency for arf19_1 and arf19_2, respectively). As these two crRNAs targeted near-by regions at the same genomic locus, it seems unlikely that these differences were caused by variation in the target DNA, such as the heterochromatin/euchromatin state of the two target regions. While epigenetic modifications at the target loci cannot be formally ruled out, it seems more likely that the differences are caused by the variation between the two crRNA sequences. The specific features of crRNA sequences that influence the editing efficiency of Cpf1 in plants are unknown. Recently, a high-through-put screen was conducted to measure the mutagenic frequencies for thousands of AsCpf1 crRNAs in mammalian cells (Kim *et al.*, 2017b). From this large data set, a software was designed to enable prediction of mutation frequencies based on the sequence of the crRNA and associated PAM sequence (Kim *et al*., 2017b). To assess whether this predictive software could be beneficial for rational design of Cpf1 crRNAs for plant editing purposes, data from published Cpf1-editing experiments in plants (Hu *et al*., 2017, Wang *et al*., 2017b, Xu *et al.*, 2017) were collated. The crRNA sequences and their respective PAMs were used to generate predicted mutation frequencies, using the published software developed for mammalian cells (http://big.hanyang.ac.kr/cindel). As the published predictive model requires crRNAs to have a GC content in the range of 30-70% (Kim *et al*., 2017b), GC contents for each of the crRNAs were quantified to ensure that this criterion was satisfied. The experimentally measured (observed) mutation frequencies for each of the Cpf1 crRNAs were extracted from the relevant publications. These data are summarised in Table 6.4.

crRNA name	Cpf1 Orthologue	Predicted efficiency (%)	Observed efficiency (%)	GC content (%)	Reference
OsRLK-798	Fn	7.5	75.0	56.5	(Wang *et al*., 2017b)
OsRLK-799	Fn	32.8	62.5	60.9	
OsRLK-802	Fn	9.5	65.6	60.9	
OsRLK-803	Fn	37.3	43.8	65.2	
OsBEL-230	Lb	25.2	60.0	43.5	
OsBEL-240	Lb	17.0	53.3	47.8	
OsBEL-250	Lb	7.4	40	47.8	
OsBEL-260	Lb	34.3	56.7	52.2	
C1	Lb	32.7	12.1	43.5	(Hu *et al*., 2017)
C2	Lb	33.8	5.0	43.5	
C4	Lb	16.1	0.0	43.5	
C5	Lb	15.7	5.0	52.2	
OsBEL	Lb	39.5	41.2	52.2	(Xu *et al*., 2017)
OsPDS	Lb	23.7085	21.4	43.5	
arf19_1	Lb	10.6641	71.4	47.8	(Pyott, unpublished)
arf19_2	Lb	16.5448	0.0	39.1	

Table 6.4 List of crRNAs used for *in planta* genome editing studies. Observed efficiencies reflects the proportion of plants transformed with the relevant Cpf1 + crRNA construct containing genomic mutations of any type (heterozygous/biallelic/homozygous). Predicted efficiencies and GC contents for each crRNA sequence were obtained using the following software http://big.hanyang.ac.kr/cindel (Kim *et al*., 2017b).

To determine whether the software can be used to accurately predict the *in planta* editing efficiency of Cpf1 crRNAs, scatterplots and linear regression analyses were conducted to assess the strength of correlation between the predicted and observed mutation frequencies for the crRNAs presented in Table 6.4. The R^2 value provides a means to statistically quantify the correlation between two variables using a linear regression model. Briefly, the 0-1 range in R^2 values reflects the range of 0%-100% correlation between the two variables. Plotting the predicted and observed mutation frequencies for all of the published crRNAs from Table 6.4 together indicated a very low correlation (R^2=0.006953) between the predicted and observed efficiency values (Figure 6.11A). It is possible that differences in methods used to score the observed mutation frequencies in the various publications may confound this global correlative analysis. To resolve this, individual correlation analyses were performed, grouping data according to their respective publications. Furthermore, one publication (Wang *et al*., 2017b) provided data for both LbCpf1 and Francisella-(Fn)Cpf1. These data were divided according to the two Cpf1 orthologues to prevent this variation confounding the correlation analysis. While dividing the data in this way is likely to reduce errors caused by experimental/biological variation it considerably reduces the statistical power of the analyses as each correlative assessment is based on only a few data points. For the FnCpf1 data, a negative correlation was observed indicating that the model based on AsCpf1 activity in mammalian cells does not accurately predict the *in planta* efficiency of crRNAs for FnCpf1 (Figure 6.11B). The LbCpf1 data sets containing more than 2 data points each exhibited a positive correlation between the predicted and observed efficiencies with R^2 values of 0.706 and 0.5625 (Figure 6.11C,D), indicating that the software may be at least partially competent for predicting the *in planta* editing efficiencies of LbCpf1 crRNAs. For data sets with only 2 data points, it is not possible to compute a meaningful R^2 value. Nevertheless, a positive correlation can be observed in the scatterplot for the LbCpf1 data from (Xu *et al*., 2017)(Figure 6.11E). Moreover, close inspection of the corresponding values in Table 6.4 reveals that the predicted and observed efficiencies for this data set were very closely matched, although statistical support for this observation is lacking due to the small data set. Strangely, there was a negative correlation for the two *ARF19*-targeting crRNAs used in this study (Figure 6.11F). This is possibly because the proposed microhomology-

mediated, non-random mutations observed for arf19_1 resulted in a much higher observed mutation frequency than could be predicted by the software, which does not take microhomology domains into account. Another noteworthy source of variation between the data sets is the method used to express and process the crRNA. Most of the data sets (Figure 6.11 B,C,D,E) delivered crRNAs as polycistrons, whereas the *ARF19*-targeting crRNAs used in this study were delivered as mature crRNAs. Recent reports have suggested that delivering crRNAs as polycistronic sequences results in higher *in planta* mutation frequencies compared to delivering mature crRNAs (Hu *et al*., 2017, Wang *et al*., 2017b).

.

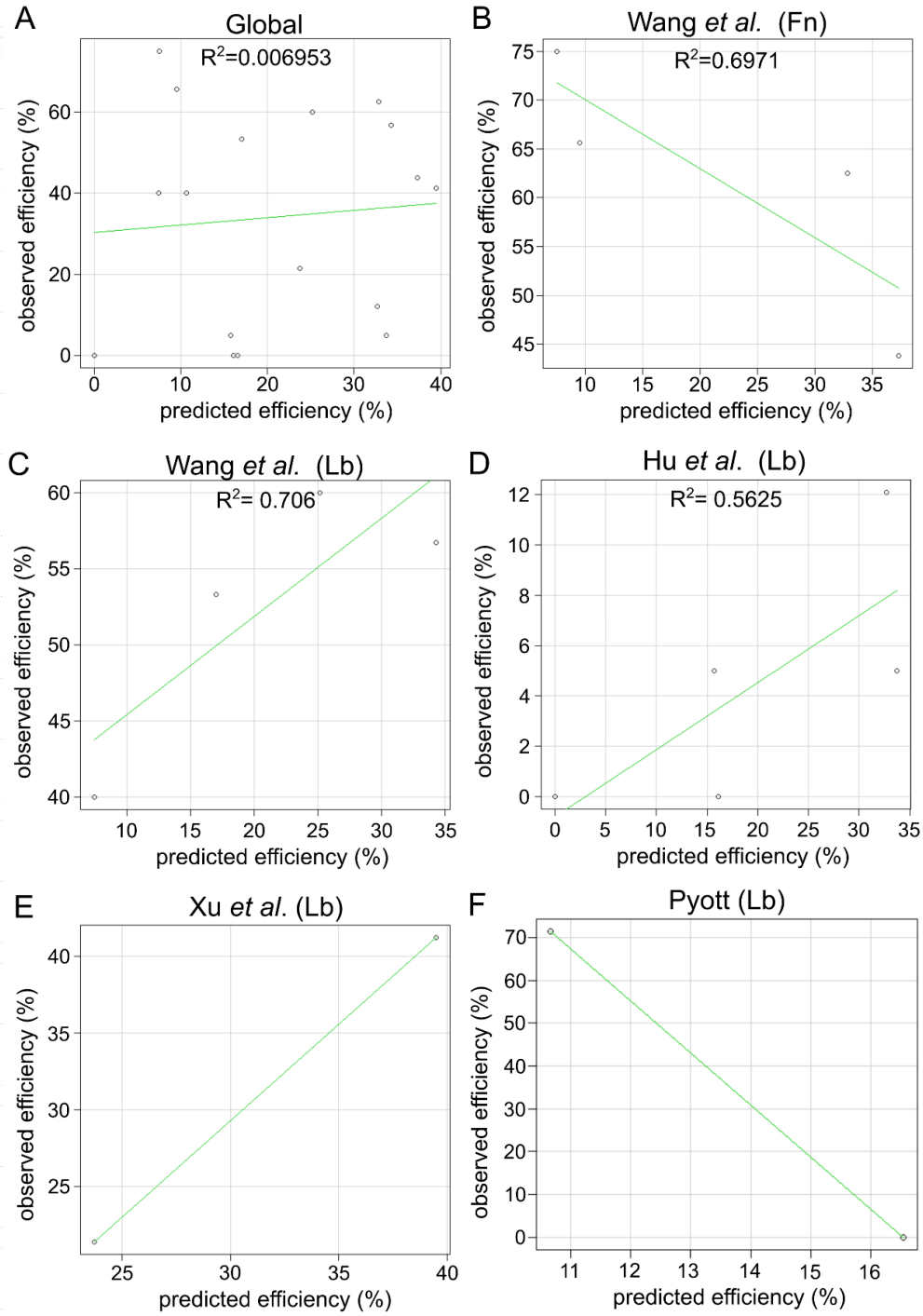

Figure 6.11 Scatter plots for assessing correlations between the observed and predicted *in planta* editing efficiencies for Cpf1 crRNAs. A linear regression function (green line) was used to statistically assess the correlation between the two variables. R^2 values for each linear regression function are indicated within the plot axes. **A** Global analysis of all crRNA sequences from Table 6.4. **B-F** Local analyses grouping crRNAs according to their respective publications and Cpf1 orthologues.

To further investigate features of the Cpf1 crRNA sequences which may act as editing efficiency determinants, the secondary structures of full crRNA sequences (direct repeats and spacers) were analysed using an online tool (http://rna.tbi.univie.ac.at/) that predicts RNA secondary structures based on minimum free energy calculations (Gruber *et al.*, 2008). The sequences for LbCpf1 crRNAs presented in Table 6.4 were used for this analysis. The FnCpf1 crRNAs were omitted because differences in the direct repeats of LbCpf1 and FnCpf1 crRNAs would complicate the analysis. Similarly, the crRNAs which were processed using a tRNA-based polycistron system (Hu *et al*., 2017) were not analysed, as the requirement for the RNAse P/Z recognition sequences within the crRNA could affect the secondary structure and hence complicate interpretations. Diagrammatic representations of the secondary structures for the remaining crRNAs are presented in Figure 6.12. The crRNAs were grouped according to the types of secondary structures they contained. The first group (Figure 6.12A, B,C,D) contained a hairpin structure within the guide sequence (highlighted green) of the crRNA. The second group (Figure 6.12E,F,G) was defined by an extensive fold-back structure formed by pairing between the direct repeat leader sequence (highlighted red) and the guide sequence (highlighted green). The final group (Figure 6.12H) exhibited an extended hairpin structure in the direct repeat (highlighted blue) due to base pairing between the start of the direct repeat sequence and the start of the guide sequence. Each of the groups contained crRNAs with both high and low *in planta* mutation efficiencies, and there was no significant correlation (R^2=0.007144) between the minimum free energy of folded RNA structures and the observed editing efficiency. These results indicate that the secondary structures for this set of LbCpf1 crRNAs are not major factors determining the efficiency at which they induce mutations *in planta*.

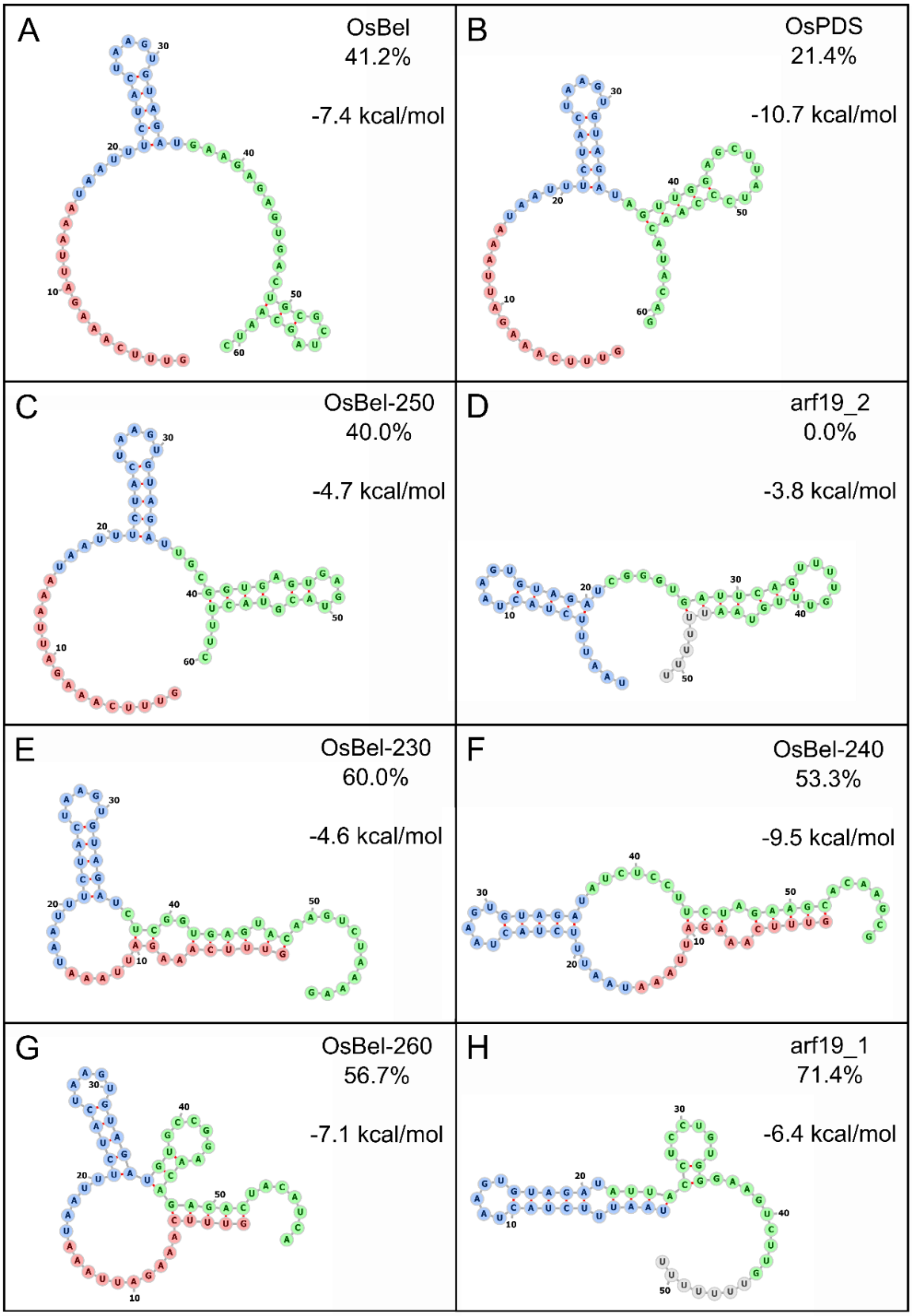

Figure 6.12 *(Figure legend overleaf)*

Figure 6.12 Predicted secondary structures for LbCpf1 crRNAs presented in Table 6.4 using online 'RNA-fold' software (http://rna.tbi.univie.ac.at/). The repeat-leader and stem-loop regions of the crRNA direct repeats are highlighted in red and blue, respectively. The guide sequence of each crRNA is highlighted in green. The *ARF19*-targeting crRNAs (**D**,**H**) lack the repeat leader sequence (red) as they were expressed as mature crRNAs. Additionally, they contain a 6X polyU tail (grey) at their 3'terminus as a result of RNA polymerase III transcription of the mature crRNA. The observed editing frequencies (as listed in Table 6.4) and minimum free energy values are indicated for each crRNA.

Taken together, the currently available software for designing Cpf1 crRNAs may not be appropriate for FnCpf1 crRNAs but has limited predictive capacity for LbCpf1 crRNAs that are delivered as polycistronic sequences. However, there may still be several determinants affecting the *in planta* editing efficiency of Cpf1 crRNAs that are yet to be deduced. The preliminary analysis presented herein suggests that the secondary structure of the crRNA is not a major determinant for editing efficiency. An interesting avenue for future research would be to conduct a high-throughput screen to measure crRNA editing efficiencies in plant cells, to allow plant-specific predictive software to be developed.

6.4 Microhomology-mediated targeted DNA replacement using Cpf1

The putative microhomology-mediated DNA repair resulting in the 'T-insertion' mutation for the arf19_1 target led to speculation that Cpf1-induced microhomology-mediated repair could be exploited for efficient targeted KI of exogenous DNA. To explore the potential of this novel DNA KI technique, a vector was prepared with the aim of delivering an exogenous 'donor DNA' sequence from a binary vector into the *GFP* locus of the 16c line of *N. benthamiana*. A pair of LbCpf1 guides targeting the *GFP* locus were cloned, each individually driven by *U6* promoters, into a pK7FWG2 vector containing the LbCpf1 coding sequence, flanked by *UBIQUITIN* promoter/terminator sequences (Figure 6.13A). The *GFP*-targeting crRNAs were designed such that DSBs created by the pair of crRNAs would excise a fragment of

approximately 280bp from the *GFP* locus. Into the same vector, a ~100bp random sequence was cloned to act as a 'donor DNA' substrate. This donor DNA was flanked by rationally designed target sequences to allow its excision by the action of another pair of LbCpf1 crRNAs, which were also cloned into the same vector, each individually driven by *U6* promoters. These crRNAs and their targets flanking the donor DNA were designed to create cohesive ends between the excised donor DNA fragment and the two staggered DSBs created by the *GFP*-targeting crRNAs (Figure 6.13). Importantly, the *GFP*-targeting crRNAs were designed to target regions between the two Cpf1-induced DSBs (Figure 6.13B). Similarly, the donor-DNA-targeting crRNAs were designed such that the target sequences were distal to the Cpf1-excised donor sequence (Figure 6.13C). This was to ensure that integration of the donor DNA to the *GFP* target locus would result in a stable recombinant KI sequence without target sequences complementary to the guides of any of the 4 crRNAs (Figure 6.13D).

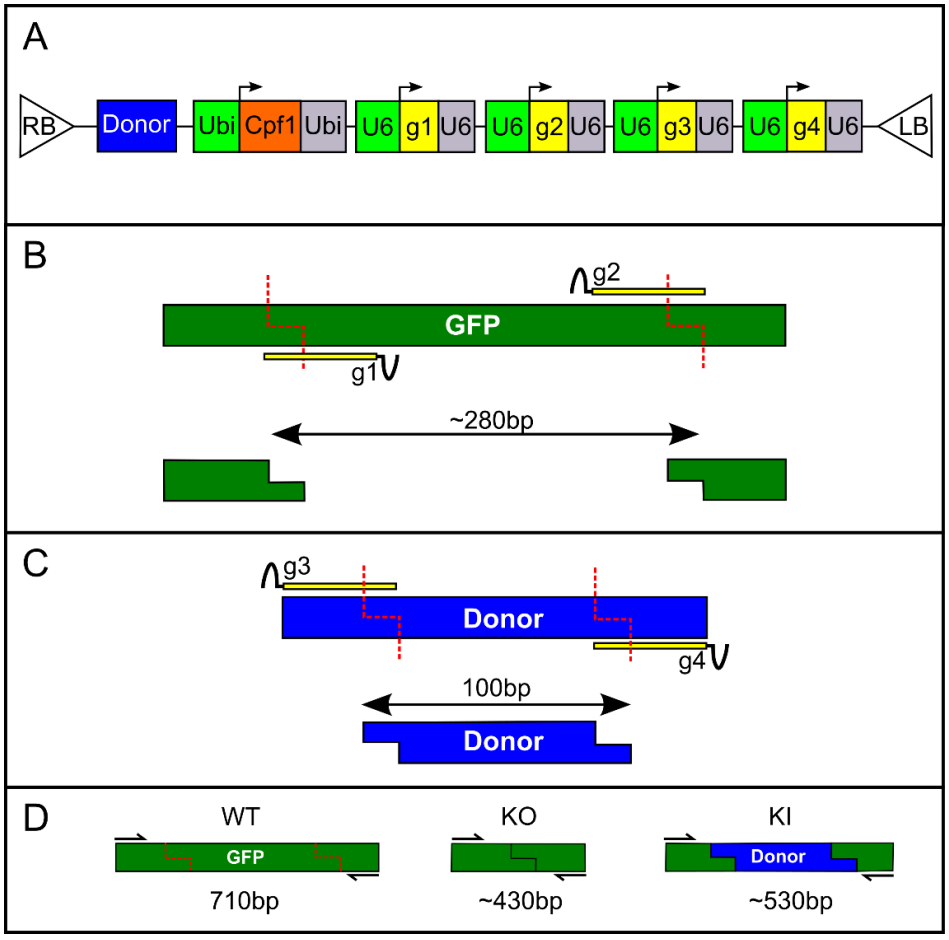

Figure 6.13 Schematic of the strategy for targeted KI of a donor DNA sequence into the GFP locus of *N. benthamiana* 16c genomes. **A** LbCpf1, donor DNA, and a cassette of four crRNAs are delivered as a single T-DNA. LB/RB denote the left/right border sequences that demarcate the T-DNA. Promoter/terminator sequences are shown in green/grey, respectively. *Ubi* = *UBIQUITIN*. crRNA sequences (g1-4) are shown in yellow. Donor DNA is shown in blue. Arrows indicate transcriptional start sites. **B** 2 crRNAs (g1/g2) guide Cpf1 to excise a ~280bp fragment from the middle of the *GFP* ORF. crRNAs were designed such that their target sequences and PAMs are removed from the *GFP* locus by excision of the ~280bp fragment. **C** 2 crRNAs (g3/g4) guide Cpf1 to excise the donor DNA from the integrated T-DNA sequence. These crRNAs were designed such that their target sequences and PAMs were not present on the excised donor DNA. The donor DNA and g3/4 crRNAs were designed to generate compatible overhangs to the staggered Cpf1-induced DSBs generated at the *GFP* locus by g1/g2 crRNAs. **D** Illustration of the PCR assay used to distinguish WT, KO, and KI alleles, whereby a pair of distal primers would result in 710bp, ~430bp, and ~530bp PCR amplicons for WT, KO, and KI alleles, respectively. Primer positions are indicated by half arrows.

This ‘KI inducer’ construct (Figure 6.13A) was delivered into leaves of 4-week old 16c plants by syringe infiltration with *Agrobacterium tumefaciens*. Infiltrated tissue was collected at 7dpi to analyse KI and KO editing events from extracted DNA. A PCR assay was devised to be able to discern: wild type alleles, alleles with a ~280bp sequenced excised (KO), and alleles with 100bp donor sequence inserted into the *GFP* locus (KI). Primers flanking the two *GFP*-crRNA target sites were designed for this assay. PCR amplification with these primers would yield a 710bp amplicon from wild type alleles. The same primer pair would result in a ~430bp PCR product for KO alleles due to the excision of a ~280bp fragment in the template DNA. Amplification from KI alleles would result in a ~530bp amplicon due to the excision of a ~280bp fragment, and sequential insertion of a 100bp donor DNA fragment (Figure 6.13D). Performing this PCR assay for the infiltrated samples yielded the expected ~700bp wild type amplicon for all the samples. Additionally, a ~430bp fragment was visible for the sample taken from plants infiltrated with constructs containing all 4 crRNAs, or all 4 crRNAs and the donor DNA sequence (Figure 6.14). This indicates that a KO allele had been created by the excision of a ~280bp fragment in between the 2 *GFP*-targeting crRNAs. No evidence for a KI mutation event was obtained by this assay, however.

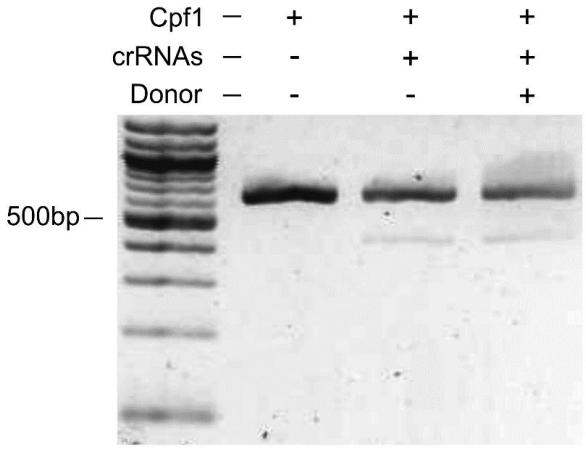

Figure 6.14 PCR assay to distinguish wild type/KO/KI alleles of *GFP* in 16c *N. benthamiana* plants. Constructs expressing either; Cpf1 alone, Cpf1 + all four crRNAs, or Cpf1 + all four crRNAs + donor DNA, were infiltrated into separate leaves of 16c plants as controls for the three different wild type/KO/KI allelic variants. DNA was extracted from infiltrated patches at 7dpi. Products for the diagnostic PCR outlined in Figure 6.13D were separated on a 2% agarose gel, alongside a 100bp DNA ladder.

To confirm the interpretation from the PCR assay, the ~430bp fragment was gel-purified and cloned into a pGEM-T-Easy (Promega Catalogue# A1360) vector to allow the fragment to be sequenced by Sanger sequencing. Sequence analysis of this cloned fragment confirmed that the KO allele resulted from the excision of a 279bp sequence by the dual action of the 2 *GFP*-targeting crRNAs (Figure 6.15A). Hence, this truncated *GFP* allele was named Δ279. Interestingly, close inspection of the Δ279 allele suggests that it may have been generated by microhomology directed repair (Figure 6.15B-F), analogous to that previously described for the LbCpf1-induced *arf19-T* mutation (Figure 6.9). Furthermore, 4/5 of the sequenced clones were genetically identical. Interestingly, the single clone that differed from the other 4 harboured a repeat of the 'CCTTA' overhang at the g1 crRNA target site. These data indicate that the putative Cpf1-induced MMEJ repair pathway results in non-random mutations, as previously observed for the *arf19-T* mutants.

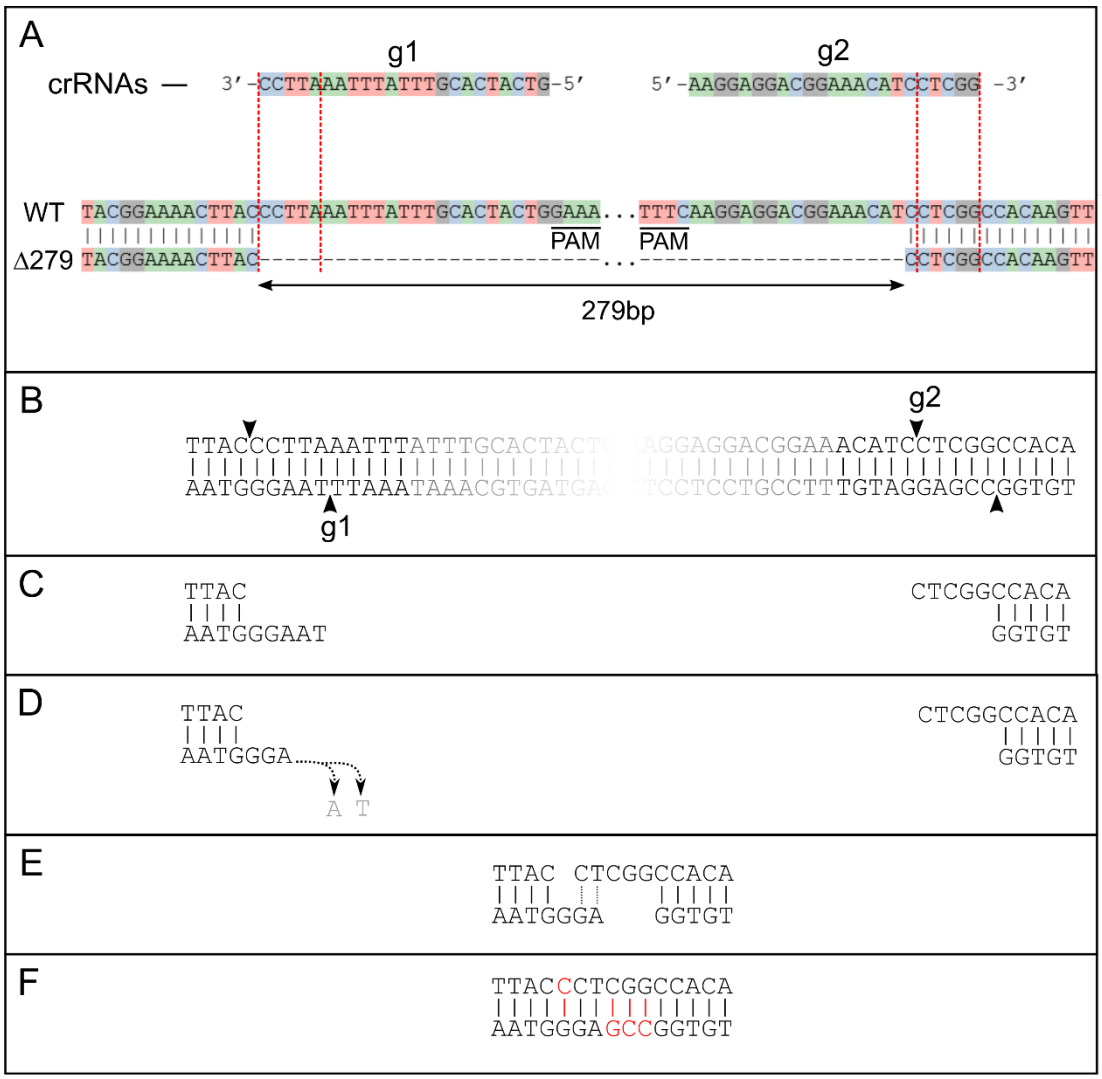

Figure 6.15 Generation of a *GFP* KO allele (Δ279) by Cpf1+(g1/g2)crRNA-induced DSBs at the *GFP* locus. **A** Sequence alignment of the truncated *GFP* allele, to the Wild Type (WT) allele sequence. The guide sequences of the g1/g2 crRNAs are indicated and their respective PAM sequences are underlined. Red dashed lines indicate the predicted Cpf1 cleavage sites for the g1/g2 targets. **B-F** Proposed model of MMEJ-directed DNA repair generating the Δ279 truncated *GFP* allele. **B** Arrow heads represent the predicted Cpf1 cleavage sites for g1/g2 crRNA targets, flanking a ~280bp fragment of the *GFP* ORF (shaded white). **C** A pair of Cpf1-induced DSBs excise a 279bp sequence from the *GFP* ORF, leaving 5' overhangs either side of the excised fragment. **D** Resection of one of the overhangs results in a dinucleotide microhomology domain between the two flanking 5' overhangs. MMEJ DNA repair results in ligation of the DNA ends flanking the excised 279bp fragment, creating the recombinant *GFP* sequence observed (Figure 6.15A).

To test whether the lack detectable of KI events could have occurred because the donor DNA was not sufficiently excised from the T-DNA an additional PCR assay was performed. For this, primers flanking the donor DNA sequence were designed. PCR from un-edited T-DNA would result in a 486bp amplicon. Alternatively, PCR amplification of T-DNA in which the donor DNA fragment had been excised would result in a 365bp amplicon (Figure 6.16A). The presence of the shorter 365bp amplicon (in addition to the 486bp amplicon) in the DNA extracted from a plant infiltrated with a construct containing both the donor DNA and all 4 crRNAs, indicates that excision of this donor DNA had occurred *in planta* (Figure 6.16B). As a negative control, the same PCR was performed using the plasmid used to deliver the T-DNA.

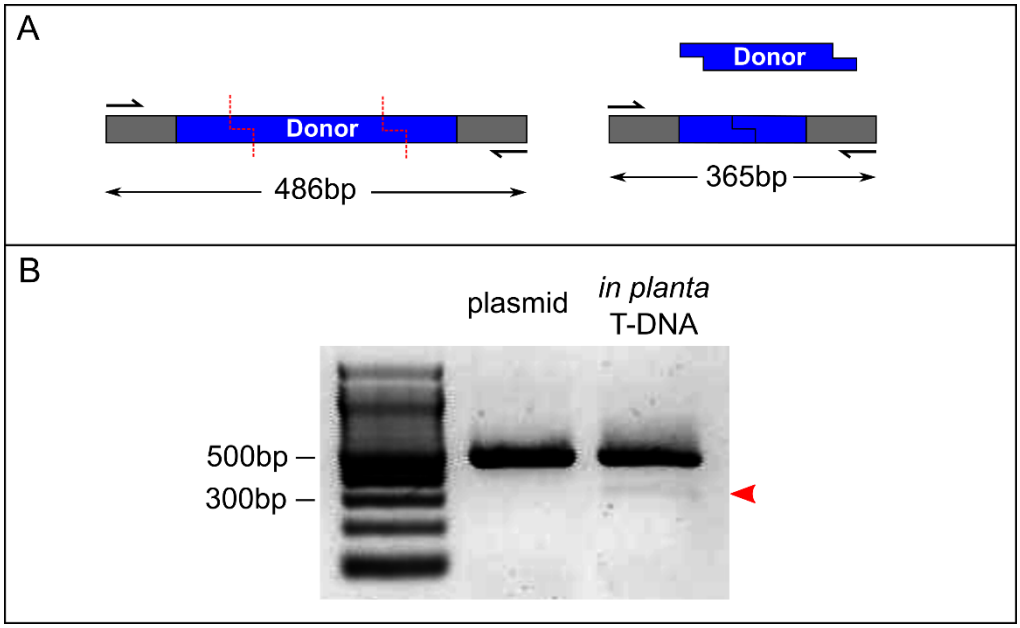

Figure 6.16 PCR assay to assess excision of the donor DNA from the T-DNA. **A** Primers distal to the predicted Cpf1 cleavage sites (red dashed lines) result in a 486bp amplicon for DNA templates where no excision of the donor DNA occurred. Excision of the donor DNA sequence, and subsequent ligation of the flanking DNA ends, results in a 365bp PCR product using the same primer pair. Primer positions are marked by half-arrows. Donor DNA is shown in blue and flanking T-DNA is shown in grey. **B** PCR products resolves on a 2% agarose gel, alongside a 100bp DNA ladder. A control PCR was performed with the binary vector (plasmid) used to deliver the T-DNA (as shown in Figure 6.13A). Red arrow head indicates the faint ~365bp product indicative of donor DNA excision.

Results from this experiment indicate that further technological adjustments are required to fully explore the possibility of using microhomology-mediated repair pathways to deliver exogenous DNA to plant genomes. It is likely that the lack of KI events is a result of an insufficient quantity of donor DNA being present in close proximity to the staggered DSBs at the *GFP* target sites. Therefore, future experiments should aim to increase the abundance of donor DNA to further test whether a similar design of microhomology-mediated repair could offer an efficient method for targeted KI events. Flanking the donor DNA sequence by trans-acting transposon replication elements could offer a method for amplifying the donor DNA. Alternatively, DNA viruses could be used to deliver donor DNA in high quantities. This method has already been utilised for targeted KI of exogenous DNA in plant genomes (Baltes *et al*., 2014, Wang *et al*., 2017a), though not in conjunction with microhomology-directed DNA repair.

6.5 Chapter Discussion

6.5.1 The LbCpf1 orthologue can induce heritable mutations in plants

This chapter set out to explore the potential of a novel CRISPR-associated endonuclease, Cpf1, for plant biotechnological applications. First, by comparing the two Cpf1 orthologues which had previously been shown to induce mutations in mammalian cells (Zetsche *et al*., 2015), it was established that LbCpf1 is superior to AsCpf1 for *in planta* CRISPR/Cpf1 genome editing (Figure 6.3). This result has been independently confirmed by other research groups (Hu *et al*., 2017, Kim *et al*., 2017a, Tang *et al*., 2017). Next, mutagenic screen was devised to compare the frequency of heritable mutations produced by SpCas9- and LbCpf1-based editing systems. It was initially hoped that the haploinsufficient phenotype resulting from mutations at *ARF19* would enable a sensitive and high-thoughput screening method to compare SpCas9- and LbCpf1-editing efficiencies. However, testing this phenotype using the previously described (Okushima *et al*., 2005) *arf19-1* homozygous mutant indicated that the haploinsufficient 2,4-D insensitivity phenotype was not reliable enough, under the assay conditions used, for accurate phenotypic screening. Therefore, DNA sequencing of T_1 seedlings was used to assess the mutagenic efficiencies of SpCas9 and LbCpf1.

The initial goal of comparing the activities of the two nucleases, by scoring the mutation frequencies for *ARF19*-targeting sgRNAs/crRNAs with nearly identical guide sequences, was not possible as neither of the 'SpCas9 + sgRNA' transgenic lines harboured any *ARF19* mutations. However, it was possible to conclude that, at least for certain guide sequences, LbCpf1 can induce heritable mutations at frequencies greater than any currently reported (using analogous methods) for SpCas9 systems. A model was devised to explain the high-frequency and non-random nature of the LbCpf1-induced mutations. This model proposed that microhomology regions within the staggered DSBs induced by Cpf1 could result in frequent, non-random mutations. If this proposed model is correct, this knowledge could help to improve the frequency and determinism of CRISPR/Cpf1 genome editing platforms, by designing crRNAs with microhomology domains within the target DNA sequence.

6.5.2 The function determinants of LbCpf1 crRNAs are yet to be adequately deduced

The high degree of variation in editing efficiency for different Cpf1 crRNAs, both reported herein and elsewhere (Hu *et al*., 2017, Kim *et al*., 2017a, Tang *et al*., 2017, Wang *et al*., 2017b, Xu *et al.*, 2017), indicates that gaining a better understanding of the functional determinants of LbCpf1 crRNAs is an important topic for future research. Software enabling the rational design of AsCpf1 crRNAs for genome editing in mammalian cells has recently been published (Kim *et al.*, 2017b). To test whether this software could be appropriated for designing LbCpf1 crRNAs for *in planta* genome editing, correlations between editing frequencies observed *in planta*, and the software-predicted frequencies for published Cpf1 crRNAs were assessed. While there was some correlation for some sets of published crRNAs (Figure 6.11), this analysis indicated that the published software is insufficient to accurately predict the efficiency of LbCpf1 crRNAs for *in planta* genome editing. This parallels previous reports describing considerable differences in the sequence requirements for designing efficient Cas9 sgRNAs for mammalian/plant applications (Doench *et al.*, 2014, Liang *et al.*, 2016). This highlights the need for a high-throughput comparison of *in planta* editing frequencies for multiple Cpf1 crRNAs, to enable plant-specific crRNA-design software to be developed.

6.5.3 Exploring the potential for microhomology directed sequence knock-ins

The prospect of LbCpf1-induced microhomology-directed DNA repair, observed for the *arf19_1* mutants generated in this study, led to speculation that such a DNA repair mechanism could be utilised as a novel method for targeted sequence replacements. To test this, a construct was created to attempt a targeted, microhomology-directed KI of a donor DNA sequence into the *GFP* locus that is stably integrated into the *N. benthamiana* (16c) genome. Pairs of crRNAs were designed to excise a fragment from the GFP locus, and concomitantly excise a donor DNA fragment from the vector. The donor DNA was designed such that LbCpf1 cleavage would result in staggered DSBs flanking the donor DNA with microhomology to the staggered ends generated by the *GFP*-targeting crRNAs. A PCR assay revealed that the pair of *GFP*-targeting crRNAs could efficiently excise a fragment of 279bp between the 2 LbCpf1-induced DSBs, generating a truncated *GFP* allele named Δ279 (Figure 6.15). Sanger sequencing of this Δ279 *GFP* allele confirmed that it was generated by the repair of the terminal sequences flanking the fragment excised 279bp fragment. Moreover, close inspection of the repair junction suggests that microhomology-mediated repair may have been involved in generating this KO Δ279 *GFP* allele. No evidence for successful KI of donor DNA at the target site was obtained however, despite confirming that the excision of the donor DNA fragment from the T-DNA occurred (Figure 6.16). Though entirely speculative, it is likely that the lack of KI events is a result of insufficient quantities of excised donor DNA. It would therefore be interesting in the future to attempt a similar strategy of microhomology-mediated KI with methods that increase the abundance of donor DNA templates. Replicative systems such as DNA viruses or transposon sequences could be used to this end, or alternatively biolistic delivery of DNA templates could be employed.

Taken together, this chapter has explored some of the basic features of the novel nuclease, Cpf1, though results suggest that our understanding of its *in planta* behaviour (particularly relating to crRNA design features) is still in its infancy and deserves significant attention. The intriguing possibility that the staggered DSBs generated by Cpf1 may result in a DNA repair mechanism somewhat distinct from classical NHEJ or HDR, namely microhomology directed repair, is an attractive feature of Cpf1-based genome-editing as it may result in high-frequency, non-random mutations. Further

validation of this putative mechanism is needed to fully explore its utility for plant biotechnological purposes.

Chapter 7: Conclusions and future outlook

7.1 The importance of potyvirus research in the context of an expanding global population

The world's rapidly growing human population will be met with several challenges in the coming decades. Ensuring worldwide food security is one of these challenges that has already become a high priority topic on the global agenda. To face this issue, societies, particularly in 'developed' countries, should be prepared to alter several aspects of their consumer and agricultural practices. In addition to this, there is (more than ever) an appreciable importance for plant science research. Within the broad field of plant science, research into plant virus diseases is an agriculturally relevant topic due to the ubiquity and severity of virus diseases in crops. Of all plant viruses, the family *Potyviridae* is an important focus both because of its size (number of species) and the agricultural relevance of certain potyvirus species (Gibbs & Ohshima, 2010).

This thesis endeavoured to gain a deeper understanding of natural potyvirus resistance mechanisms in plants, focussing both on viral RNA silencing pathways and a genetically recessive 'loss-of-susceptibility' resistance phenomenon. Ultimately a better understanding of these (and other) resistance mechanisms could help to ameliorate crop losses to virus diseases; by providing improved phytosanitary methods, and by biotechnological construction of resistance alleles in crops.

7.2 The effects of temperature on the potyvirus synthesis/degradation equilibrium

Thermotherapy is a widely used phytosanitary method for inducing recovery from virus infections (Panattoni *et al.*, 2013). It is particularly important for crop varieties where identification and/or introgression of resistance alleles is difficult. Additionally, it is most frequently applied to perennial crops, or crops that are vegetative propagated as both of these scenarios are conducive to the long-term persistence of viruses. In contrast, crops that are planted annually from seeds can escape this ratcheting

accumulation of viruses as most virus species are not transmitted between generations in seeds (Regenmortel & Mahy, 2009).

The molecular nature of this thermotherapy phenomenon has been shown to derive from a progressive hyper-accumulation of vsiRNAs as ambient temperatures rise (Chellappan *et al*., 2005, Ma *et al*., 2016, Szittya *et al*., 2003, Velázquez *et al*., 2010, Qu *et al*., 2005). This results in a shift in the virus synthesis/degradation equilibrium towards net degradation, eventually resulting in elimination of the virus from plant tissue. Results presented in chapter 3 (using a TuMV-*Arabidopsis* model for potyvirus infection) stress the importance of considering both the temporal infection dynamics and the virus synthesis/degradation equilibrium concept, to correctly interpret the effects of ambient temperature changes on potyvirus infections. It appears that increasing growth temperatures over a 15^{o}C-27^{o}C range initially promotes the accumulation of TuMV, despite an increase in vsiRNA production. This could arise from increased activity of enzymes involved in virus synthesis and/or increased expression of host factors required for the potyvirus lifecycle (for example general changes in translational activity, discussed further below). It was determined that incubating plants at 27^{o}C for longer periods (greater than 1 week) could result in a shift in the virus synthesis/degradation equilibrium towards net degradation, and eventual recovery from infection. Furthermore, RDR1 and RDR6, which are both involved in amplification of vsiRNAs, were both found to be necessary for the temperature-induced hyperaccumulation of TuMV-vsiRNAs.

These results are relevant to agricultural potyvirus diseases for several reasons. Firstly, a careful consideration of temporal infection dynamics and the synthesis/degradation equilibrium concept is important to determine the optimal temperature and duration of thermotherapy treatments, which is likely to vary between different virus/host combinations. Secondly, a greater knowledge of the genetic determinants for efficient thermotherapy (*RDR1* and *RDR6* in the case of the TuMV-*Arabidopsis* system) could enable breeding efforts to maximise the efficiency of thermotherapy for crops/varieties in which this phytosanitary practice is commonly used. Finally, a detailed understanding of the nuanced effects of ambient temperature changes could contribute to sophisticated models to predict the severity of potyvirus infections in fields. Such models could help to mitigate crop losses by informing when or whether to plant

certain crops by evaluating the risk of losses to different viruses according to forecasted climatic conditions. Similar models have previously been developed and (to some extent) used in agriculture, though current models primarily focus on vector dynamics (Thackray *et al.*, 2004).

7.3 'Skototherapy' and biological relevance of the influence of light signals on viral RNA silencing defence pathways

Further to exploring the well-characterised thermotherapy phenomenon, using the TuMV-*Arabidopsis* system, intriguing observations were made concerning the effects of light intensity on vsiRNA-induced recovery from TuMV infection. Specifically, it was shown that incubating plants at very low light intensities (50µE) resulted in an increase in vsiRNA accumulation, eventually resulting in elimination of TuMV from systemic tissue. Interestingly, the same genetic loci involved in the thermotherapy response (*RDR1* and *RDR6*) were found to be necessary for the low-light-intensity-induced recovery from TuMV. These findings led to the proposal that skototherapy (improved viral immunity under low light conditions) could provide a novel phytosanitary method to use in conjunction with, or as an alternative to, the well-characterised practice of thermotherapy. Future research is required to assess the potential applications of these findings, however.

To gain some insight into the biological scenario(s) underpinning the artificially-induced skototherapy phenomenon, promoters of vsiRNA-related genes were analysed. A high abundance of light-regulated elements and circadian-clock regulatory motifs were found. Moreover, several of these vsiRNA-related genes exhibited rhythmic expression with a ~24hour period. The rhythmic expression for many of these genes persisted under constant light (LL) indicating circadian, as well as diurnal, control of these loci. The fact that the loci with oscillating expression patterns all had synchronous peaks at the end of the night period suggests that viral RNA silencing may have evolved to function primarily during the night or early morning. These results strongly suggest that the 'nocturnal' expression of vsiRNA-related genes may be the biological scenario that underpins the observed skoto-therapy phenomenon.

To assess the functional properties and biological relevance of the rhythmic expression of vsiRNA-related genes, TuMV inoculations were performed at different times across day/night and day/subjective-night cycles. These experiments showed that susceptibility to TuMV is regulated by exogenous day/night signals and also by the endogenous circadian oscillator. Using a genetic approach, it was deduced that the antiviral DCL functions of DCL2, DCL3 and/or DCL4 are dispensable during the day but form an important part of the night-time antiviral response. This strengthens the argument that viral RNA silencing has evolved a nocturnal peak of activity. However, these same genetic analyses implied that the nocturnal activity of viral RNA silencing is more likely a response to, rather than the determinant of, the oscillations in TuMV susceptibility. Therefore, identifying the factors responsible for generating the daily rhythms in TuMV susceptibility is an intriguing topic worthy of future research. The rhythmic expression of *eIF(iso)4E* (an essential host factor for TuMV proliferation) renders this gene an interesting candidate that may be at least partially responsible for this phenomenon. This hypothesis could be tested using the transgenic 35S:eIF(iso)4E over-expressor line that has been previously reported (Martínez-Silva *et al*., 2012). In parallel, future research should determine whether plants exhibit similar susceptibility oscillations to other viruses within *Potyviridae* and, perhaps more interestingly still, to more distantly-related virus species.

As previously mentioned, the reduction in TuMV-GFP RNA at low light intensities may be a result of both increased vsiRNA accumulation and also a general decrease in cellular transcriptional/translational activities. Indeed, it was recently shown that the translational activity of *Arabidopsis* rosettes varies in response to light cues across diurnal cycles. Specifically, polysome loading and protein synthesis were shown to be greater during the light period compared to the dark period and it was suggested that sucrose availability may be critical to this response (Pal *et al.*, 2013). These observations of diurnal oscillations in translational activity are interesting in the context of the observed oscillations in TuMV-GFP RNA accumulation. However, directly linking these two phenomena requires future research. In particular, it would be important to accurately assess the delay (or lag phase) between time of inoculation and the initiation of exponential synthesis of TuMV to draw meaningful conclusions about the effects of translational activity oscillations on TuMV susceptibility.

7.4 The potential for biotechnological advances to create novel potyvirus resistant germplasm

In addition to utilising plant immune responses (such as RNA silencing) for phytosanitary purposes, deployment of *R genes* in crop varieties is an important and widely used strategy for combating virus diseases. Interestingly, most naturally occurring *R genes* conferring resistance to potyviruses have been found to behave in a recessive manner. A 'loss-of-susceptibility' model was originally put forward to mechanistically explain recessive virus *R genes*. This model proposed that resistance arises from homozygous loss-of-function mutations in host factors (susceptibility factors) that are required for the virus lifecycle (Fraser, 1990). Years of research mapping various recessive potyvirus *R genes* has substantiated this loss-of-susceptibility model, as most cases of recessive potyvirus resistance have been shown to result from loss-of-function (or loss-of-VPg-interaction) mutations in components of the eukaryotic translation complex, which are required for the potyvirus lifecycle.

Identification and mapping of *R genes* is the first step towards breeding for virus resistance in crops, however, the next crucial stage is to introgress these genes into the cultivated varieties. Apart from the great time and labour inputs required for the successive crosses and back-crosses associated with breeding, introgression can be hampered by genetic incompatibility between genotypes used for crossing. Furthermore, introgression of a recessive allele can be particularly difficult for crops with polyploid genomes [for instance, potato (4n) or sweet potato (6n)] because of the presence of multiple homologous/homeologous dominant paralogues.

The advent of CRISPR-based genome editing in plants offers great potential overcome some of the current limitations in crop breeding. Chapter 5 provided a proof-of-concept for the application of CRISPR-Cas9 technology for generating novel potyvirus resistance alleles in plants by targeted ablation of the eIF(iso)4E host factor that is required for TuMV proliferation. The high efficiency of this method (59% mutations in the T_2 generation) suggests that this strategy could possibly be used to overcome the problems associated with introgressing recessive *R genes* into polyploid crops, as it may be possible to simultaneously mutate all homologous and homeologous sequences at the *R gene* locus.

The *eIF(iso)4e* mutants generated from this study may also be valuable for assessing potential fitness trade-off associated with mutation at this locus. An advantage of using this *eIF(iso)4E* population is that they are more likely to be isogenic that previously described *eIF(iso)4E* mutant lines which were generated by random mutagenesis(Duprat *et al*., 2002, Lellis *et al*., 2002). Therefore, assessing the fitness the CRISPR-Cas9-generated *eIF(iso)4E* mutant population under a variety of stress conditions, or even under natural growth environments, would be an interesting and important topic for future research.

The proof-of-concept research presented in chapter 5 does not imply that further technological adjustments are not needed to effectively apply this concept to crop breeding scenarios. It is likely that subtle mutations (closely resembling natural mutations) that alter specific amino acids without resulting in complete loss-of-function may be a superior approach for generating novel potyvirus resistance alleles. This is partly because such mutations are likely to be less detrimental to plant growth if the mutation diminishes the viral usurpation of the host factor, without completely abolishing its native function. Additionally, recent research has suggested that these 'subtle' mutations may result in a broader resistance spectrum compared to complete loss-of-function mutations (Piron *et al*., 2010, Ruffel *et al*., 2005). Improved methods for delivering exogenous DNA to precise genomic targets would greatly enable creation of these 'subtle' mutations that could result in stable potyvirus resistance. Chapter 6 was dedicated to exploring new biotechnological methods which could be used to this end.

The *in planta* properties of the recently described CRISPR/Cas nuclease, Cpf1, were investigated with the ultimate goal of using the Cpf1-induced staggered DSBs for targeted insertion of exogenous DNA into plant genomes. As a first step towards this goal, it was established that the LbCpf1 orthologue is more appropriate than the AsCpf1 orthologue for targeted mutation of plant genomes. This finding was later supported by similar finding published by other research groups (Hu *et al*., 2017, Kim *et al*., 2017a, Tang *et al.*, 2017). So far, Cpf1 has not been used to induce heritable mutations in the model plant *Arabidopsis*. Therefore, a mutagenic screen was devised to assess the mutagenic ability of CRISPR/Cpf1 in *Arabidopsis*, relative to the widely used CRISPR/Cas9 system. Problems with the phenotypic screening method and

(possibly) limitations relating to the design of sgRNAs/crRNAs precluded the original goal of accurately comparing the mutagenic efficiencies of SpCas9 and LbCpf1. However, the variable editing frequencies associated with different LbCpf1 crRNAs highlighted the need for improved bioinformatic tools to allow *in silico* assessment of crRNAs for plant applications. Hence, this is a very worthwhile avenue for future research. Another interesting observation from the LbCpf1 editing experiments was the incidence of non-random homozygous mutations in the T_1 generation at a very high frequency. A microhomology directed DNA repair pathway was suggested to be the cause of this observation, but confirmation of this hypothesis requires future research.

Finally, a targeted KI of an exogenous DNA sequence into the *N. benthamiana* (16C) genome was attempted, using microhomology domains generated by Cpf1-induced staggered DSBs. KO, but not KI, alleles were generated using this system. Interestingly however, it seemed that the KO allele may have resulted from a microhomology directed repair pathway, giving support to the notion that utilisation of this pathway might enable deterministic CRISPR/Cpf1 mutations. To further explore the possibility for targeted KI using CRISPR/Cpf1-induced microhomology-directed repair pathways, methods to increase the copy number of the donor DNA should be explored.

While the ultimate goal of targeted KI of exogenous DNA into a plant genome was not achieved, some progress was made to this end. Future research into the potential for Cpf1-induced targeted insertion of DNA sequences may offer powerful new tools for plant biotechnology. The speed of progress for CRISPR-based technologies that has been observed so far suggests that improved methods for targeted KI in plants may be achieved in the near future. This would be greatly beneficial to crop breeding applications and fundamental research, alike.

7.5 Concluding remarks

This thesis has focussed on naturally occurring defences mechanisms that exist to limit potyvirus infections in plants, and explored how our knowledge of these mechanisms can be exploited in an artificial context for eliminating or preventing viruses from crops. Such an approach of concomitantly deepening our understanding of natural

defence, and widening our ability to utilise these mechanisms for agricultural purposes will be an essential step towards reducing global crop losses to viral diseases in the future.

References

Abdelkefi, H., Sugliani, M., Ke, H., Harchouni, S., Soubigou-Taconnat, L., Citerne, S.*, et al.* (2017) Guanosine tetraphosphate modulates salicylic acid signalling and the resistance of Arabidopsis thaliana to Turnip mosaic virus. *Molecular plant pathology*.

Adelman, Z. N., Anderson, M. A., Wiley, M. R., Murreddu, M. G., Samuel, G. H., Morazzani, E. M.*, et al.* (2013) Cooler temperatures destabilize RNA interference and increase susceptibility of disease vector mosquitoes to viral infection. *PLoS neglected tropical diseases,* **7,** e2239.

Agbeci, M., Grangeon, R., Nelson, R. S., Zheng, H. and Laliberté, J.-F. (2013) Contribution of Host Intracellular Transport Machineries to Intercellular Movement of Turnip Mosaic Virus. *PLOS Pathogens,* **9,** e1003683.

Alabadi, D., Oyama, T., Yanovsky, M. J., Harmon, F. G., Mas, P. and Kay, S. A. (2001) Reciprocal regulation between TOC1 and LHY/CCA1 within the Arabidopsis circadian clock. *Science,* **293,** 880-883.

Ali, Z., Abulfaraj, A., Idris, A., Ali, S., Tashkandi, M. and Mahfouz, M. M. (2015) CRISPR/Cas9-mediated viral interference in plants. *Genome biology,* **16,** 238.

Allison, R. F., Sorenson, J. C., Kelly, M. E., Armstrong, F. B. and Dougherty, W. G. (1985) Sequence determination of the capsid protein gene and flanking regions of tobacco etch virus: Evidence for synthesis and processing of a polyprotein in potyvirus genome expression. *Proceedings of the National Academy of Sciences of the United States of America,* **82,** 3969-3972.

Anagnostou, K., Jahn, M. and Perl-Treves, R. (2000) Inheritance and linkage analysis of resistance to zucchini yellow mosaic virus, watermelon mosaic virus, papaya ringspot virus and powdery mildew in melon. *Euphytica,* **116,** 265-270.

Antignus, Y., Raccah, B., Gall-on, A. and Cohen, S. (1989) Biological and Serological Characterization of Zucchini Yellow Mosaic and Watermelon Mosaic Virus-2 Isolates in Israel. *Phytoparasitica,* **17,** 289-298.

Anton, T., Leonhardt, H. and Markaki, Y. (2016) Visualization of Genomic Loci in Living Cells with a Fluorescent CRISPR/Cas9 System. *Methods in molecular biology (Clifton, N.J.),* **1411,** 407-417.

Aravin, A. A., Hannon, G. J. and Brennecke, J. (2007) The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science,* **318,** 761-764.

Aregger, M., Borah, B. K., Seguin, J., Rajeswaran, R., Gubaeva, E. G., Zvereva, A. S.*, et al.* (2012) Primary and secondary siRNAs in geminivirus-induced gene silencing. *PLoS Pathog,* **8,** e1002941.

Asad, S., Haris, W. A., Bashir, A., Zafar, Y., Malik, K. A., Malik, N. N.*, et al.* (2003) Transgenic tobacco expressing geminiviral RNAs are resistant to the serious viral pathogen causing cotton leaf curl disease. *Archives of virology,* **148,** 2341-2352.

Atanasoff, D. (1935) Plum pox. A new virus disease. *Yearbook University of Sofia,* **11,** 49-69.

Atreya, C. D. and Pirone, T. P. (1993) Mutational analysis of the helper component-proteinase gene of a potyvirus: effects of amino acid substitutions, deletions, and

gene replacement on virulence and aphid transmissibility. *Proceedings of the National Academy of Sciences of the United States of America,* **90,** 11919-11923.
Azevedo, J., Garcia, D., Pontier, D., Ohnesorge, S., Yu, A., Garcia, S.*, et al.* (2010) Argonaute quenching and global changes in Dicer homeostasis caused by a pathogen-encoded GW repeat protein. *Genes & development,* **24,** 904-915.
Bally, J., Nakasugi, K., Jia, F., Jung, H., Ho, S. Y., Wong, M.*, et al.* (2015) The extremophile Nicotiana benthamiana has traded viral defence for early vigour. *Nature plants,* **1,** 15165.
Baltes, N. J., Gil-Humanes, J., Cermak, T., Atkins, P. A. and Voytas, D. F. (2014) DNA replicons for plant genome engineering. *The Plant cell,* **26,** 151-163.
Bastet, A., Robaglia, C. and Gallois, J. L. (2017) eIF4E Resistance: Natural Variation Should Guide Gene Editing. *Trends in plant science,* **22,** 411-419.
Baudry, A., Ito, S., Song, Y. H., Strait, A. A., Kiba, T., Lu, S.*, et al.* (2010) F-Box Proteins FKF1 and LKP2 Act in Concert with ZEITLUPE to Control Arabidopsis Clock Progression. *The Plant cell,* **22,** 606-622.
Baulcombe, D. (2004) RNA silencing in plants. *Nature,* **431,** 356-363.
Baulcombe, D. C., Lloyd, J., Manoussopoulos, I. N., Roberts, I. M. and Harrison, B. D. (1993) Signal for potyvirus-dependent aphid transmission of potato aucuba mosaic virus and the effect of its transfer to potato virus X. *Journal of General Virology,* **74,** 1245-1253.
Baumberger, N. and Baulcombe, D. C. (2005) Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *Proc Natl Acad Sci U S A,* **102,** 11928-11933.
Bawden, F. C. and Kleczkowski, A. (1955) Studies on the Ability of Light to Counteract the Inactivating Action of Ultraviolet Radiation on Plant Viruses. *Microbiology (Reading, England),* **13,** 370-382.
Bawden, F. C. and Roberts, F. M. (1947) The influence of light intensity on the susceptibility of plants to certain viruses. *Annals of Applied Biology,* **34,** 286-296.
Bawden, F. C. and Roberts, F. M. (1948) Photosynthesis and Predisposition of Plants to Infection with Certain Viruses. *Annals of Applied Biology,* **35,** 418-428.
Beator, J. and Kloppstech, K. (1996) Significance of circadian gene expression in higher plants. *Chronobiology international,* **13,** 319-339.
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,** 775-782.
Bechtold, U., Karpinski, S. and Mullineaux, P. M. (2005) The influence of the light environment and photosynthesis on oxidative signalling responses in plant–biotrophic pathogen interactions. *Plant, Cell & Environment,* **28,** 1046-1055.
Beczner, L., Horváth, J., Romhányi, I. and Förster, H. (1984) Studies on the aetiology of tuber necrotic ringspot disease in potato. *Potato Research,* **27,** 339-352.
Berger, P. H. and Pirone, T. P. (1986) The effect of helper component on the uptake and localization of potyviruses in Myzus persicae. *Virology,* **153,** 256-261.
Bernstein, E., Caudy, A. A., Hammond, S. M. and Hannon, G. J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature,* **409,** 363-366.
Besong-Ndika, J., Ivanov, K. I., Hafrèn, A., Michon, T. and Mäkinen, K. (2015) Cotranslational Coat Protein-Mediated Inhibition of Potyviral RNA Translation. *Journal of virology,* **89,** 4237-4248.

Bhardwaj, V., Meier, S., Petersen, L. N., Ingle, R. A. and Roden, L. C. (2011) Defence Responses of Arabidopsis thaliana to Infection by Pseudomonas syringae Are Regulated by the Circadian Clock. *PLoS ONE,* **6,** e26968.

Blanc, S., Ammar, E. D., Garcia-Lampasona, S., Dolja, V. V., Llave, C., Baker, J.*, et al.* (1998) Mutations in the potyvirus helper component protein: effects on interactions with virions and aphid stylets. *Journal of General Virology,* **79,** 3119-3122.

Blevins, T., Rajeswaran, R., Aregger, M., Borah, B. K., Schepetilnikov, M., Baerlocher, L.*, et al.* (2011) Massive production of small RNAs from a non-coding region of Cauliflower mosaic virus in plant defense and viral counter-defense. *Nucleic acids research,* **39,** 5003-5014.

Blevins, T., Rajeswaran, R., Shivaprasad, P. V., Beknazariants, D., Si-Ammour, A., Park, H. S.*, et al.* (2006) Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic acids research,* **34,** 6233-6246.

Bonfim, K., Faria, J. C., Nogueira, E. O., Mendes, E. A. and Aragao, F. J. (2007) RNAi-mediated resistance to Bean golden mosaic virus in genetically engineered common bean (Phaseolus vulgaris). *Molecular plant-microbe interactions : MPMI,* **20,** 717-726.

Borsani, O., Zhu, J., Verslues, P. E., Sunkar, R. and Zhu, J. K. (2005) Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell,* **123,** 1279-1291.

Broadbent, L. and Hollings, M. (1951) The influence of heat on some aphids. *Annals of Applied Biology,* **38,** 577-581.

Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y. Y., Sieburth, L.*, et al.* (2008) Widespread translational inhibition by plant miRNAs and siRNAs. *Science,* **320,** 1185-1190.

Brooks, C., Nekrasov, V., Lippman, Z. B. and Van Eck, J. (2014) Efficient gene editing in tomato in the first generation using the clustered regularly interspaced short palindromic repeats/CRISPR-associated9 system. *Plant physiology,* **166,** 1292-1297.

Browning, K. S. (1996) The plant translational apparatus. *Plant molecular biology,* **32,** 107-144.

Bunning, E. (1936) Die endogene Tagesrhythmik als Grundlage der photoperiodischen Reaktion. *Ber. Dtsch. Bot. Ges,* **54,** 590-607.

Burgyan, J. and Havelda, Z. (2011) Viral suppressors of RNA silencing. *Trends in plant science,* **16,** 265-272.

Bush, M. S., Hutchins, A. P., Jones, A. M., Naldrett, M. J., Jarmolowski, A., Lloyd, C. W.*, et al.* (2009) Selective recruitment of proteins to 5' cap complexes during the growth cycle in Arabidopsis. *The Plant journal : for cell and molecular biology,* **59,** 400-412.

Cai, C. Q., Doyon, Y., Ainley, W. M., Miller, J. C., Dekelver, R. C., Moehle, E. A.*, et al.* (2009) Targeted transgene integration in plant cells using designed zinc finger nucleases. *Plant molecular biology,* **69,** 699-709.

Cai, Y., Chen, L., Liu, X., Sun, S., Wu, C., Jiang, B.*, et al.* (2015) CRISPR/Cas9-Mediated Genome Editing in Soybean Hairy Roots. *PLoS One,* **10,** e0136064.

Calarco, J. P., Borges, F., Donoghue, M. T., Van Ex, F., Jullien, P. E., Lopes, T.*, et al.* (2012) Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. *Cell,* **151,** 194-205.

Callot, C. and Gallois, J.-L. (2014) Pyramiding resistances based on translation initiation factors in Arabidopsis is impaired by male gametophyte lethality. *Plant Signaling & Behavior,* **9,** e27940.

Cambra, M., Capote, N., Myrta, A. and Llácer, G. (2006) Plum pox virus and the estimated costs associated with sharka disease. *EPPO Bulletin,* **36,** 202-204.

Carbonell, A., Fahlgren, N., Garcia-Ruiz, H., Gilbert, K. B., Montgomery, T. A., Nguyen, T.*, et al.* (2012) Functional Analysis of Three Arabidopsis ARGONAUTES Using Slicer-Defective Mutants. *The Plant cell,* **24,** 3613.

Carrington, J. C., Freed, D. D. and Sanders, T. C. (1989) Autocatalytic processing of the potyvirus helper component proteinase in Escherichia coli and in vitro. *Journal of virology,* **63,** 4459-4463.

Carrington, J. C., Haldeman, R., Dolja, V. V. and Restrepo-Hartwig, M. A. (1993) Internal cleavage and trans-proteolytic activities of the VPg-proteinase (NIa) of tobacco etch potyvirus in vivo. *Journal of virology,* **67,** 6995-7000.

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 : for cell and molecular biology,* **14,** 393-400.

Cavatorta, J., Perez, K. W., Gray, S. M., Van Eck, J., Yeam, I. and Jahn, M. (2011) Engineering virus resistance using a modified potato gene. *Plant biotechnology journal,* **9,** 1014-1021.

Chandra-Shekara, A. C., Gupte, M., Navarre, D., Raina, S., Raina, R., Klessig, D.*, et al.* (2006) Light-dependent hypersensitive response and resistance signaling against Turnip Crinkle Virus in Arabidopsis. *The Plant journal : for cell and molecular biology,* **45,** 320-334.

Chandrasekaran, J., Brumin, M., Wolf, D., Leibman, D., Klap, C., Pearlsman, M.*, et al.* (2016) Development of broad virus resistance in non-transgenic cucumber using CRISPR/Cas9 technology. *Molecular plant pathology,* **17,** 1140-1153.

Chapman, E. J., Prokhnevsky, A. I., Gopinath, K., Dolja, V. V. and Carrington, J. C. (2004) Viral RNA silencing suppressors inhibit the microRNA pathway at an intermediate step. *Genes & development,* **18,** 1179-1186.

Chavez, A., Scheiman, J., Vora, S., Pruitt, B. W., Tuttle, M., E, P. R. I.*, et al.* (2015) Highly efficient Cas9-mediated transcriptional programming. *Nat Methods,* **12,** 326-328.

Chellappan, P., Vanitharani, R., Ogbe, F. and Fauquet, C. M. (2005) Effect of temperature on geminivirus-induced RNA silencing in plants. *Plant physiology,* **138,** 1828-1841.

Chen, W. Q. and Sherwood, J. L. (1991) Evaluation of Tip Culture, Thermotherapy and Chemotherapy for Elimination of Peanut Mottle Virus from Arachis hypogaea. *Journal of Phytopathology,* **132,** 230-236.

Cheng, C. Y., Gilson, T., Dallaire, F., Ketner, G., Branton, P. E. and Blanchette, P. (2011) The E4orf6/E1B55K E3 ubiquitin ligase complexes of human adenoviruses exhibit heterogeneity in composition and substrate specificity. *J Virol.,* **85**.

Chisholm, S. T., Mahajan, S. K., Whitham, S. A., Yamamoto, M. L. and Carrington, J. C. (2000) Cloning of the Arabidopsis RTM1 gene, which controls restriction of long-distance movement of tobacco etch virus. *Proceedings of the National Academy of Sciences of the United States of America,* **97,** 489-494.

Chisholm, S. T., Parra, M. A., Anderberg, R. J. and Carrington, J. C. (2001) Arabidopsis RTM1 and RTM2 genes function in phloem to restrict long-distance movement of tobacco etch virus. *Plant physiology,* **127,** 1667-1675.

Choi, D., Park, J., Oh, S. and Cheong, H. (2016) Autophagy induction in tobacco leaves infected by potato virus YO and its putative roles. *Biochemical and Biophysical Research Communications,* **474,** 606-611.

Chu, V. T., Weber, T., Wefers, B., Wurst, W., Sander, S. and Rajewsky, K. (2015) Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat Biotechnol.,* **33**.

Chung, B. N., Canto, T., Tenllado, F., Choi, K. S., Joa, J. H., Ahn, J. J.*, et al.* (2016) The Effects of High Temperature on Infection by Potato virus Y, Potato virus A, and Potato leafroll virus. *The plant pathology journal,* **32,** 321-328.

Chung, B. N., Choi, K. S., Ahn, J. J., Joa, J. H., Do, K. S. and Park, K. S. (2015) Effects of Temperature on Systemic Infection and Symptom Expression of Turnip mosaic virus in Chinese cabbage (Brassica campestris). *The plant pathology journal,* **31,** 363-370.

Chung, B. Y., Miller, W. A., Atkins, J. F. and Firth, A. E. (2008) An overlapping essential gene in the Potyviridae. *Proc Natl Acad Sci U S A,* **105,** 5897-5902.

Collmer, C. W., Marston, M. F., Taylor, J. C. and Jahn, M. (2000) The I gene of bean: a dosage-dependent allele conferring extreme resistance, hypersensitive resistance, or spreading vascular necrosis in response to the potyvirus Bean common mosaic virus. *Molecular plant-microbe interactions : MPMI,* **13,** 1266-1270.

Conrath, U. (2006) Systemic Acquired Resistance. *Plant Signaling & Behavior,* **1,** 179-184.

Cosson, P., Schurdi-Levraud, V., Le, Q. H., Sicard, O., Caballero, M., Roux, F.*, et al.* (2012) The RTM Resistance to Potyviruses in Arabidopsis thaliana: Natural Variation of the RTM Genes and Evidence for the Implication of Additional Genes. *PLOS ONE,* **7,** e39169.

Cosson, P., Sofer, L., Hien Le, Q., Léger, V., Schurdi-Levraud, V., Whitham, S. A.*, et al.* (2010) RTM3, Which Controls Long-Distance Movement of Potyviruses, Is a Member of a New Plant Gene Family Encoding a Meprin and TRAF Homology Domain-Containing Protein. *Plant physiology,* **154,** 222.

Cotton, S., Grangeon, R., Thivierge, K., Mathieu, I., Ide, C., Wei, T.*, et al.* (2009) Turnip mosaic virus RNA replication complex vesicles are mobile, align with microfilaments, and are each derived from a single viral genome. *Journal of virology,* **83,** 10460-10471.

Covington, M. F. and Harmer, S. L. (2007) The Circadian Clock Regulates Auxin Signaling and Responses in Arabidopsis. *PLOS Biology,* **5,** e222.

Cronin, S., Verchot, J., Haldeman-Cahill, R., Schaad, M. C. and Carrington, J. C. (1995) Long-distance movement factor: a transport function of the potyvirus helper component proteinase. *The Plant cell,* **7,** 549-559.

Cuevas, J. M., Delaunay, A., Visser, J. C., Bellstedt, D. U., Jacquot, E. and Elena, S. F. (2012) Phylogeography and Molecular Evolution of Potato virus Y. *PLOS ONE,* **7,** e37853.

Cui, H. and Wang, A. (2016) Plum Pox Virus 6K1 Protein Is Required for Viral Replication and Targets the Viral Replication Complex at the Early Stage of Infection. *Journal of virology,* **90,** 5119-5131.

Cui, X., Wei, T., Chowda-Reddy, R. V., Sun, G. and Wang, A. (2010) The Tobacco etch virus P3 protein forms mobile inclusions via the early secretory pathway and traffics along actin microfilaments. *Virology,* **397,** 56-63.

D'Halluin, K., Vanderstraeten, C., Van Hulle, J., Rosolowska, J., Van Den Brande, I., Pennewaert, A.*, et al.* (2013) Targeted molecular trait stacking in cotton through targeted double-strand break induction. *Plant biotechnology journal,* **11,** 933-941.

Dangl, J. L. and Jones, J. D. G. (2001) Plant pathogens and integrated defence responses to infection. *Nature,* **411,** 826-833.

de Mairan, J. (1729) Observation botanique. *Hist. Acad. Roy. Sci.***,** 35-36.

de Ronde, D., Butterbach, P. and Kormelink, R. (2014) Dominant resistance against plant viruses. *Front Plant Sci,* **5,** 307.

Dean, R. A. and Kuć, J. (1986) Induced systemic protection in cucumbers: the source of the "signal". *Physiological and Molecular Plant Pathology,* **28,** 227-233.

Decroocq, V., Salvador, B., Sicard, O., Glasa, M., Cosson, P., Svanella-Dumas, L.*, et al.* (2009) The determinant of potyvirus ability to overcome the RTM resistance of Arabidopsis

thaliana maps to the N-terminal region of the coat protein. *Molecular plant-microbe interactions : MPMI,* **22,** 1302-1311.
Decroocq, V., Sicard, O., Alamillo, J. M., Lansac, M., Eyquard, J. P., Garcia, J. A.*, et al.* (2006) Multiple resistance traits control Plum pox virus infection in Arabidopsis thaliana. *Molecular plant-microbe interactions : MPMI,* **19,** 541-549.
Dehairs, J., Talebi, A., Cherifi, Y. and Swinnen, J. V. (2016) CRISP-ID: decoding CRISPR mediated indels by Sanger sequencing. *Scientific reports,* **6,** 28973.
Del Toro, F. J., Aguilar, E., Hernández-Walias, F. J., Tenllado, F., Chung, B.-N. and Canto, T. (2015) High Temperature, High Ambient CO2 Affect the Interactions between Three Positive-Sense RNA Viruses and a Compatible Host Differentially, but not Their Silencing Suppression Efficiencies. *PLOS ONE,* **10,** e0136062.
Deleris, A., Gallego-Bartolome, J., Bao, J., Kasschau, K. D., Carrington, J. C. and Voinnet, O. (2006) Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. *Science,* **313,** 68-71.
Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A.*, et al.* (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature,* **471,** 602-607.
den Boon, J. A. and Ahlquist, P. (2010) Organelle-like membrane compartmentalization of positive-strand RNA virus replication factories. *Annual review of microbiology,* **64,** 241-256.
Dijkstra, J. and de Jager, C. P. (1998) Determination of the Thermal Inactivation Point. In: *Practical Plant Virology: Protocols and Exercises.* (Dijkstra, J. and de Jager, C. P., eds.). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 102-104.
Doench, J. G., Hartenian, E., Graham, D. B., Tothova, Z., Hegde, M., Smith, I.*, et al.* (2014) Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nature biotechnology,* **32,** 1262-1267.
Doherty, C. J. and Kay, S. A. (2010) Circadian Control of Global Gene Expression Patterns. *Annual review of genetics,* **44,** 419-444.
Dolja, V. V., Haldeman, R., Robertson, N. L., Dougherty, W. G. and Carrington, J. C. (1994) Distinct functions of capsid protein in assembly and movement of tobacco etch potyvirus in plants. *The EMBO Journal,* **13,** 1482-1491.
Dombrovsky, A., Gollop, N., Chen, S., Chejanovsky, N. and Raccah, B. (2007) In vitro association between the helper component-proteinase of zucchini yellow mosaic virus and cuticle proteins of Myzus persicae. *The Journal of general virology,* **88,** 1602-1610.
Dombrovsky, A., Huet, H., Chejanovsky, N. and Raccah, B. (2005) Aphid transmission of a potyvirus depends on suitability of the helper component and the N terminus of the coat protein. *Archives of virology,* **150,** 287-298.
Dong, D., Ren, K., Qiu, X., Zheng, J., Guo, M., Guan, X.*, et al.* (2016) The crystal structure of Cpf1 in complex with CRISPR RNA. *Nature,* **532,** 522-526.
Dougherty, W. G., Lindbo, J. A., Smith, H. A., Parks, T. D., Swaney, S. and Proebsting, W. M. (1994) RNA-mediated virus resistance in transgenic plants: exploitation of a cellular pathway possibly involved in RNA degradation. *Molecular plant-microbe interactions : MPMI,* **7,** 544-552.
Duan, H., Richael, C. and Rommens, C. M. (2012) Overexpression of the wild potato eIF4E-1 variant Eva1 elicits Potato virus Y resistance in plants silenced for native eIF4E-1. *Transgenic research,* **21,** 929-938.
Dunoyer, P., Himber, C. and Voinnet, O. (2005) DICER-LIKE 4 is required for RNA interference and produces the 21-nucleotide small interfering RNA component of the plant cell-to-cell silencing signal. *Nature genetics,* **37,** 1356-1360.

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 : for cell and molecular biology,* **32,** 927-934.

Edgar, R. S., Stangherlin, A., Nagy, A. D., Nicoll, M. P., Efstathiou, S., O'Neill, J. S.*, et al.* (2016) Cell autonomous regulation of herpes and influenza virus infection by the circadian clock. *Proc Natl Acad Sci U S A,* **113,** 10085-10090.

Edwards, K. D., Anderson, P. E., Hall, A., Salathia, N. S., Locke, J. C., Lynn, J. R.*, et al.* (2006) FLOWERING LOCUS C mediates natural variation in the high-temperature response of the Arabidopsis circadian clock. *The Plant cell,* **18,** 639-650.

Elena, S. F. and Rodrigo, G. (2012) Towards an integrated molecular model of plant–virus interactions. *Current Opinion in Virology,* **2,** 719-724.

Endres, M. W., Gregory, B. D., Gao, Z., Foreman, A. W., Mlotshwa, S., Ge, X.*, et al.* (2010) Two plant viral suppressors of silencing require the ethylene-inducible host transcription factor RAV2 to block RNA silencing. *PLoS Pathog,* **6,** e1000729.

Engler, C., Youles, M., Gruetzner, R., Ehnert, T. M., Werner, S., Jones, J. D.*, et al.* (2014) A golden gate modular cloning toolbox for plants. *ACS synthetic biology,* **3,** 839-843.

Esvelt, K. M., Mali, P., Braff, J. L., Moosburner, M., Yaung, S. J. and Church, G. M. (2013) Orthogonal Cas9 Proteins for RNA-Guided Gene Regulation and Editing. *Nature methods,* **10,** 1116-1121.

Fan, M., Bai, M. Y., Kim, J. G., Wang, T., Oh, E., Chen, L.*, et al.* (2014) The bHLH transcription factor HBI1 mediates the trade-off between growth and pathogen-associated molecular pattern-triggered immunity in Arabidopsis. *The Plant cell,* **26,** 828-841.

Farboud, B. and Meyer, B. J. (2015) Dramatic enhancement of genome editing by CRISPR/Cas9 through improved guide RNA design. *Genetics,* **199,** 959-971.

Farre, E. M., Harmer, S. L., Harmon, F. G., Yanovsky, M. J. and Kay, S. A. (2005) Overlapping and distinct roles of PRR7 and PRR9 in the Arabidopsis circadian clock. *Current biology : CB,* **15,** 47-54.

Fauser, F., Schiml, S. and Puchta, H. (2014) Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. *The Plant journal : for cell and molecular biology,* **79,** 348-359.

Feng, Z., Mao, Y., Xu, N., Zhang, B., Wei, P., Yang, D. L.*, et al.* (2014) Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. *Proc Natl Acad Sci U S A,* **111,** 4632-4637.

Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D. L., Wei, P.*, et al.* (2013) Efficient genome editing in plants using a CRISPR/Cas system. *Cell research,* **23,** 1229-1232.

Fernandez-Calvino, L., Goytia, E., Lopez-Abella, D., Giner, A., Urizarna, M., Vilaplana, L.*, et al.* (2010) The helper-component protease transmission factor of tobacco etch potyvirus binds specifically to an aphid ribosomal protein homologous to the laminin receptor precursor. *The Journal of general virology,* **91,** 2862-2873.

Flis, A., Fernandez, A. P., Zielinski, T., Mengin, V., Sulpice, R., Stratford, K.*, et al.* (2015) Defining the robust behaviour of the plant clock gene circuit with absolute RNA timeseries and open infrastructure. *Open biology,* **5**.

Flood, J. (2010) The importance of plant health to food security. *Food Security,* **2,** 215-231.

Flor, H. H. (1971) Current Status of the Gene-For-Gene Concept. *Annual review of phytopathology,* **9,** 275-296.

Fonfara, I., Le Rhun, A., Chylinski, K., Makarova, K. S., Lecrivain, A. L., Bzdrenga, J.*, et al.* (2014) Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic acids research,* **42,** 2577-2590.

Fonfara, I., Richter, H., Bratovic, M., Le Rhun, A. and Charpentier, E. (2016) The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature,* **532,** 517-521.

Foxman, E. F., Storer, J. A., Vanaja, K., Levchenko, A. and Iwasaki, A. (2016) Two interferon-independent double-stranded RNA-induced host defense strategies suppress the common cold virus at warm temperature. *Proceedings of the National Academy of Sciences,* **113,** 8496-8501.

Fraser, R. S. S. (1990) The Genetics of Resistance to Plant Viruses. *Annual review of phytopathology,* **28,** 179-200.

Fulton, R. W. (1984) Pioneer Leaders in Plant Pathology: James Johnson. *Annual review of phytopathology,* **22,** 27-34.

Gaffney, T., Friedrich, L., Vernooij, B., Negrotto, D., Nye, G., Uknes, S.*, et al.* (1993) Requirement of salicylic Acid for the induction of systemic acquired resistance. *Science,* **261,** 754-756.

Gaj, T., Gersbach, C. A. and Barbas, C. F., 3rd (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in biotechnology,* **31,** 397-405.

Gal-On, A. (2000) A Point Mutation in the FRNK Motif of the Potyvirus Helper Component-Protease Gene Alters Symptom Expression in Cucurbits and Elicits Protection Against the Severe Homologous Virus. *Phytopathology,* **90,** 467-473.

Gallois, J. L., Charron, C., Sanchez, F., Pagny, G., Houvenaghel, M. C., Moretti, A.*, et al.* (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. *The Journal of general virology,* **91,** 288-293.

Gao, L., Tuo, D., Shen, W., Yan, P., Li, X. and Zhou, P. (2015) NIa-Pro of Papaya ringspot virus interacts with Carica papaya eukaryotic translation initiation factor 3 subunit G (CpeIF3G). *Virus Genes,* **50,** 97-103.

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. *The Plant journal : for cell and molecular biology,* **40,** 376-385.

Garcia-Ruiz, H., Carbonell, A., Hoyer, J. S., Fahlgren, N., Gilbert, K. B., Takeda, A.*, et al.* (2015) Roles and programming of Arabidopsis ARGONAUTE proteins during Turnip mosaic virus infection. *PLoS Pathog,* **11,** e1004755.

Garcia-Ruiz, H., Takeda, A., Chapman, E. J., Sullivan, C. M., Fahlgren, N., Brempelis, K. J.*, et al.* (2010) Arabidopsis RNA-dependent RNA polymerases and dicer-like proteins in antiviral defense and small interfering RNA biogenesis during Turnip Mosaic Virus infection. *The Plant cell,* **22,** 481-496.

Garcia, D., Garcia, S. and Voinnet, O. (2014) Nonsense-mediated decay serves as a general viral restriction mechanism in plants. *Cell host & microbe,* **16,** 391-402.

Gasiunas, G., Barrangou, R., Horvath, P. and Siksnys, V. (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A,* **109,** E2579-2586.

Gauffier, C., Lebaron, C., Moretti, A., Constant, C., Moquet, F., Bonnet, G.*, et al.* (2016) A TILLING approach to generate broad-spectrum resistance to potyviruses in tomato is hampered by eIF4E gene redundancy. *The Plant journal : for cell and molecular biology,* **85,** 717-729.

Geng, C., Yan, Z. Y., Cheng, D. J., Liu, J., Tian, Y. P., Zhu, C. X.*, et al.* (2017) Tobacco vein banding mosaic virus 6K2 Protein Hijacks NbPsbO1 for Virus Replication. *Scientific reports,* **7,** 43455.

Genoud, T., Buchala, A. J., Chua, N. H. and Metraux, J. P. (2002) Phytochrome signalling modulates the SA-perceptive pathway in Arabidopsis. *The Plant journal : for cell and molecular biology,* **31,** 87-95.

Gerik, J. S., Duffus, J. E., Perry, R., Stenger, D. C. and Maren, A. F. v. (1990) Etiology of tomato plant decline in the California desert. *Phytopathology,* **80,** 1352-1356.

Gibbs, A. and Ohshima, K. (2010) Potyviruses and the digital revolution. *Annual review of phytopathology,* **48,** 205-223.

Gilbert, L. A., Horlbeck, M. A., Adamson, B., Villalta, J. E., Chen, Y., Whitehead, E. H.*, et al.* (2014) Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell,* **159,** 647-661.

Globus, R. and Qimron, U. (2017) A Technological and Regulatory Outlook on CRISPR Crop Editing. *Journal of cellular biochemistry*.

Godfray, H. C., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F.*, et al.* (2010) Food security: the challenge of feeding 9 billion people. *Science,* **327,** 812-818.

Gonsalves, D. (1998) CONTROL OF PAPAYA RINGSPOT VIRUS IN PAPAYA: A Case Study. *Annual review of phytopathology,* **36,** 415-437.

Goodspeed, D., Chehab, E. W., Covington, M. F. and Braam, J. (2013) Circadian control of jasmonates and salicylates: The clock role in plant defense. *Plant Signaling & Behavior,* **8,** e23123.

Goodspeed, D., Chehab, E. W., Min-Venditti, A., Braam, J. and Covington, M. F. (2012) Arabidopsis synchronizes jasmonate-mediated defense with insect circadian behavior. *Proceedings of the National Academy of Sciences of the United States of America,* **109,** 4674-4677.

Govier, D. A. and Kassanis, B. (1974a) Evidence that a component other than the virus particle is needed for aphid transmission of potato virus Y. *Virology,* **57,** 285-286.

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

Graf, A., Schlereth, A., Stitt, M. and Smith, A. M. (2010) Circadian control of carbohydrate availability for growth in Arabidopsis plants at night. *Proceedings of the National Academy of Sciences,* **107,** 9458-9463.

Grangeon, R., Agbeci, M., Chen, J., Grondin, G., Zheng, H. and Laliberte, J. F. (2012a) Impact on the endoplasmic reticulum and Golgi apparatus of turnip mosaic virus infection. *Journal of virology,* **86,** 9255-9265.

Grangeon, R., Jiang, J. and Laliberte, J. F. (2012b) Host endomembrane recruitment for plant RNA virus replication. *Curr Opin Virol,* **2,** 683-690.

Granier, F., Durand-Tardif, M., Casse-Delbart, F., Lecoq, H. and Robaglia, C. (1993) Mutations in zucchini yellow mosaic virus helper component protein associated with loss of aphid transmissibility. *The Journal of general virology,* **74 (Pt 12),** 2737-2742.

Griebel, T. and Zeier, J. (2008) Light Regulation and Daytime Dependency of Inducible Plant Defenses in Arabidopsis: Phytochrome Signaling Controls Systemic Acquired Resistance Rather Than Local Defense. *Plant physiology,* **147,** 790-801.

Gruber, A. R., Lorenz, R., Bernhart, S. H., Neuböck, R. and Hofacker, I. L. (2008) The Vienna RNA Websuite. *Nucleic acids research,* **36,** W70-W74.

Haberl, H., Erb, K. H., Krausmann, F., Gaube, V., Bondeau, A., Plutzar, C.*, et al.* (2007) Quantifying and mapping the human appropriation of net primary production in earth's terrestrial ecosystems. *Proceedings of the National Academy of Sciences,* **104,** 12942-12947.

Hafren, A., Hofius, D., Ronnholm, G., Sonnewald, U. and Makinen, K. (2010) HSP70 and its cochaperone CPIP promote potyvirus infection in Nicotiana benthamiana by regulating viral coat protein functions. *The Plant cell,* **22,** 523-535.

Hajimorad, M. R., Eggenberger, A. L. and Hill, J. H. (2003) Evolution of Soybean mosaic virus-G7 molecularly cloned genome in Rsv1-genotype soybean results in emergence of a mutant capable of evading Rsv1-mediated recognition. *Virology,* **314,** 497-509.

Haldeman-Cahill, R., Daros, J. A. and Carrington, J. C. (1998) Secondary structures in the capsid protein coding sequence and 3' nontranslated region involved in amplification of the tobacco etch virus genome. *Journal of virology,* **72,** 4072-4079.

Hameed, A., Tahir, M. N., Asad, S., Bilal, R., Van Eck, J., Jander, G.*, et al.* (2017) RNAi-Mediated Simultaneous Resistance Against Three RNA Viruses in Potato. *Molecular biotechnology,* **59,** 73-83.

Hari, V. (1981) The RNA of Tobacco etch virus: Further characterization and detection of protein linked to RNA. *Virology,* **112,** 391-399.

Hari, V., Siegel, A., Rozek, C. and Timberlake, W. E. (1979) The RNA of tobacco etch virus contains poly(A). *Virology,* **92,** 568-571.

Harmer, S. L., Hogenesch, J. B., Straume, M., Chang, H. S., Han, B., Zhu, T.*, et al.* (2000) Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. *Science,* **290,** 2110-2113.

Harrison, B. D. (1956) Studies on the effect of temperature on virus multiplication in inoculated leaves. *Annals of Applied Biology,* **44,** 215-226.

Hasan, M. S. and Rashid, M. M. (2015) Viral infections in potato fields in relation to aphid population. *World Applied Sciences Journal,* **33,** 63-68.

Hasiow-Jaroszewska, B., Stachecka, J., Minicka, J., Sowinski, M. and Borodynko, N. (2015) Variability of Potato virus Y in Tomato Crops in Poland and Development of a Reverse-Transcription Loop-Mediated Isothermal Amplification Method for Virus Detection. *Phytopathology,* **105,** 1270-1276.

Havelda, Z., Hornyik, C., Valoczi, A. and Burgyan, J. (2005) Defective interfering RNA hinders the activity of a tombusvirus-encoded posttranscriptional gene silencing suppressor. *Journal of virology,* **79,** 450-457.

Helms, K. and McIntyre, G. A. (1967) Light-induced susceptibility of Phaseolus vulgaris L. to tobacco mosaic virus infection. I. Effects of light intensity, temperature, and the length of the preinoculation dark period. *Virology,* **31,** 191-196.

Hine, R. B., Osborne, W. E. and Dennis, R. E. (1970) Elevation & Temperature Effects on Severity of Maize Dwarf Mosaic Virus in Sorghum. *Plant Disease Reporter,* **54,** 1064-1068.

Hofius, D., Maier, A. T., Dietrich, C., Jungkunz, I., Bornke, F., Maiss, E.*, et al.* (2007) Capsid protein-mediated recruitment of host DnaJ-like proteins is required for Potato virus Y infection in tobacco plants. *Journal of virology,* **81,** 11870-11880.

Hollings, M. and Stone, O. M. (1969) Viruses in fungi. *Science Progress (1933-),* **57,** 371-391.

Holmes, F. O. (1929) Local Lesions in Tobacco Mosaic. *Botanical Gazette,* **87,** 39-55.

Hong, S. M., Bahn, S. C., Lyu, A., Jung, H. S. and Ahn, J. H. (2010) Identification and testing of superior reference genes for a starting pool of transcript normalization in Arabidopsis. *Plant & cell physiology,* **51,** 1694-1706.

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

Horvath, P., Romero, D. A., Coute-Monvoisin, A. C., Richards, M., Deveau, H., Moineau, S.*, et al.* (2008) Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus. *Journal of bacteriology,* **190,** 1401-1412.

Hosford, R. M. (1967) Transmission of Plant Viruses by Dodder. *Botanical Review,* **33,** 387-406.

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V.*, et al.* (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature biotechnology,* **31,** 827-832.

Hu, X., Wang, C., Liu, Q., Fu, Y. and Wang, K. (2017) Targeted mutagenesis in rice using CRISPR-Cpf1 system. *Journal of genetics and genomics = Yi chuan xue bao,* **44,** 71-73.

Huang, H. and Nusinow, D. A. (2016) Into the Evening: Complex Interactions in the Arabidopsis Circadian Clock. *Trends in Genetics,* **32,** 674-686.

Hudson, M. E. and Quail, P. H. (2003) Identification of Promoter Motifs Involved in the Network of Phytochrome A-Regulated Gene Expression by Combined Analysis of Genomic Sequence and Microarray Data. *Plant physiology,* **133,** 1605-1616.

Hwang, J., Li, J., Liu, W.-Y., An, S.-J., Cho, H., Her, N. H.*, et al.* (2009) Double mutations in eIF4E and eIFiso4E confer recessive resistance to Chilli veinal mottle virus in pepper. *Molecules and Cells,* **27,** 329-336.

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D.*, et al.* (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature biotechnology,* **31,** 227-229.

Ilardi, V. and Tavazza, M. (2015) Biotechnological strategies and tools for Plum pox virus resistance: trans-, intra-, cis-genesis, and beyond. *Front Plant Sci,* **6,** 379.

Ito, Y., Nishizawa-Yokoi, A., Endo, M., Mikami, M. and Toki, S. (2015) CRISPR/Cas9-mediated mutagenesis of the RIN locus that regulates tomato fruit ripening. *Biochem Biophys Res Commun,* **467,** 76-82.

Ivanov, K. I., Puustinen, P., Gabrenaite, R., Vihinen, H., Ronnstrand, L., Valmu, L.*, et al.* (2003) Phosphorylation of the potyvirus capsid protein by protein kinase CK2 and its relevance for virus infection. *The Plant cell,* **15,** 2124-2139.

Jacobs, T. B., LaFayette, P. R., Schmitz, R. J. and Parrott, W. A. (2015) Targeted genome modifications in soybean with CRISPR/Cas9. *BMC biotechnology,* **15,** 16.

Jay, F., Wang, Y., Yu, A., Taconnat, L., Pelletier, S., Colot, V.*, et al.* (2011) Misregulation of AUXIN RESPONSE FACTOR 8 underlies the developmental abnormalities caused by three distinct viral silencing suppressors in Arabidopsis. *PLoS Pathog,* **7,** e1002035.

Jeeva, M. L., Balakrishnan, S., Edison, S. and Rajmohan, K. (2004) Meristem culture and thermotherapy in the management of sweet potato feathery mottle virus (SPFMV). *Journal of Root Crops,* **30,** 135-142.

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 : MPMI,* **23,** 1498-1505.

Jenner, C. E., Sanchez, 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 : MPMI,* **13,** 1102-1108.

Jenner, C. E., Tomimura, K., Ohshima, K., Hughes, S. L. and Walsh, J. A. (2002) Mutations in Turnip mosaic virus P3 and cylindrical inclusion proteins are separately required to overcome two Brassica napus resistance genes. *Virology,* **300,** 50-59.

Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B. and Weeks, D. P. (2013) Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic acids research,* **41,** e188.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. and Charpentier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science,* **337,** 816-821.

Johnson, J. (1921) The relation of air temperature to certain plant diseases. *Phytopathology,* **11,** 446-458.

Johnson, J. (1922) The relation of air temperature to the mosaic disease of potatoes and other plants. *Phytopathology,* **12,** 438-440.

Jones, A. L., Thomas, C. L. and Maule, A. J. (1998) De novo methylation and co-suppression induced by a cytoplasmically replicating plant RNA virus. *Embo j,* **17,** 6385-6393.

Jones, L. R. (1924) The Relation of Environment to Disease in Plants. *American Journal of Botany,* **11,** 601-609.

Jones, R. W., Jackson, A. O. and Morris, T. J. (1990) Defective-interfering RNAs and elevated temperatures inhibit replication of tomato bushy stunt virus in inoculated protoplasts. *Virology,* **176,** 539-545.

Josse, E.-M. and Halliday, K. J. (2008) Skotomorphogenesis: The Dark Side of Light Signalling. *Current Biology,* **18,** R1144-R1146.

Kamenova, I. and Milusheva, S. (2005) Sharka Disease in Bulgaria: Past, Present and Future. *Biotechnology & Biotechnological Equipment,* **19,** 22-40.

Kang, B. C., Yeam, I. and Jahn, M. M. (2005) 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.*, et al.* (2005) Evidence that the recessive bymovirus resistance locus rym4 in barley corresponds to the eukaryotic translation initiation factor 4E gene. *Molecular plant pathology,* **6,** 449-458.

Karasev, A. V. and Gray, S. M. (2013) Continuous and emerging challenges of Potato virus Y in potato. *Annual review of phytopathology,* **51,** 571-586.

Karimi, M., Inze, D. and Depicker, A. (2002) GATEWAY vectors for Agrobacterium-mediated plant transformation. *Trends in plant science,* **7,** 193-195.

Kassanis, B. (1952) Some effects of high temperature on the susceptibility of plants to infection with viruses. *Annals of Applied Biology,* **39,** 358-369.

Kassanis, B. (1954) Heat-therapy of virus-infected plants. *Annals of Applied Biology,* **41,** 470-474.

Kassanis, B. and Govier, D. A. (1971) The Role of the Helper Virus in Aphid Transmission of Potato Aucuba Mosaic Virus and Potato Virus C. *Journal of General Virology,* **13,** 221-228.

Kasschau, K. D. and Carrington, J. C. (1998) A Counterdefensive Strategy of Plant Viruses. *Cell,* **95,** 461-470.

Kasschau, K. D., Cronin, S. and Carrington, J. C. (1997) Genome amplification and long-distance movement functions associated with the central domain of tobacco etch potyvirus helper component-proteinase. *Virology,* **228,** 251-262.

Kasschau, K. D., Xie, Z., Allen, E., Llave, C., Chapman, E. J., Krizan, K. A.*, et al.* (2003) P1/HC-Pro, a viral suppressor of RNA silencing, interferes with Arabidopsis development and miRNA unction. *Developmental cell,* **4,** 205-217.

Katiyar-Agarwal, S., Morgan, R., Dahlbeck, D., Borsani, O., Villegas, A., Jr., Zhu, J. K.*, et al.* (2006) A pathogen-inducible endogenous siRNA in plant immunity. *Proc Natl Acad Sci U S A,* **103,** 18002-18007.

Kawoosa, T., Gahlan, P., Devi, A. S. and Kumar, S. (2014) The GATA and SORLIP motifs in the 3-hydroxy-3-methylglutaryl-CoA reductase promoter of Picrorhiza kurrooa for the control of light-mediated expression. *Functional & Integrative Genomics,* **14,** 191-203.

Kegler, H., Fuchs, E., Gruntzig, M. and Schwarz, S. (1998) Some results of 50 years of research on the resistance to plum pox virus. *Acta virologica,* **42,** 200-215.
Kerlan, C., Nikolaeva, O. V., Hu, X., Meacham, T., Gray, S. M. and Karasev, A. V. (2011) Identification of the molecular make-up of the Potato virus Y strain PVY(Z): genetic typing of PVY(Z)-NTN. *Phytopathology,* **101,** 1052-1060.
Kim, B., Masuta, C., Matsuura, H., Takahashi, H. and Inukai, T. (2008) Veinal necrosis induced by turnip mosaic virus infection in Arabidopsis is a form of defense response accompanying HR-like cell death. *Molecular plant-microbe interactions : MPMI,* **21,** 260-268.
Kim, H., Kim, S. T., Ryu, J., Kang, B. C., Kim, J. S. and Kim, S. G. (2017a) CRISPR/Cpf1-mediated DNA-free plant genome editing. *Nature communications,* **8,** 14406.
Kim, H. K., Song, M., Lee, J., Menon, A. V., Jung, S., Kang, Y. M*., et al.* (2017b) In vivo high-throughput profiling of CRISPR-Cpf1 activity. *Nat Methods,* **14,** 153-159.
Kim, J., Kang, W. H., Hwang, J., Yang, H. B., Dosun, K., Oh, C. S*., et al.* (2014) Transgenic Brassica rapa plants over-expressing eIF(iso)4E variants show broad-spectrum Turnip mosaic virus (TuMV) resistance. *Molecular plant pathology,* **15,** 615-626.
Kim, S.-B., Lee, H.-Y., Seo, S., Lee, J. H. and Choi, D. (2015) RNA-Dependent RNA Polymerase (NIb) of the Potyviruses Is an Avirulence Factor for the Broad-Spectrum Resistance Gene Pvr4 in Capsicum annuum cv. CM334. *PLoS ONE,* **10,** e0119639.
Klein, P. G., Klein, R. R., Rodriguez-Cerezo, E., Hunt, A. G. and Shaw, J. G. (1994) Mutational analysis of the tobacco vein mottling virus genome. *Virology,* **204,** 759-769.
Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. and Liu, D. R. (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature,* **533,** 420-424.
Kotakis, C., Vrettos, N., Daskalaki, M. G., Kotzabasis, K. and Kalantidis, K. (2011) DCL3 and DCL4 are likely involved in the light intensity-RNA silencing cross talk in Nicotiana benthamiana. *Plant Signal Behav,* **6,** 1180-1182.
Kotakis, C., Vrettos, N., Kotsis, D., Tsagris, M., Kotzabasis, K. and Kalantidis, K. (2010) Light intensity affects RNA silencing of a transgene in Nicotiana benthamiana plants. *BMC plant biology,* **10,** 220.
Kumar, S. V. and Wigge, P. A. (2010) H2A.Z-containing nucleosomes mediate the thermosensory response in Arabidopsis. *Cell,* **140,** 136-147.
Kunkel, L. O. (1936) Heat treatments for the cure of yellows and other virus diseases of peach. *Phytopathology,* **26,** 809-810.
Kunkel, L. O. (1941) Heat cure of aster yellows in periwinkles. *American Journal of Botany,* **28,** 761-762.
Kurihara, T., Fukuhara, T., Ono, C., Yamamoto, S., Uemura, K., Okamoto, T*., et al.* (2017) Suppression of HBV replication by the expression of nickase- and nuclease dead-Cas9. *Scientific reports,* **7,** 6122.
Lain, S., Riechmann, J. L. and Garcia, J. A. (1990) RNA helicase: a novel activity associated with a protein encoded by a positive strand RNA virus. *Nucleic acids research,* **18,** 7003-7006.
Lakatos, L., Csorba, T., Pantaleo, V., Chapman, E. J., Carrington, J. C., Liu, Y. P*., et al.* (2006) Small RNA binding is a common strategy to suppress RNA silencing by several viral suppressors. *Embo j,* **25,** 2768-2780.
Laliberte, J. F. and Sanfacon, H. (2010) Cellular remodeling during plant virus infection. *Annual review of phytopathology,* **48,** 69-91.
Le Romancer, M. and Nedellec, M. (1997) Effect of plant genotype, virus isolate and temperature on the expression of the potato tuber necrotic ringspot disease (PTNRD). *Plant Pathology,* **46,** 104-111.

Lebaron, C., Rosado, A., Sauvage, C., Gauffier, C., German-Retana, S., Moury, B.*, et al.* (2016) A new eIF4E1 allele characterized by RNAseq data mining is associated with resistance to potato virus Y in tomato albeit with a low durability. *The Journal of general virology,* **97,** 3063-3072.

Lee, S.-C., Wu, M. and Wong, S.-M. (1993) Nucleotide sequence of a Singapore isolate of zucchini yellow mosaic virus coat protein gene revealed an altered DAG motif. *Virus Genes,* **7,** 381-387.

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 : CB,* **12,** 1046-1051.

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

Li, J.-F., Aach, J., Norville, J. E., McCormack, M., Zhang, D., Bush, J.*, et al.* (2013a) Multiplex and homologous recombination-mediated plant genome editing via guide RNA/Cas9. *Nature biotechnology,* **31,** 688-691.

Li, J., Dai, X. and Zhao, Y. (2006) A role for auxin response factor 19 in auxin and ethylene signaling in Arabidopsis. *Plant physiology,* **140,** 899-908.

Li, W., Teng, F., Li, T. and Zhou, Q. (2013b) Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nature biotechnology,* **31,** 684-686.

Li, X. H., Valdez, P., Olvera, R. E. and Carrington, J. C. (1997) Functions of the tobacco etch virus RNA polymerase (NIb): subcellular transport and protein-protein interaction with VPg/proteinase (NIa). *Journal of virology,* **71,** 1598-1607.

Liang, G., Zhang, H., Lou, D. and Yu, D. (2016) Selection of highly efficient sgRNAs for CRISPR/Cas9-based plant genome editing. *Scientific reports,* **6,** 21451.

Liang, Z., Zhang, K., Chen, K. and Gao, C. (2014) Targeted mutagenesis in Zea mays using TALENs and the CRISPR/Cas system. *Journal of genetics and genomics = Yi chuan xue bao,* **41,** 63-68.

Lindbo, J. A., Silva-Rosales, L., Proebsting, W. M. and Dougherty, W. G. (1993) Induction of a Highly Specific Antiviral State in Transgenic Plants: Implications for Regulation of Gene Expression and Virus Resistance. *The Plant cell,* **5,** 1749.

Liu, Y., Gao, Q., Wu, B., Ai, T. and Guo, X. (2009) NgRDR1, an RNA-dependent RNA polymerase isolated from Nicotiana glutinosa, was involved in biotic and abiotic stresses. *Plant physiology and biochemistry : PPB,* **47,** 359-368.

Loebenstein, G. (2009) Local lesions and induced resistance. *Advances in virus research,* **75,** 73-117.

Loebenstein, G. (2015) Control of sweet potato virus diseases. *Advances in virus research,* **91,** 33-45.

López-Delgado, H., Mora-Herrera, M. E., Zavaleta-Mancera, H. A., Cadena-Hinojosa, M. and Scott, I. M. (2004) Salicylic acid enhances heat tolerance and potato virus X (PVX) elimination during thermotherapy of potato microplants. *American Journal of Potato Research,* **81,** 171-176.

Louie, R. and Knoke, J. K. (1975) Strains of maize dwarf mosaic virus. *Plant Disease Reporter,* **59,** 518-522.

Ma, J., Hou, X., Xiao, D., Qi, L., Wang, F., Sun, F.*, et al.* (2010) Cloning and Characterization of the BcTuR3 Gene Related to Resistance to Turnip Mosaic Virus (TuMV) from Non-heading Chinese Cabbage. *Plant Molecular Biology Reporter,* **28,** 588-596.

Ma, L., Huang, X., Yu, R., Jing, X. L., Xu, J., Wu, C. A.*, et al.* (2016) Elevated Ambient Temperature Differentially Affects Virus Resistance in Two Tobacco Species. *Phytopathology,* **106,** 94-100.

MacLean, D., Elina, N., Havecker, E. R., Heimstaedt, S. B., Studholme, D. J. and Baulcombe, D. C. (2010) Evidence for large complex networks of plant short silencing RNAs. *PLoS One,* **5,** e9901.

Mahajan, S., 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,** 4370-4379.

Mahajan, S. K., Chisholm, S. T., Whitham, S. A. and Carrington, J. C. (1998) Identification and characterization of a locus (RTM1) that restricts long-distance movement of tobacco etch virus in Arabidopsis thaliana. *The Plant journal : for cell and molecular biology,* **14,** 177-186.

Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J.*, et al.* (2015) An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Micro,* **13,** 722-736.

Mäkinen, K. and Hafren, A. (2014) Intracellular coordination of potyviral RNA functions in infection. *Frontiers in Plant Science,* **5**.

Maldonado, A. M., Doerner, P., Dixon, R. A., Lamb, C. J. and Cameron, R. K. (2002) A putative lipid transfer protein involved in systemic resistance signalling in Arabidopsis. *Nature,* **419,** 399-403.

Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S.*, et al.* (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature biotechnology,* **31,** 833-838.

Manacorda, C. A., Mansilla, C., Debat, H. J., Zavallo, D., Sanchez, F., Ponz, F.*, et al.* (2013) Salicylic acid determines differential senescence produced by two Turnip mosaic virus strains involving reactive oxygen species and early transcriptomic changes. *Molecular plant-microbe interactions : MPMI,* **26,** 1486-1498.

Manfre, A., Glenn, M., Nunez, A., Moreau, R. A. and Dardick, C. (2011) Light quantity and photosystem function mediate host susceptibility to Turnip mosaic virus via a salicylic acid-independent mechanism. *Molecular plant-microbe interactions : MPMI,* **24,** 315-327.

Manganaris, G. A., Economou, A. S., Boubourakas, I. N. and Katis, N. I. (2003) Elimination of PPV and PNRSV through thermotherapy and meristem-tip culture in nectarine. *Plant cell reports,* **22,** 195-200.

Mangrauthia, S. K., Singh Shakya, V. P., Jain, R. K. and Praveen, S. (2009) Ambient temperature perception in papaya for papaya ringspot virus interaction. *Virus Genes,* **38,** 429-434.

Martínez-Silva, A. V., Aguirre-Martínez, C., Flores-Tinoco, C. E., Alejandri-Ramírez, N. D. and Dinkova, T. D. (2012) Translation Initiation Factor AteIF(iso)4E Is Involved in Selective mRNA Translation in Arabidopsis Thaliana Seedlings. *PLOS ONE,* **7,** e31606.

Martinez, F., Rodrigo, G., Aragones, V., Ruiz, M., Lodewijk, I., Fernandez, U.*, et al.* (2016) Interaction network of tobacco etch potyvirus NIa protein with the host proteome during infection. *BMC genomics,* **17,** 87.

Martinez, F., Sardanyes, J., Elena, S. F. and Daros, J. A. (2011) Dynamics of a plant RNA virus intracellular accumulation: stamping machine vs. geometric replication. *Genetics,* **188,** 637-646.

Matthews, R. E. F. (1953a) Factors affecting the production of local lesions by plant viruses. I. The effect of time of day of inoculation *Annals of Applied Biology,* **40,** 377-383.

Matthews, R. E. F. (1953b) Factors affecting the production of local lesions by plant viruses. II. Some effects of light, darkness and temperature. *Annals of Applied Biology,* **40,** 377-383.
Matzke, M. A. and Mosher, R. A. (2014) RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nature reviews. Genetics,* **15,** 394-408.
Maule, A. J. (2008) Plasmodesmata: structure, function and biogenesis. *Current opinion in plant biology,* **11,** 680-686.
Maule, A. J., Caranta, C. and Boulton, M. I. (2007) Sources of natural resistance to plant viruses: status and prospects. *Molecular plant pathology,* **8,** 223-231.
McCallum, C. M., Comai, L., Greene, E. A. and Henikoff, S. (2000) Targeted screening for induced mutations. *Nature biotechnology,* **18,** 455-457.
McClung, C. R. (2006) Plant Circadian Rhythms. *The Plant cell,* **18,** 792.
McDonald, M. J. and Rosbash, M. (2001) Microarray analysis and organization of circadian gene expression in Drosophila. *Cell,* **107,** 567-578.
McVey, M. and Lee, S. E. (2008) MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends in genetics : TIG,* **24,** 529-538.
Meinke, D. W. (2013) A survey of dominant mutations in Arabidopsis thaliana. *Trends in plant science,* **18,** 84-91.
Merrow, M., Spoelstra, K. and Roenneberg, T. (2005) The circadian cycle: daily rhythms from behaviour to genes: First in the Cycles Review Series. *EMBO Reports,* **6,** 930-935.
Mette, M. F., Aufsatz, W., van der Winden, J., Matzke, M. A. and Matzke, A. J. (2000) Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *Embo j,* **19,** 5194-5201.
Michael, T. P. and McClung, C. R. (2003) Enhancer trapping reveals widespread circadian clock transcriptional control in Arabidopsis. *Plant physiology,* **132,** 629-639.
Michael, T. P., Mockler, T. C., Breton, G., McEntee, C., Byer, A., Trout, J. D.*, et al.* (2008) Network discovery pipeline elucidates conserved time-of-day-specific cis-regulatory modules. *PLoS genetics,* **4,** e14.
Mikami, M., Toki, S. and Endo, M. (2016) Precision Targeted Mutagenesis via Cas9 Paired Nickases in Rice. *Plant and Cell Physiology,* **57,** 1058-1068.
Miller, S. and Krijnse-Locker, J. (2008) Modification of intracellular membrane structures for virus replication. *Nature reviews. Microbiology,* **6,** 363-374.
Mink, G. I. W., R. Howell, W.E. (1998) Heat treatment of perennial plants to eliminate phytoplasmas, viruses and viroids while maintaining plant survival. Breeding for resistance to plant viruses. *Plant Virus Disease Control***,** 332-345.
Mockler, T. C., Michael, T. P., Priest, H. D., Shen, R., Sullivan, C. M., Givan, S. A.*, et al.* (2007) The DIURNAL project: DIURNAL and circadian expression profiling, model-based pattern matching, and promoter analysis. *Cold Spring Harbor symposia on quantitative biology,* **72,** 353-363.
Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J. and Almendros, C. (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology (Reading, England),* **155,** 733-740.
Moldovan, D., Spriggs, A., Dennis, E. S. and Wilson, I. W. (2010) The hunt for hypoxia responsive natural antisense short interfering RNAs. *Plant Signal Behav,* **5,** 247-251.
Molnar, A., Csorba, T., Lakatos, L., Varallyay, E., Lacomme, C. and Burgyan, J. (2005) Plant virus-derived small interfering RNAs originate predominantly from highly structured single-stranded viral RNAs. *Journal of virology,* **79,** 7812-7818.

Murphy, J. F., Rychlik, W., Rhoads, R. E., Hunt, A. G. and Shaw, J. G. (1991) A tyrosine residue in the small nuclear inclusion protein of tobacco vein mottling virus links the VPg to the viral RNA. *Journal of virology,* **65,** 511-513.

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. *Molecular plant pathology,* **11,** 255-263.

Nakamichi, N., Kiba, T., Henriques, R., Mizuno, T., Chua, N. H. and Sakakibara, H. (2010) PSEUDO-RESPONSE REGULATORS 9, 7, and 5 are transcriptional repressors in the Arabidopsis circadian clock. *The Plant cell,* **22,** 594-605.

Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J. D. and Kamoun, S. (2013) Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. *Nature biotechnology,* **31,** 691-693.

Newman, J. A. (2004) Climate change and cereal aphids: the relative effects of increasing CO2 and temperature on aphid population dynamics. *Global Change Biology,* **10,** 5-15.

Nicaise, V., Gallois, J. L., Chafiai, F., Allen, L. M., Schurdi-Levraud, V., Browning, K. S.*, et al.* (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., Sanjuan, R., Dubrana, M. P., Mazier, M., Maisonneuve, B.*, et al.* (2003) The eukaryotic translation initiation factor 4E controls lettuce susceptibility to the Potyvirus Lettuce mosaic virus. *Plant physiology,* **132,** 1272-1282.

Nieto, C., Morales, M., Orjeda, G., Clepet, C., Monfort, A., Sturbois, B.*, et al.* (2006) An eIF4E allele confers resistance to an uncapped and non-polyadenylated RNA virus in melon. *The Plant journal : for cell and molecular biology,* **48,** 452-462.

Nozue, K., Covington, M. F., Duek, P. D., Lorrain, S., Fankhauser, C., Harmer, S. L.*, et al.* (2007) Rhythmic growth explained by coincidence between internal and external cues. *Nature,* **448,** 358-361.

Nusinow, D. A., Helfer, A., Hamilton, E. E., King, J. J., Imaizumi, T., Schultz, T. F.*, et al.* (2011) The ELF4-ELF3-LUX complex links the circadian clock to diurnal control of hypocotyl growth. *Nature,* **475,** 398-402.

Okushima, Y., Overvoorde, P. J., Arima, K., Alonso, J. M., Chan, A., Chang, C.*, et al.* (2005) Functional genomic analysis of the AUXIN RESPONSE FACTOR gene family members in Arabidopsis thaliana: unique and overlapping functions of ARF7 and ARF19. *The Plant cell,* **17,** 444-463.

Oldroyd, G. E. and Staskawicz, B. J. (1998) Genetically engineered broad-spectrum disease resistance in tomato. *Proc Natl Acad Sci U S A,* **95,** 10300-10305.

Ouibrahim, L., Rubio, A. G., Moretti, A., Montane, M. H., Menand, B., Meyer, C.*, et al.* (2015) Potyviruses differ in their requirement for TOR signalling. *The Journal of general virology,* **96,** 2898-2903.

Pan, C., Ye, L., Qin, L., Liu, X., He, Y., Wang, J.*, et al.* (2016) CRISPR/Cas9-mediated efficient and heritable targeted mutagenesis in tomato plants in the first and later generations. *Scientific reports,* **6,** 24765.

Panattoni, A., Luvisi, A. and Triolo, E. (2013) Review. Elimination of viruses in plants: twenty years of progress. *Spanish Journal of Agricultural Research,* **11,** 173-188.

Patrick, R. M., Mayberry, L. K., Choy, G., Woodard, L. E., Liu, J. S., White, A.*, et al.* (2014) Two Arabidopsis loci encode novel eukaryotic initiation factor 4E isoforms that are functionally distinct from the conserved plant eukaryotic initiation factor 4E. *Plant physiology,* **164,** 1820-1830.

Peng, Y. H., Kadoury, D., Gal-On, A., Huet, H., Wang, Y. and Raccah, B. (1998) Mutations in the HC-Pro gene of zucchini yellow mosaic potyvirus: effects on aphid transmission and binding to purified virions. *The Journal of general virology,* **79 (Pt 4),** 897-904.

Pfeiffer, P., Goedecke, W. and Obe, G. (2000) Mechanisms of DNA double-strand break repair and their potential to induce chromosomal aberrations. *Mutagenesis,* **15,** 289-302.

Piron, F., Nicolai, M., Minoia, S., Piednoir, E., Moretti, A., Salgues, A.*, et al.* (2010) An induced mutation in tomato eIF4E leads to immunity to two potyviruses. *PLoS One,* **5,** e11313.

Pirone, T. P. and Megahed, E.-S. (1966) Aphid transmissibility of some purified viruses and viral RNA's. *Virology,* **30,** 631-637.

Pittendrigh, C. S. (1993) Temporal organization: reflections of a Darwinian clock-watcher. *Annual review of physiology,* **55,** 16-54.

Pittendrigh, C. S. and Minis, D. H. (1964) The Entrainment of Circadian Oscillations by Light and Their Role as Photoperiodic Clocks. *The American Naturalist,* **98,** 261-294.

Pokhilko, A., Fernandez, A. P., Edwards, K. D., Southern, M. M., Halliday, K. J. and Millar, A. J. (2012) The clock gene circuit in Arabidopsis includes a repressilator with additional feedback loops. *Molecular systems biology,* **8,** 574.

Pooggin, M., Shivaprasad, P. V., Veluthambi, K. and Hohn, T. (2003) RNAi targeting of DNA virus in plants. *Nature biotechnology,* **21,** 131-132.

Pooggin, M. M. (2013) How can plant DNA viruses evade siRNA-directed DNA methylation and silencing? *International journal of molecular sciences,* **14,** 15233-15259.

Poque, S., Pagny, G., Ouibrahim, L., Chague, A., Eyquard, J. P., Caballero, M.*, et al.* (2015) Allelic variation at the rpv1 locus controls partial resistance to Plum pox virus infection in Arabidopsis thaliana. *BMC plant biology,* **15,** 159.

Poulicard, N., Pacios, L. F., Gallois, J. L., Pinero, D. and Garcia-Arenal, F. (2016) Human Management of a Wild Plant Modulates the Evolutionary Dynamics of a Gene Determining Recessive Resistance to Virus Infection. *PLoS genetics,* **12,** e1006214.

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

Puchta, H. (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *Journal of experimental botany,* **56,** 1-14.

Pumplin, N., Sarazin, A., Jullien, P. E., Bologna, N. G., Oberlin, S. and Voinnet, O. (2016) DNA Methylation Influences the Expression of DICER-LIKE4 Isoforms, Which Encode Proteins of Alternative Localization and Function. *The Plant cell,* **28,** 2786.

Puustinen, P. and Makinen, K. (2004) Uridylylation of the potyvirus VPg by viral replicase NIb correlates with the nucleotide binding capacity of VPg. *The Journal of biological chemistry,* **279,** 38103-38110.

Pyott, D. E. and Molnar, A. (2015) Going mobile: non-cell-autonomous small RNAs shape the genetic landscape of plants. *Plant biotechnology journal,* **13,** 306-318.

Pyott, D. E., Sheehan, E. and Molnar, A. (2016) Engineering of CRISPR/Cas9-mediated potyvirus resistance in transgene-free Arabidopsis plants. *Molecular plant pathology,* **17,** 1276-1288.

Qi, X., Bao, F. S. and Xie, Z. (2009) Small RNA deep sequencing reveals role for Arabidopsis thaliana RNA-dependent RNA polymerases in viral siRNA biogenesis. *PLoS One,* **4,** e4971.

Qu, F., Ye, X., Hou, G., Sato, S., Clemente, T. E. and Morris, T. J. (2005) RDR6 has a broad-spectrum but temperature-dependent antiviral defense role in Nicotiana benthamiana. *Journal of virology,* **79,** 15209-15217.

Qu, F., Ye, X. and Morris, T. J. (2008) Arabidopsis DRB4, AGO1, AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. *Proc Natl Acad Sci U S A,* **105,** 14732-14737.

Ramgareeb, S., Snyman, S. J., van Antwerpen, T. and Rutherford, R. S. (2010) Elimination of virus and rapid propagation of disease-free sugarcane (Saccharum spp. cultivar NCo376) using apical meristem culture. *Plant Cell, Tissue and Organ Culture (PCTOC),* **100,** 175-181.

Ran, F. A., Hsu, P. D., Lin, C.-Y., Gootenberg, J. S., Konermann, S., Trevino, A.*, et al.* (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell,* **154,** 1380-1389.

Ratcliff, F. G., MacFarlane, S. A. and Baulcombe, D. C. (1999) Gene silencing without DNA. rna-mediated cross-protection between viruses. *The Plant cell,* **11,** 1207-1216.

Regenmortel, M. and Mahy, B. (2009) Desk Encyclopedia of Plant and Fungal Virology. Elsevier, pp. 426-430.

Reinhart, B. J., Weinstein, E. G., Rhoades, M. W., Bartel, B. and Bartel, D. P. (2002) MicroRNAs in plants. *Genes & development,* **16,** 1616-1626.

Riechmann, J. L., Cervera, M. T. and Garcia, J. A. (1995) Processing of the plum pox virus polyprotein at the P3-6K1 junction is not required for virus viability. *The Journal of general virology,* **76 (Pt 4),** 951-956.

Riechmann, J. L., Lain, S. and Garcia, J. A. (1992) Highlights and prospects of potyvirus molecular biology. *The Journal of general virology,* **73 (Pt 1),** 1-16.

Rimbaud, L., Dallot, S., Gottwald, T., Decroocq, V., Jacquot, E., Soubeyrand, S.*, et al.* (2015) Sharka epidemiology and worldwide management strategies: learning lessons to optimize disease control in perennial plants. *Annual review of phytopathology,* **53,** 357-378.

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

Robaglia, C., Durand-Tardif, M., Tronchet, M., Boudazin, G., Astier-Manifacier, S. and Casse-Delbart, F. (1989) Nucleotide sequence of potato virus Y (N Strain) genomic RNA. *The Journal of general virology,* **70 (Pt 4),** 935-947.

Robert-Seilaniantz, A., Grant, M. and Jones, J. D. (2011) Hormone crosstalk in plant disease and defense: more than just jasmonate-salicylate antagonism. *Annual review of phytopathology,* **49,** 317-343.

Robert, F., Barbeau, M., Éthier, S., Dostie, J. and Pelletier, J. (2015) Pharmacological inhibition of DNA-PK stimulates Cas9-mediated genome editing. *Genome Medicine,* **7,** 93.

Roberts, A. G. and Oparka, K. J. (2003) Plasmodesmata and the control of symplastic transport. *Plant, Cell & Environment,* **26,** 103-124.

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.

Rodgers, K. and McVey, M. (2016) Error-Prone Repair of DNA Double-Strand Breaks. *Journal of Cellular Physiology,* **231,** 15-24.

Rodriguez-Cerezo, E., Findlay, K., Shaw, J. G., Lomonossoff, G. P., Qiu, S. G., Linstead, P.*, et al.* (1997) The coat and cylindrical inclusion proteins of a potyvirus are associated with connections between plant cells. *Virology,* **236,** 296-306.

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,** 465-473.

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 proteins. *Virology,* **237,** 283-295.

Rojo, E., Solano, R. and Sánchez-Serrano, J. J. (2003) Interactions Between Signaling Compounds Involved in Plant Defense. *Journal of Plant Growth Regulation,* **22,** 82-98.

Romancer, M. L., Kerlan, C. and Nedellec†, M. (1994) Biological characterisation of various geographical isolates of potato virus Y inducing superficial necrosis on potato tubers. *Plant Pathology,* **43,** 138-144.

Rosbash, M. (2009) The Implications of Multiple Circadian Clock Origins. *PLOS Biology,* **7,** e1000062.

Ross, A. F. (1961) Localized acquired resistance to plant virus infection in hypersensitive hosts. *Virology,* **14,** 329-339.

Ruffel, S., Dussault, M. H., Palloix, A., Moury, B., Bendahmane, A., Robaglia, C*., et al.* (2002) A natural recessive resistance gene against potato virus Y in pepper corresponds to the eukaryotic initiation factor 4E (eIF4E). *The Plant journal : for cell and molecular biology,* **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 : MGG,* **274,** 346-353.

Ruiz, M. T., Voinnet, O. and Baulcombe, D. C. (1998) Initiation and maintenance of virus-induced gene silencing. *The Plant cell,* **10,** 937-946.

Sahana, N., Kaur, H., Jain, R. K., Palukaitis, P., Canto, T. and Praveen, S. (2014) The asparagine residue in the FRNK box of potyviral helper-component protease is critical for its small RNA binding and subcellular localization. *The Journal of general virology,* **95,** 1167-1177.

Samach, A. and Gover, A. (2001) Photoperiodism: The consistent use of CONSTANS. *Current Biology,* **11,** R651-R654.

Samach, A., Onouchi, H., Gold, S. E., Ditta, G. S., Schwarz-Sommer, Z., Yanofsky, M. F*., et al.* (2000) Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. *Science,* **288,** 1613-1616.

Sambrook, J. and Russell, D. W. (2006) The inoue method for preparation and transformation of competent e. Coli: "ultra-competent" cells. *CSH protocols,* **2006**.

Sanchez, F., Martinez-Herrera, D., Aguilar, I. and Ponz, F. (1998) Infectivity of turnip mosaic potyvirus cDNA clones and transcripts on the systemic host Arabidopsis thaliana and local lesion hosts. *Virus research,* **55,** 207-219.

Sastry, K. S. and Zitter, T. A. (2014) Management of Virus and Viroid Diseases of Crops in the Tropics. In: *Plant Virus and Viroid Diseases in the Tropics: Volume 2: Epidemiology and Management.* (Sastry, K. S. and A. Zitter, T., eds.). Dordrecht: Springer Netherlands, pp. 149-480.

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 letters,* **579,** 1167-1171.

Sauer, N. J., Narvaez-Vasquez, J., Mozoruk, J., Miller, R. B., Warburg, Z. J., Woodward, M. J*., et al.* (2016) Oligonucleotide-Mediated Genome Editing Provides Precision and Function to Engineered Nucleases and Antibiotics in Plants. *Plant physiology,* **170,** 1917-1928.

Saunders, D. S. (2009) Circadian rhythms and the evolution of photoperiodic timing in insects. *Physiological Entomology,* **34,** 301-308.

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.

Schaad, M. C., Jensen, P. E. and Carrington, J. C. (1997) Formation of plant RNA virus replication complexes on membranes: role of an endoplasmic reticulum-targeted viral protein. *The EMBO Journal,* **16,** 4049-4059.

Schiml, S., Fauser, F. and Puchta, H. (2014) The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. *The Plant journal : for cell and molecular biology,* **80,** 1139-1150.

Schmittgen, T. D. and Livak, K. J. (2008) Analyzing real-time PCR data by the comparative CT method. *Nat. Protocols,* **3,** 1101-1108.

Schoelz, J. E., Harries, P. A. and Nelson, R. S. (2011) Intracellular transport of plant viruses: finding the door out of the cell. *Molecular plant,* **4,** 813-831.

Scholthof, K. B., Adkins, S., Czosnek, H., Palukaitis, P., Jacquot, E., Hohn, T*., et al.* (2011) Top 10 plant viruses in molecular plant pathology. *Molecular plant pathology,* **12,** 938-954.

Schott, G., Mari-Ordonez, A., Himber, C., Alioua, A., Voinnet, O. and Dunoyer, P. (2015) Differential effects of viral silencing suppressors on siRNA and miRNA loading support the existence of two distinct cellular pools of ARGONAUTE1. *Embo j,* **34,** 2593-2594.

Schwach, F., Vaistij, F. E., Jones, L. and Baulcombe, D. C. (2005) An RNA-Dependent RNA Polymerase Prevents Meristem Invasion by Potato Virus X and Is Required for the Activity But Not the Production of a Systemic Silencing Signal. *Plant physiology,* **138,** 1842.

Scorza, R., Callahan, A., Levy, L., Damsteegt, V., Webb, K. and Ravelonandro, M. (2001) Post-transcriptional gene silencing in plum pox virus resistant transgenic European plum containing the plum pox potyvirus coat protein gene. *Transgenic research,* **10,** 201-209.

Scorza, R., Callahan, A., Ravelonandro, M. and Braverman, M. (2012) Development and Regulation of the Plum Pox Virus Resistant Transgenic Plum 'HoneySweet'. In: *Regulation of Agricultural Biotechnology: The United States and Canada.* (Wozniak, C. A. and McHughen, A., eds.). Dordrecht: Springer Netherlands, pp. 269-280.

Scorza, R., Ravelonandro, M., Callahan, A. M., Cordts, J. M., Fuchs, M., Dunez, J*., et al.* (1994) Transgenic plums (Prunus domestica L.) express the plum pox virus coat protein gene. *Plant cell reports,* **14,** 18-22.

Seaton, D. D., Smith, R. W., Song, Y. H., MacGregor, D. R., Stewart, K., Steel, G*., et al.* (2015) Linked circadian outputs control elongation growth and flowering in response to photoperiod and temperature. *Molecular systems biology,* **11,** 776.

Seo, P. J. and Mas, P. (2015) STRESSing the role of the plant circadian clock. *Trends in plant science,* **20,** 230-237.

Shan, Q., Wang, Y., Li, J. and Gao, C. (2014) Genome editing in rice and wheat using the CRISPR/Cas system. *Nature protocols,* **9,** 2395-2410.

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z*., et al.* (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. *Nature biotechnology,* **31,** 686-688.

Sharma, O. P., Sharma, P. P. and Chowfla, S. C. (1989) Inheritance of resistance to potato virus Y in garden pepper (Capsicum annuum L.). *Euphytica,* **42,** 31-33.

Sharma, P. N., Sharma, V., Sharma, A., Rajput, K. and Sharma, S. K. (2015) Identification and molecular characterization of Bean yellow mosaic virus infecting French bean in Himachal Pradesh. *Virusdisease,* **26,** 315-318.

Shattuck, V. I. (2010) The Biology, Epidemiology, and Control of Turnip Mosaic Virus. In: *Horticultural Reviews.* John Wiley & Sons, Inc., pp. 199-238.

Shen, B., Zhang, W., Zhang, J., Zhou, J., Wang, J., Chen, L.*, et al.* (2014) Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat Meth,* **11,** 399-402.

Shi, J., Gao, H., Wang, H., Lafitte, H. R., Archibald, R. L., Yang, M.*, et al.* (2017) ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. *Plant biotechnology journal,* **15,** 207-216.

Shiboleth, Y. M., Haronsky, E., Leibman, D., Arazi, T., Wassenegger, M., Whitham, S. A.*, et al.* (2007) The conserved FRNK box in HC-Pro, a plant viral suppressor of gene silencing, is required for small RNA binding and mediates symptom development. *Journal of virology,* **81,** 13135-13148.

Shimada, T. L., Shimada, T. and Hara-Nishimura, I. (2010) A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of Arabidopsis thaliana. *The Plant journal : for cell and molecular biology,* **61,** 519-528.

Shimura, H., Pantaleo, V., Ishihara, T., Myojo, N., Inaba, J.-i., Sueda, K.*, et al.* (2011) A Viral Satellite RNA Induces Yellow Symptoms on Tobacco by Targeting a Gene Involved in Chlorophyll Biosynthesis using the RNA Silencing Machinery. *PLOS Pathogens,* **7,** e1002021.

Shivaprasad, P. V., Chen, H. M., Patel, K., Bond, D. M., Santos, B. A. and Baulcombe, D. C. (2012) A microRNA superfamily regulates nucleotide binding site-leucine-rich repeats and other mRNAs. *The Plant cell,* **24,** 859-874.

Shopan, J., Mou, H., Zhang, L., Zhang, C., Ma, W., Walsh, J. A.*, et al.* (2017) Eukaryotic translation initiation factor 2B-beta (eIF2Bbeta), a new class of plant virus resistance gene. *The Plant journal : for cell and molecular biology,* **90,** 929-940.

Shukla, V. K., Doyon, Y., Miller, J. C., DeKelver, R. C., Moehle, E. A., Worden, S. E.*, et al.* (2009) Precise genome modification in the crop species Zea mays using zinc-finger nucleases. *Nature,* **459,** 437-441.

Singh, M. N., Paul Khurana, S. M., Nagaich, B. B. and Agrawal, H. O. (1988) Environmental factors influencing aphid transmission of potato virus Y and potato leafroll virus. *Potato Research,* **31,** 501-509.

Singh, R. P., Valkonen, J. P. T., Gray, S. M., Boonham, N., Jones, R. A. C., Kerlan, C.*, et al.* (2008) Discussion paper: The naming of Potato virus Y strains infecting potato. *Archives of virology,* **153,** 1-13.

Smith, K. M. (1931) On the Composite Nature of Certain Potato Virus Diseases of the Mosaic Group as Revealed by the use of Plant Indicators and Selective Methods of Transmission. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character,* **109,** 251.

Smith, K. M. and Bald, J. C. (1935) A description of a necrotic virus disease affecting tobacco and other plants. *Parasitology,* **27,** 231-232.

Smith, N. A., Eamens, A. L. and Wang, M. B. (2011) Viral small interfering RNAs target host genes to mediate disease symptoms in plants. *PLoS Pathog,* **7,** e1002022.

Sochor, J., Babula, P., Adam, V., Krska, B. and Kizek, R. (2012) Sharka: the past, the present and the future. *Viruses,* **4,** 2853-2901.

Sofer, L., Cabanillas, D. G., Gayral, M., Teplier, R., Pouzoulet, J., Ducousso, M.*, et al.* (2017) Identification of host factors potentially involved in RTM-mediated resistance during potyvirus long distance movement. *Archives of virology,* **162,** 1855-1865.

Soumounou, Y. and Laliberte, J. F. (1994) Nucleic acid-binding properties of the P1 protein of turnip mosaic potyvirus produced in Escherichia coli. *The Journal of general virology,* **75 (Pt 10),** 2567-2573.

Spence, N. J., Phiri, N. A., Hughes, S. L., Mwaniki, A., Simons, S., Oduor, G.*, et al.* (2007) Economic impact of Turnip mosaic virus, Cauliflower mosaic virus and Beet mosaic virus in three Kenyan vegetables. *Plant Pathology,* **56,** 317-323.

Stein, N., Perovic, D., Kumlehn, J., Pellio, B., Stracke, S., Streng, S.*, et al.* (2005) The eukaryotic translation initiation factor 4E confers multiallelic recessive Bymovirus resistance in Hordeum vulgare (L.). *The Plant Journal,* **42,** 912-922.

Steinert, J., Schiml, S., Fauser, F. and Puchta, H. (2015) Highly efficient heritable plant genome engineering using Cas9 orthologues from Streptococcus thermophilus and Staphylococcus aureus. *The Plant journal : for cell and molecular biology,* **84,** 1295-1305.

Stemmer, M., Thumberger, T., Del Sol Keyer, M., Wittbrodt, J. and Mateo, J. L. (2015) CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLoS One,* **10,** e0124633.

Stenger, D. C., Hein, G. L., Gildow, F. E., Horken, K. M. and French, R. (2005) Plant virus HC-Pro is a determinant of eriophyid mite transmission. *Journal of virology,* **79,** 9054-9061.

Sternberg, S. H., LaFrance, B., Kaplan, M. and Doudna, J. A. (2015) Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature,* **527,** 110-113.

Sun, X., Hu, Z., Chen, R., Jiang, Q., Song, G., Zhang, H.*, et al.* (2015) Targeted mutagenesis in soybean using the CRISPR-Cas9 system. *Scientific reports,* **5,** 10342.

Sun, Y., Zhang, X., Wu, C., He, Y., Ma, Y., Hou, H.*, et al.* (2016) Engineering Herbicide-Resistant Rice Plants through CRISPR/Cas9-Mediated Homologous Recombination of Acetolactate Synthase. *Molecular plant,* **9,** 628-631.

Svitashev, S., Schwartz, C., Lenderts, B., Young, J. K. and Mark Cigan, A. (2016) Genome editing in maize directed by CRISPR-Cas9 ribonucleoprotein complexes. *Nature communications,* **7,** 13274.

Svitashev, S., Young, J. K., Schwartz, C., Gao, H., Falco, S. C. and Cigan, A. M. (2015) Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA. *Plant physiology,* **169,** 931-945.

Szittya, G., Silhavy, D., Molnar, A., Havelda, Z., Lovas, A., Lakatos, L.*, et al.* (2003) Low temperature inhibits RNA silencing-mediated defence by the control of siRNA generation. *Embo j,* **22,** 633-640.

Tan, R., Wang, L., Hong, N. and Wang, G. (2010) Enhanced efficiency of virus eradication following thermotherapy of shoot-tip cultures of pear. *Plant Cell, Tissue and Organ Culture (PCTOC),* **101,** 229-235.

Tang, G., Reinhart, B. J., Bartel, D. P. and Zamore, P. D. (2003) A biochemical framework for RNA silencing in plants. *Genes & development,* **17,** 49-63.

Tang, X., Lowder, L. G., Zhang, T., Malzahn, A. A., Zheng, X., Voytas, D. F.*, et al.* (2017) A CRISPR-Cpf1 system for efficient genome editing and transcriptional repression in plants. *Nature plants,* **3,** 17103.

Taylor, C. E. and Robertson, W. M. (1974) Electron Microscopy Evidence for the Association of Tobacco Severe Etch Virus with the Maxillae in Myzus persicae (Sulz.). *Journal of Phytopathology,* **80,** 257-266.

Thackray, D. J., Diggle, A. J., Berlandier, F. A. and Jones, R. A. (2004) Forecasting aphid outbreaks and epidemics of Cucumber mosaic virus in lupin crops in a Mediterranean-type environment. *Virus research,* **100,** 67-82.

Thain, S. C., Murtas, G., Lynn, J. R., McGrath, R. B. and Millar, A. J. (2002) The Circadian Clock That Controls Gene Expression in Arabidopsis Is Tissue Specific. *Plant physiology,* **130,** 102-110.

Thivierge, K., Cotton, S., Dufresne, P. J., Mathieu, I., Beauchemin, C., Ide, C.*, et al.* (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.

Thomma, B. P., Penninckx, I. A., Broekaert, W. F. and Cammue, B. P. (2001) The complexity of disease signaling in Arabidopsis. *Current opinion in immunology,* **13,** 63-68.

Thomson, A. D. and Wright, D. S. C. (1966) Incidence and some effects of potato virus Y on new zealand flue-cured tobacco. *New Zealand journal of agricultural research,* **9,** 886--893.

Thornbury, D. W., Patterson, C. A., Dessens, J. T. and Pirone, T. P. (1990) Comparative sequence of the helper component (HC) region of potato virus Y and a HC-defective strain, potato virus C. *Virology,* **178,** 573-578.

Tian, Y. P. and Valkonen, J. P. (2013) Genetic determinants of Potato virus Y required to overcome or trigger hypersensitive resistance to PVY strain group O controlled by the gene Ny in potato. *Molecular plant-microbe interactions : MPMI,* **26,** 297-305.

Tillet (1755) "Dissertation sur la cause qui corrompt et noircit les grans de ble dans les epis; et sur les moyens de prevenir ces accidents. Bordeaux" English Translation by H.B. Humphrey. *Phytopathology Classics,* **5,** 191-192.

Tilman, D., Balzer, C., Hill, J. and Befort, B. L. (2011) Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences,* **108,** 20260-20264.

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

Tompkins, C. M. (1939) Two mosaic diseases of annual stock. *Journal of Agricultural Research,* **58,** 63-77.

Torres, A. C., Fajardo, T. V., Dusi, A. N., Resende, R. d. O. and Buso, J. A. (2000) Shoot tip culture and thermotherapy for recovering virus-free plants of garlic. *Horticultura Brasileira,* **18,** 192-195.

Ueda, H. R., Matsumoto, A., Kawamura, M., Iino, M., Tanimura, T. and Hashimoto, S. (2002) Genome-wide transcriptional orchestration of circadian rhythms in Drosophila. *The Journal of biological chemistry,* **277,** 14048-14052.

Usenik, V., Kastelec, D., Stampar, F. and Marn, M. V. (2015) Effect of Plum pox virus on chemical composition and fruit quality of plum. *Journal of agricultural and food chemistry,* **63,** 51-60.

Usenik, V. and Marn, M. V. (2017) Sugars and organic acids in plum fruit affected by Plum pox virus. *Journal of the Science of Food and Agriculture,* **97,** 2154-2158.

van den Heuvel, J. F. J. M., van der Vlugt, R. A. A., Verbeek, M., de Haan, P. T. and Huttinga, H. (1994) Characteristics of a resistance-breaking isolate of potato virus Y causing potato tuber necrotic ringspot disease. *European Journal of Plant Pathology,* **100,** 347-356.

van der Ploeg, J. R. (2009) Analysis of CRISPR in Streptococcus mutans suggests frequent occurrence of acquired immunity against infection by M102-like bacteriophages. *Microbiology (Reading, England),* **155,** 1966-1976.

Várallyay, É., Válóczi, A., Ágyi, Á., Burgyán, J. and Havelda, Z. (2010) Plant virus-mediated induction of miR168 is associated with repression of ARGONAUTE1 accumulation. *The EMBO Journal,* **29,** 3507-3519.

Vaucheret, H., Vazquez, F., Crete, P. and Bartel, D. P. (2004) The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes & development,* **18,** 1187-1197.

Vazquez, F., Gasciolli, V., Crete, P. and Vaucheret, H. (2004) The nuclear dsRNA binding protein HYL1 is required for microRNA accumulation and plant development, but not posttranscriptional transgene silencing. *Current biology : CB,* **14,** 346-351.

Velázquez, K., Renovell, A., Comellas, M., Serra, P., García, M. L., Pina, J. A.*, et al.* (2010) Effect of temperature on RNA silencing of a negative-stranded RNA plant virus: Citrus psorosis virus. *Plant Pathology,* **59,** 982-990.

Verchot, J. (2011) Wrapping membranes around plant virus infection. *Curr Opin Virol,* **1,** 388-395.

Verchot, J., Herndon, K. L. and Carrington, J. C. (1992) Mutational analysis of the tobacco etch potyviral 35-kDa proteinase: identification of essential residues and requirements for autoproteolysis. *Virology,* **190,** 298-306.

Vidal, S., Cabrera, H., Andersson, R. A., Fredriksson, A. and Valkonen, J. P. (2002) Potato gene Y-1 is an N gene homolog that confers cell death upon infection with potato virus Y. *Molecular plant-microbe interactions : MPMI,* **15,** 717-727.

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,** e1002639.

Vitousek, P. M., Ehrlich, P. R., Ehrlich, A. H. and Matson, P. A. (1986) Human Appropriation of the Products of Photosynthesis. *BioScience,* **36,** 368-373.

Voinnet, O. (2008) Use, tolerance and avoidance of amplified RNA silencing by plants. *Trends in plant science,* **13,** 317-328.

Walsh, J. A. and Jenner, C. E. (2002) Turnip mosaic virus and the quest for durable resistance. *Molecular plant pathology,* **3,** 289-300.

Wan, J., Cabanillas, D. G., Zheng, H. and Laliberte, J. F. (2015) Turnip mosaic virus moves systemically through both phloem and xylem as membrane-associated complexes. *Plant physiology,* **167,** 1374-1388.

Wang, D., Macfarlane, S. A. and Maule, A. J. (1997) Viral Determinants of Pea Early Browning Virus Seed Transmission in Pea. *Virology,* **234,** 112-117.

Wang, D. and Maule, A. J. (1994) A Model for Seed Transmission of a Plant Virus: Genetic and Structural Analyses of Pea Embryo Invasion by Pea Seed-Borne Mosaic Virus. *The Plant cell,* **6,** 777.

Wang, G.-Y., Shi, J.-L., Ng, G., Battle, S. L., Zhang, C. and Lu, H. (2011) Circadian Clock-Regulated Phosphate Transporter PHT4;1 Plays an Important Role in Arabidopsis Defense. *Molecular plant,* **4,** 516-526.

Wang, G., Zhang, C., Battle, S. and Lu, H. (2014) The phosphate transporter PHT4;1 is a salicylic acid regulator likely controlled by the circadian clock protein CCA1. *Frontiers in Plant Science,* **5,** 701.

Wang, H., Yang, H., Shivalila, C. S., Dawlaty, M. M., Cheng, A. W., Zhang, F.*, et al.* (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell,* **153,** 910-918.

Wang, M., Lu, Y., Botella, J. R., Mao, Y., Hua, K. and Zhu, J. K. (2017a) Gene Targeting by Homology-Directed Repair in Rice Using a Geminivirus-Based CRISPR/Cas9 System. *Molecular plant,* **10,** 1007-1010.

Wang, M., Mao, Y., Lu, Y., Tao, X. and Zhu, J.-k. (2017b) Multiplex Gene Editing in Rice Using the CRISPR-Cpf1 System. *Molecular plant,* **10,** 1011-1013.

Wang, M. B., Masuta, C., Smith, N. A. and Shimura, H. (2012) RNA silencing and plant viral diseases. *Molecular plant-microbe interactions : MPMI,* **25,** 1275-1285.

Wang, X., Ullah, Z. and Grumet, R. (2000) Interaction between zucchini yellow mosaic potyvirus RNA-dependent RNA polymerase and host poly-(A) binding protein. *Virology,* **275,** 433-443.

Wang, Z. P., Xing, H. L., Dong, L., Zhang, H. Y., Han, C. Y., Wang, X. C*., et al.* (2015) Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. *Genome biology,* **16,** 144.

Wang, Z. Y. and Tobin, E. M. (1998) Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. *Cell,* **93,** 1207-1217.

Wardlaw, I. F. and Bagnall, D. (1981) Phloem Transport and the Regulation of Growth of Sorghum bicolor (Moench) at Low Temperature. *Plant physiology,* **68,** 411-414.

Wassenegger, M., Heimes, S., Riedel, L. and Sanger, H. L. (1994) RNA-directed de novo methylation of genomic sequences in plants. *Cell,* **76,** 567-576.

Waterhouse, P. M., Wang, M. B. and Lough, T. (2001) Gene silencing as an adaptive defence against viruses. *Nature,* **411,** 834-842.

Wei, T., Huang, T. S., McNeil, J., Laliberte, J. F., Hong, J., Nelson, R. S*., et al.* (2010a) Sequential recruitment of the endoplasmic reticulum and chloroplasts for plant potyvirus replication. *Journal of virology,* **84,** 799-809.

Wei, T. and Wang, A. (2008) Biogenesis of cytoplasmic membranous vesicles for plant potyvirus replication occurs at endoplasmic reticulum exit sites in a COPI- and COPII-dependent manner. *Journal of virology,* **82,** 12252-12264.

Wei, T., Zhang, C., Hong, J., Xiong, R., Kasschau, K. D., Zhou, X*., et al.* (2010b) Formation of Complexes at Plasmodesmata for Potyvirus Intercellular Movement Is Mediated by the Viral Protein P3N-PIPO. *PLOS Pathogens,* **6,** e1000962.

Wei, T., Zhang, C., Hou, X., Sanfaçon, H. and Wang, A. (2013) The SNARE Protein Syp71 Is Essential for Turnip Mosaic Virus Infection by Mediating Fusion of Virus-Induced Vesicles with Chloroplasts. *PLOS Pathogens,* **9,** e1003378.

Wen, R. H. and Hajimorad, M. R. (2010) Mutational analysis of the putative pipo of soybean mosaic virus suggests disruption of PIPO protein impedes movement. *Virology,* **400,** 1-7.

White, J. L. and Kaper, J. M. (1989) A simple method for detection of viral satellite RNAs in small plant tissue samples. *Journal of virological methods,* **23,** 83-93.

Whitham, S. A., Anderberg, R. J., Chisholm, S. T. and Carrington, J. C. (2000) Arabidopsis RTM2 gene is necessary for specific restriction of tobacco etch virus and encodes an unusual small heat shock-like protein. *The Plant cell,* **12,** 569-582.

Whitham, S. A., Yamamoto, M. L. and Carrington, J. C. (1999) Selectable viruses and altered susceptibility mutants in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America,* **96,** 772-777.

Wiltshire, G. H. (1956) The effect of darkening on the susceptibility of plants to infection with viruses. *Annals of Applied Biology,* **44,** 233-248.

Wittmann, S., Chatel, H., Fortin, M. G. and Laliberte, 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.

Woo, J. W., Kim, J., Kwon, S. I., Corvalan, C., Cho, S. W., Kim, H*., et al.* (2015) DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nature biotechnology,* **33,** 1162-1164.

Xie, K. and Yang, Y. (2013) RNA-guided genome editing in plants using a CRISPR-Cas system. *Molecular plant,* **6,** 1975-1983.

Xie, Z., Fan, B., Chen, C. and Chen, Z. (2001) An important role of an inducible RNA-dependent RNA polymerase in plant antiviral defense. *Proceedings of the National Academy of Sciences,* **98,** 6516-6521.

Xie, Z., Kasschau, K. D. and Carrington, J. C. (2003) Negative feedback regulation of Dicer-Like1 in Arabidopsis by microRNA-guided mRNA degradation. *Current biology : CB,* **13,** 784-789.

Xu, R., Qin, R., Li, H., Li, D., Li, L., Wei, P.*, et al.* (2017) Generation of targeted mutant rice using a CRISPR-Cpf1 system. *Plant biotechnology journal,* **15,** 713-717.

Yamamura, Y. and Scholthof, H. B. (2005) Tomato bushy stunt virus: a resilient model system to study virus-plant interactions. *Molecular plant pathology,* **6,** 491-502.

Yamano, T., Nishimasu, H., Zetsche, B., Hirano, H., Slaymaker, I. M., Li, Y.*, et al.* (2016) Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA. *Cell,* **165,** 949-962.

Yang, S. J., Carter, S. A., Cole, A. B., Cheng, N. H. and Nelson, R. S. (2004a) A natural variant of a host RNA-dependent RNA polymerase is associated with increased susceptibility to viruses by Nicotiana benthamiana. *Proc Natl Acad Sci U S A,* **101,** 6297-6302.

Yang, Y., Sherwood, T. A., Patte, C. P., Hiebert, E. and Polston, J. E. (2004b) Use of Tomato yellow leaf curl virus (TYLCV) Rep Gene Sequences to Engineer TYLCV Resistance in Tomato. *Phytopathology,* **94,** 490-496.

Yarwood, C. E. (1952a) Latent Period and Generation Time for Two Plant Viruses. *American Journal of Botany,* **39,** 613-618.

Yarwood, C. E. (1952b) Some Relations of Carbohydrate Level of the Host to Plant Virus Infections. *American Journal of Botany,* **39,** 119-124.

Yoo, B.-C., Kragler, F., Varkonyi-Gasic, E., Haywood, V., Archer-Evans, S., Lee, Y. M.*, et al.* (2004) A Systemic Small RNA Signaling System in Plants. *The Plant cell,* **16,** 1979.

Yu, D., Fan, B., MacFarlane, S. A. and Chen, Z. (2003) Analysis of the involvement of an inducible Arabidopsis RNA-dependent RNA polymerase in antiviral defense. *Molecular plant-microbe interactions : MPMI,* **16,** 206-216.

Zeier, J., Pink, B., Mueller, M. J. and Berger, S. (2004) Light conditions influence specific defence responses in incompatible plant-pathogen interactions: uncoupling systemic resistance from salicylic acid and PR-1 accumulation. *Planta,* **219,** 673-683.

Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P.*, et al.* (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell,* **163,** 759-771.

Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E. M.*, et al.* (2017) Multiplex gene editing by CRISPR-Cpf1 through autonomous processing of a single crRNA array. *Nature biotechnology,* **35,** 31-34.

Zhang, C., Xie, Q., Anderson, R. G., Ng, G., Seitz, N. C., Peterson, T.*, et al.* (2013a) Crosstalk between the Circadian Clock and Innate Immunity in Arabidopsis. *PLOS Pathogens,* **9,** e1003370.

Zhang, H., Zhang, J., Wei, P., Zhang, B., Gou, F., Feng, Z.*, et al.* (2014) The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant biotechnology journal,* **12,** 797-807.

Zhang, X., Zhang, X., Singh, J., Li, D. and Qu, F. (2012) Temperature-dependent survival of Turnip crinkle virus-infected arabidopsis plants relies on an RNA silencing-based defense that requires dcl2, AGO2, and HEN1. *Journal of virology,* **86,** 6847-6854.

Zhang, Y., Heidrich, N., Ampattu, B. J., Gunderson, C. W., Seifert, H. S., Schoen, C.*, et al.* (2013b) Processing-independent CRISPR RNAs limit natural transformation in Neisseria meningitidis. *Molecular cell,* **50,** 488-503.

Zhao, Y., Zhang, C., Liu, W., Gao, W., Liu, C., Song, G.*, et al.* (2016) An alternative strategy for targeted gene replacement in plants using a dual-sgRNA/Cas9 design. *Scientific reports,* **6,** 23890.

Zhou, H., Liu, B., Weeks, D. P., Spalding, M. H. and Yang, B. (2014) Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic acids research,* **42,** 10903-10914.

Zhu, J., Song, N., Sun, S., Yang, W., Zhao, H., Song, W.*, et al.* (2016) Efficiency and Inheritance of Targeted Mutagenesis in Maize Using CRISPR-Cas9. *Journal of genetics and genomics = Yi chuan xue bao,* **43,** 25-36.

Zhu, L., Mon, H., Xu, J., Lee, J. M. and Kusakabe, T. (2015) CRISPR/Cas9-mediated knockout of factors in non-homologous end joining pathway enhances gene targeting in silkworm cells. *Scientific reports,* **5,** 18103.

Zinn, J. G. (1759) Von dem Schlafe der Pflanzen. *Hamburg. Mag. ,* **22,** 40-50.