A molecular genetic analysis of resistance to poleroviruses in sugar beet and oilseed rape

Eileen Wallace Gallagher

A thesis submitted to the University of East Anglia for the degree of Doctor of Philosophy

University of East Anglia
School of Biological Sciences
Norwich

September 2013

Abstract

Beet mild yellowing virus (BMYV) and *Turnip yellows virus* (TuYV) are both poleroviruses that cause significant reduction in the yields of sugar beet and oilseed rape respectively. Both viruses are transmitted by the aphid vector *Myzus persicae*. Current control methods rely heavily on the use of insecticides for controlling the aphids which can spread these viruses to a wide range of host plants. Recent EU guidelines have tightened control on the use of some of these pesticides, meaning it is becoming increasingly important to find alternative control methods. It is widely agreed in the scientific community, that the best control method would be to generate durable genetically resistant crop plants. In order to achieve this gene targets, either for active or passive resistance, would need to be identified.

This study has built on a project that identified a naturally BMYV resistant *A. thaliana* ecotype, Sna-1. Crosses of the susceptible ecotype (Col-0) to the resistant ecotype Sna-1 identified the resistance as 'passive', where susceptibility was dominant, and conditioned by a monogenic trait. This study began by characterising the gene responsible for susceptibility by bulked segregant analysis and AFLP™. This identified a region of ca. 5Mbp region on *A. thaliana* chromosome 4. This region contains the Arabidopsis elongation initiation factor 4E (*eIF4E*) gene which has already been implicated in susceptibility to other viruses. This gene has frequently been shown to be important for viral infection in plants, and naturally occurring mutations can result in resistance to other viruses. Further investigation revealed a 12 bp duplicated sequence in the Sna-1 *eIF4E* allele, located in a region that encodes the cap-binding pocket of eIF4E. The same region has been shown to be required for virus infection in other species. Infections were therefore carried out using mutants in this gene, using TAS-ELISA. Previously susceptible Col-0 plants containing a T-DNA insert, or EMS point mutations in the *eIF4E* gene were found to be resistant to BMYV infection. Functional complementation with the Col-0 *eIF4E* allele into a stock that contained Sna-1 *eIF4E* resulted in susceptibility to BMYV, confirming its role as a susceptibility factor.

As BMYV and TuYV are closely related viruses it was hypothesised they would share a similar infection strategy. The mutation in *eIF4E* was not enough to prevent virus infection, and the method of infection of the UK-BB TuYV isolate remains to be elucidated as infection studies in mutants with defective components of the eukaryotic translation initiation factors, including *eIF(Iso)4E* gene, has so far failed to identify any requirements for UK-BB TuYV infection. Several T-DNA insertion lines in the *eIF(iso)4E* gene were tested but it was not possible to verify that any of these lines were true knock-outs. However, the molelcular tools for future verification have been developed. A recent report has implicated eIF(iso)4G components in TuYV infection of Arabidopsis but this result could not be repeated in this study.

Further study is required to fully understand the mode of infection of both viruses. It is expected that the identification of essential host genes required for virus infection will aid in the breeding of genetically resistant crops, and reduce the current dependence on harmful pesticides.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank several people for their help throughout this project.

I am grateful for the help, support and guidance from Dr Colwyn Thomas, Dr Mark Stevens, Dr Charles Brearley and Dr Gabriella Kelemen. I would like to particularly thank Colwyn for his perseverance and encouragement throughout this project.

I would also like to thank colleagues at the NRP, and Brooms Barn. I would like to thank Dr Lawrence Percival-Alwyn for all the work he put into this original project, and helping me take it over. I would like to thank Phillipa Hallsworth, Kevin Sawford, and Felicita Vigano for their help with experiments at Brooms Barn Applied Crop Sciences and I would also like to thank Matthew Smoker and Jodie Pike at the Sainsbury Laboratory and Pam Wells at UEA, for their help and guidance. I would like to thank Kendall Baker at UEA for being a great friend and colleague in the lab. Mostly I would like to thank Hayley Moore for her wonderful friendship throughout my time of study, and for being a fantastic ‘team-mate’, and so patient and understanding and helping me through.

I am grateful to all my friends for their support throughout my studies, particularly Hannah Chalke who has been a wonderful friend throughout, and will continue to be for long after. My biggest thanks must be to my family whose support has been irreplaceable. To my mum, dad and brother for all the encouragement and distractions they have given me, and to my granny in Scotland for her continued encouragement. I would like to dedicate this thesis to my Gran and Grandpa who are sadly not able to see me complete this milestone.

Chapter 1 Introduction

1.1 Introduction

Sugar beet (*Beta vulgaris)* and oilseed rape (*Brassica napus*) are crops grown widely in the UK and around the world. Sugar beet is one of the most valuable crops grown in the UK, and contributes 55% of the UK sugar use and almost one third of the worlds sugar production. It is estimated that 7.5 million tonnes of sugar beet is produced annually in the UK (www.britishsugar.co.uk). As well as its commercial uses, sugar beet is a valuable break crop used in crop rotations. Oilseed rape is also a "break crop" and is increasingly being grown around the UK. An estimated 2 million tonnes at a value of £700 million was produced in the UK in 2010 (www.ukagriculture.com).

As with many crops commonly grown around the globe, it is not just abiotic environmental conditions that farmers are battling with, but also biotic factors such as insects, fungi and viruses. Sugar beet and oilseed rape are no exception to this, as both are subject to a variety of insect pests and diseases (Cooke and Scott, 1993; Kimber and McGregor, 1995). Luteovirade are a family of plant viruses, which include the Poleroviruses *Beet mild yellowing virus* (BMYV) and *Turnip yellows virus* (TuYV). Infection with either of these viruses, carried by the green peach aphid (*Myzus persicae*) can result in significant crop losses to farmers (Stevens et al., 2005). Current pesticide control methods are not sufficiently effective for dealing with these diseases, therefore research into TuYV and BMYV, their infection strategy, and the investigation of innovative forms of resistance to viruses are of great interest to growers (Stevens et al., 2008).

1.1.2 Sugar Beet

1.1.2.1 History

B. vulgaris can be subdivided into four main categories: leaf beets, garden beets, fodder beets, and sugar beet. Leaf beets comprise two categories spinach beet and Swiss chard. Both these varieties are cultivated for salads. Garden beets are also grown for human consumption and include varieties such as beetroot and fodder beets, as the name suggests, are grown for livestock feeds. Sugar beet is the most widely grown variety of beet, and commercially grown for production of sugar as an alternative to sugar cane (Winner, 1993).

Beet has been cultivated as a food source for centuries. There is evidence from Greek and Roman civilisations that it was originally cultivated to use the leaves in a similar way that we commonly use Swiss chard (Ford-Lloyd and Williams, 1975). The use of beets for food spread throughout Europe in the Middle Ages as the root started to be prepared for food alongside the leaves, as it was noticed that when the root was cooked a sweet syrup was produced (Winner, 1993). It wasn't until 1747 when Andreas Sigismund Marggaraf, a German chemist, realised that sugar was able to be extracted from the root. Marggaraf was able to produce sugar crystals similar to that extracted from sugar cane. This process was not able to produce the same quantities of sugar as sugar cane, so was largely thought unsustainable at the time (Marggraf, 1749).

One of Marggrafs students, Franz Carl Achard, is commonly considered 'the father of the beet industry', as he recognised the potential for beet cultivation. He analysed the sugar content of various beets given to him by local farmers who were growing them for livestock feed. He identified a variety of beet that had a white skin and a conical shape, to be the variety that contained the most sugar. This variety became known as 'White Silesian beet', and is the progenitor of most sugar beet grown worldwide (Winner, 1993). In 1799 Achard published his findings and was soon awarded funding to establish the first sugar beet factory where beet was cultivated and processed. Achard is

also considered to be the first sugar beet breeder, as he was able to recognise that abiotic conditions, such as soil type, alongside the cultivation methods, had an effect on the sugars produced within the beet. Throughout the 19th and early 20th century sugar beet farming spread around the world as mechanisation of harvesting and processing of the sugar improved. By 1990 the worldwide annual sugar beet production was estimated to be around 40 million tonnes (Winner, 1993).

1.1.2.1 Growth and harvesting

Sugar beet now commonly cultivated in Europe has been improved by plant breeding. Roots of modern beet varieties contain a fresh weight sugar concentration of around 18 %, and a dry weight sucrose concentration of 75 % (Elliot and Weston, 1993). This has made sugar beet a considerably more productive source of sugar than the earlier varieties. Plant breeders have selected for qualities such as increased mass alongside crops that are more resilient to abiotic and biotic stresses, and its suitability as a break crop and its value make it an attractive crop for UK farmers to grow.

Biotic stresses, such as diseases, can cause a reduction in root yield. Sugar beet is commonly infected with a variety of diseases including viruses, bacterial and fungal pathogens (Cooke and Scott, 1993). Viruses infecting sugar beet including yellows viruses, such as *Beet yellows virus* (BYV), the UK isolate *Beet mild yellowing virus* (BMYV) and its US counterpart *Beet western yellows virus* (BWYV) are all transmitted by aphids. Viral infections in plants lead to increased attack by fungal pathogens such as *Alternaria* species (Russel, 1960). Current control mechanisms adopted to prevent the spread of these viruses ensures that nearby plants being grown are crops that are resistant to the viruses. Other methods of control include the use of pesticides to prevent the aphid vector feeding on the crops. Some viruses, such as *Beet yellow stunt virus* (BYSV) are so prevalent in North America, that control is virtually impossible. It is widely agreed that the only way to eliminate these pathogens is to use genetically engineered/bred virus resistant crops (Cooke and Scott, 1993).

In the UK most sugar beet is harvested from September-March, and then transported to sugar beet processing factories (Asadi, 2007). When the sugar beet arrives at the factories, a sample of the crop is taken to analyse sucrose, potassium and nitrogen content of the taproot. The main bulk of the sugar beet is washed and sliced into cossetes, to create a larger surface area for sucrose to be extracted. The cossetes are pumped into a series of three diffusion towers where they are mixed with hot water (around 70 °C) and agitated. The cossetes are then pressed producing sugar water known as 'raw juice', and a solid pulp. The raw juice is then heated through recovery systems, and mixed with calcium hydroxide and carbon dioxide. The combination of these molecules produces calcium carbonate that removes any remaining impurities from the raw juice by forming precipitates. The precipitates can be removed using a filter, ensuring that excess magnesium and potassium are removed from the raw juice. The purified juice is now called 'thin juice', and enters an evaporation process where water is boiled off in six sequential evaporator vessels. This process turns the 'thin juice' into 'thick juice' by increasing the solid content of the juice from 16 % to 65 %. The thick juice is then crystallised by boiling, and the 'seeding' of small sugar crystals on which larger sugar crystals can grow. The product is a sugar-syrup mixture called massecuite. The crystallised sugar is removed by centrifugation, and cleaned. The syrup undergoes evaporation, and the seeding process again to produce more sugar crystals. This process is repeated a further two times, with the quality of sugar reducing each time. The first two batches are sold as white sugar and fine sugar, whilst the second two, and the molasses are re-used within the factory to produce more sugar during the seeding process (www.britishsugar.co.uk).

1.1.2.2 Secondary products

The process of extracting sugar from sugar beet gives rise to many alternative products. At the start of the process, the tops of the sugar beet (comprising of the leaves and the crown of the root) are removed. The tops are not wasted, as they can provide a cheap animal fodder to farmers for a variety of livestock. They can either be used fresh or wilted, or stored as silage to be used

throughout the winter (Hartland, 1993). During cleaning of the sugar beet roots, soil and stones are removed. These are separated and dried before the soil is sold as topsoil to landscapers, and the stones as cheap aggregate (www.britishsugar.co.uk).

The pulp produced during the sugar extraction process of pressing the cossetes after being through the diffusion tower can also be reused. They are mixed with molasses (also a product of the process) and dried at around 880 °C. The dried material is pressed into pellets which can be used as feed for livestock (Hartland, 1993; www.britishsugar.co.uk). Another direct product is produced during purification of the raw juice. When the juice is filtered with calcium carbonate to remove other elements, potassium and magnesium are often removed. These can be used as constituents in fertilisers and can be sold back to farmers (www.britishsugar.co.uk).

In September 2007, British Sugar became the first company to start producing bioethanol in the UK. It is manufactured by the fermentation of sugars from the sugar beet root, followed by distillation to produce a pure alcohol. Currently, petrol vehicles can run on fuel containing up to 5 % v/v bioethanol, but it is hoped that future models will be able to use a higher percentage of bioethanol, in order to reduce reliance upon fossil fuels (britishsugar.co.uk).

1.1.3 Oilseed Rape

1.1.3.1 History

Brassica napus, commonly known as oilseed rape or rapeseed, belongs to the Brassicaceae family. Brassicas are widely cultivated around the world, and can be divided into two subspecies; swedes (*ssp. napobrassica*), and *ssp. napus,* which include oilseed and vegetable rape. It is thought that the species arose in the Mediterranean by the crossing of turnip rape (*Brassica rapa*) and cabbage (*Brassica oleracea*), where both these species are native. It is also believed that the oilseeds were some of the first plants grown as crops, with records in India suggesting that oilseed brassicas were being cultivated in around 4000 BC, and later spread to China and Japan. By the 16th century

rapeseed was used as a source of lamp oil throughout Europe, and was heavily cultivated from the 18th century. Throughout the 20th century, rapeseed oil was only produced to serve as lubricant. The early oils produced from oilseed rape during this time were not suitable as a food oil, as they contained large amounts of erucic acid (up to 50%), which can lead to cardiac disease, as well as the fact that they tasted bitter. (Kole, 2007)

In the 1970s oilseed rape '0' varieties containing less than 1% v/v erucic acid were developed, meaning that their suitability as a food oil increased dramatically. Other issues with rapeseed prevented it from becoming a widely used crop. High glucosinolate levels were present in many varieties, making it unsuitable for livestock feeds as the glucosinolate could lead to kidney and liver damage in the animals. This problem was averted in 1969 when the Polish spring rape variety 'Bronowski' was found to have low levels of glucosinolate. This low level was due to three recessive genes in this variety. In order to reduce the levels of glucosinolate in other varieties, a backcrossing programme was founded to introduce the recessive forms of these genes into low erucic acid varieties. The result of these backcrosses was the spring rapeseed Tower, released in 1974. This variety contained 0 % erucic acid and low glucosinolate levels and has become one of the most important oil crop varieties produced (Kole, 2007).

Oilseed rape is now the most cultivated crop in Europe, and produced heavily in North America and China. It is the world's third leading source of vegetable oil (Kole, 2007).

1.1.3.2 Growth and harvesting

Brassica crops are one of the few oil crops that thrive at low temperatures. Their resilience to high elevations and cool growing conditions make them ideal for growing as a winter crop. There are four main types of *Brassica* grown worldwide for oil production; *Brassica napus*, *Brassica rapa*, *Brassica juncea*, and *Brassica carinata*. Of these four, *Brassica napus* is most

commonly grown in Europe as a spring and winter crop (Kimber and McGregor, 1995).

This bright yellow flowering plant produces mainly black seeds (sometimes yellow or brown). Vegetable oils can be extracted from these seeds as they contain up to 40 % oils. Once oil has been extracted, the meal residue contains up to 44 % protein and can be used in animal feeds. Harvesting of these crops is critical, as too early and the quality of the crop is reduced, too late and the pods containing seeds have shattered. Once harvested the seed crop is dried and pressed to release the oil, then filtered and bottled to be sold commercially (Kimber and McGregor, 1995; Kole, 2007; Stevens et al., 2008).

A wealth of competition, disease and insect pests are contended with when growing this crop. Weeds growing in the fields delay harvesting and reduce the quality of the crop due to contamination (Orson, 1995). As well as weeds, the bright yellow flowers attract many insects, many of which can cause serious damage to the stem, leaf, pod or root of the plant (Ekbom, 1995). Diseases such as stem rot, stem canker and leaf and pod spots, amongst others reduce yield and quality of the oil but viruses carried by insects pose a bigger problem for oilseed rape growers (Rimmer and Buchwaldt, 1995). Whilst weeds can be managed with herbicides and crop management, and insects and many diseases can be controlled with the addition of chemicals such as fungicides and insecticides, the treatment of viruses becomes more difficult (Ekbom, 1995; Orson, 1995). The pesticide treatment of the plants is sometimes effective, but in order to gain sustainable resistant to viruses, which can greatly reduce the yield of the crop, it is widely agreed that genetic resistance is required (Kole, 2007; Rimmer and Buchwaldt, 1995; Stevens et al., 2008).

As well as producing high yield, quality oilseed rape, pant breeders have been investigating resistance to viruses. With few varieties identified as resistant to these oilseed rape viruses, the genes responsible for resistance remain to be elucidated and utilised to produce a sustainably resistant option for plant growers (Kimber and McGregor, 1995).

1.1.3.3 Products of oilseed rape

The most well known product of oilseed rape is vegetable oil. These oils, for human consumption, are a valuable source of the essential fatty acids omega-6 and omega-3. These acids form important cell membrane components and play a role as precursors of several biologically active compounds (McDonald, 1995). After extraction of oil from the seeds, the remaining seeds (around 60 % of its original weight) are used as a supplement for animal feeds. This feed, mainly for cows but also pigs and chickens, is an excellent source of amino-acid balanced protein. Because of its high fibre content, it cannot be solely used as animal feed, but since the breeding of low glucosinolate varieties it has even been used as a food supplement for fish (Bell, 1995; Kole, 2007).

Other less-well known uses for oilseed rape crop also exist. Just as breeding varieties of oilseed rape containing low levels of erucic acid was possible to allow for human consumption, varieties containing high quantities of erucic acid were also developed. Erucic acid and behenic acids extracted from oilseed rape have thousands of applications including in shampoos, photographic film production, and even the production of nylon. Many of these products are not useful as a sustainable produce of oilseed rape, but contribute to the use of by-products from oil production (Sonntag, 1995).

A final use for the oil is the creation of Biodiesel. In the 16th century rapeseed oil was commonly used as a source of fuel for oil lamps, and in the early 1900's was used in combustion engines. As fossil fuels become increasingly expensive interest has returned to the use of rapeseed oils and their use as a renewable form of energy (Mittelbach, 1996). Biodiesel is particularly attractive because of its many environmentally friendly properties. This low toxic form of fuel is highly biodegradable, has a reduced risk of fire and low emissions associated with it, making it an increasingly attractive alternative to standard diesel (Korbitz, 1995).

1.1.4 Arabidopsis

Arabidopsis thaliana, or mouse ear cress is a model plant, widely used in molecular, genetic, and physiological study of plants. It is part of the Brassicaceae (mustard) family, which also contains cultivated species such as radish and cabbage. In its natural environment however, this plant could be regarded as an insignificant weed. It grows in the wild during the winter, and is thought to be indigenous to Europe, spread to North America by European settlers. There are many different phenotypic (and genotypic) varieties of Arabidopsis that vary with responses to factors such as day length, vernalisation requirements and dormancy (Réidi, 1975). More than 750 ecotypes of *Arabidopsis thaliana* have been collected from around the world, many of which are available at seed stock centres; Arabidopsis Biological Resource Centre (ABRC) and the European (Nottingham) Arabidopsis Seed Centre (NASC). These seed stock centres hold information about the accessions, development and physiological information (leaf shape, flowering time, disease resistance etc.) for researchers worldwide to use, learn from and update information (Swarbreck et al., 2008).

1.1.4.1 History of *Arabidopsis* research

Friedrich Laibach is widely considered to be the founder of research into *Arabidopsis*. In 1907 he published his findings stating the number of chromosomes in *Arabidopsis* (2n = 10) (Rédi, 1993; Laibach, 1907). Laibach then revisited Arabidopsis research in the 1940, putting forward an argument for the use of Arabidopsis in genetic research. Laibach also pointed out the extent of natural variation between *Arabidopsis thaliana* ecotypes as a tool for studying genetics, and began the study of mutagenizing plants using X-rays (Koornneef and Meinke, 2010).

Arabidopsis research grew in importance and significance steadily throughout the 1950s and 60s. George Réidi became the leading scientist in Arabidopsis research in the 1950s, when a scientific community was set up to include researchers from Germany, the Netherlands, the Czech Republic, Belgium

and the USA. This gave rise to an international Arabidopsis conference, the Arabidopsis Information Service (AIS) newsletter and most importantly a seed stock centre containing many accessions including those used by Laibach, as well as recently generated mutant lines (Meyerowitz, 2001; Somerville and Koornneef, 2002).

At this time however, it was believed that the future of plant research lay with tissue culture, so the field of Arabidopsis research did not advance quickly. It gained renewed interest in the search for a model organism for plant biology (Koornneef and Meinke, 2010). The publication of a paper about the suitability of Arabidopsis as a genetic tool by Rédei in 1975 jumpstarted interest in Arabidopsis as a model organism for molecular genetics. Amongst the advantages of Arabidopsis listed, were the relatively small chromosome number, short lifecycle, the ability to cross plants, the high number of seeds produced, and the small size of the plants (Rédei, 1975). In the late 1980's mutant screening of Arabidopsis was being carried out worldwide for analysis of it's physiology and biochemistry. This increased interest was matched with increased funding, allowing the drive of Arabidopsis research to continue and expand (Koornneef and Meinke, 2010).

An important step in the history of Arabidopsis research was the ability to successfully transform it using *Agrobacterium tumefaciens*. This procedure was first carried out using tissue culture methods (Lloyd et al., 1986; Valvekens et al., 1988), but soon developed into the whole-plant vacuum infiltration method (Bechtold and Pelletier, 1998), and more recently has been further improved in the 'floral-dip' method (Clough and Bent, 1998). Transformation technology has allowed the generation of a wide range of T-DNA insertion mutant lines, adding to the already large collection of naturally occurring ecotypes, to create a wealth of potential genetic resources. Transformation has also enabled the ability to introduce reporter genes to plants, allowing investigation of localised gene expression in specific areas of plants (Cutler et. al., 2000; Haseloff and Amos, 1995). Although there is no specified 'wild-type' Arabidopsis ecotype, it has been widely agreed that

ecotype Columbia (Col) is the reference ecootype, which was originally collected by Rédei (Koornneef and Meinke, 2010; Réidei, 1992).

Genetic linkage maps of the Arabidopsis genome allowed estimations to be made about the positions of genes along chromosomes. A completed map, containing 76 markers along the five chromosomes was published in 1983 (Koornneef and Van der Veen, 1983), closely followed by the introduction of molecular markers. Techniques such as restriction fragment length polymorphism (RFLP) and cleaved amplified polymorphic sequence (CAPS) markers allowed the enhancement of genetic linkage maps. More recently genomic sequencing has become more important for defining the location and order of genes, however molecular markers are still important in order for map-based cloning to be carried out, and the current number of molecular markers that have been mapped in the Arabidopsis genome is 335 (Maarten Koornneef and Meinke, 2010; Meinke, et.al 2009).

The Arabidopsis genome sequencing project was initiated in 1989, and aimed for the sequence to be completed by 2000. Scientists from across the globe collaborated on the project, managing to publish the complete genome on time and within its budget (AGI, 2000). The next challenge involved elucidating the function of all the sequenced genes. This has been carried out using a number of research techniques, including knock-out studies and microarray chips. A centralised database; 'The Arabidopsis Information Resource' (TAIR) holds much of the information concerning the genome sequence and its function. This information is freely available meaning that laboratories worldwide, with a diverse range of interests, are able to utilise and enhance the existing data (Koornneef and Meinke, 2010).

There is, of course, still a lot of knowledge to be discovered regarding Arabidopsis, but this organism is used widely as a model for other plants. By using Arabidopsis information and understanding its importance, comparisons can be drawn with other plants, and can prove particularly useful to breeders of crop plants. Many of the crop plants currently grown have complicated genomes, and complex molecular interactions. By first understanding

molecular mechanisms in the relatively simple Arabidopsis, the function of similar mechanisms in more complex plants can be more fully understood.

1.2 Plant Viruses

The first purification of plant virus particles was carried out by Chamberland in the late 19th Century. His experiment used a filtration method, passing plant extract through porous porcelain. As the pores became smaller, bacteria and other small organisms were filtered out. At the time it was thought that all plant diseases were caused by micro-organisms, such as bacteria, and therefore unable to pass through the smallest pores, but a disease causing agent was found in the most pure filtrate (van der Want and Dijkstra, 2006). The pattern of the disease being caused by the filtrate in tobacco plants was described by Adolph Mayer in 1882, and named 'mosaic disease'. Mayer described that no fungus was present in the plant, and although the disease was able to be mechanically transmitted to other plants, no bacteria were able to be cultured from the sample. Mayer postulated that a "soluble enzyme-like infectious principle" was responsible (Smith, 1894).

Beijerinck (1898) carried out similar experiments, and was the first to recognise the causative agent as a new type of micro-organism. After storing sap of an infected plant for three months, he found the disease-causing agent to be still active when re-infecting other tobacco plants. Beijerinck was the first to use the term 'virus', noting that the virus must have multiplied, as a small amount of sap from one leaf, was capable of infecting multiple additional leaves (Beijerinck, 1898; van der Want and Dijkstra, 2006). We now define viruses as sub-microscopic infective particles, made primarily of a protein coat containing nucleic acid information, making it capable of replication inside a host cell. There are three categories that distinguish viruses from other cells, these are: the lack of a continuous membrane that separates the virus from the host, the lack of protein synthesis machinery within the virus, and replication is carried out by the synthesis of a series of proteins (Hull, 2009; Matthews, 1991).

The virus discovered by Chamberland is now commonly referred to as *Tobacco mosaic virus* (TMV), and in 1928 Helen Beale established that sap from diseased plants contained antigens, drawing the conclusion that viruses acted like proteins, as antigens (Beale, 1931). This presented a new method for studying viruses, using serological techniques that were utilised to study the infection levels of crops. The ability to purify and characterise this virus has lead to the discovery of thousands of other viruses infecting plants and animals (van der Want and Dijkstra, 2006; Van Slogteren and Van Slogteren, 1957).

One of the main reasons research into plant viruses is so driven, is because of their damage to crop plants. It is often difficult to estimate the economic losses caused by infection with plant viruses. This value is also usually underestimated and variable, due to the difficulty of testing plants, variation between regions, and the different methods used to determine virus titres. There are many types of damage that viruses can cause plants, either directly or indirectly. Direct methods include reduction in growth and crop yield, poor quality yield, and a reduced ability to seed, or propagate. Indirect damage leads to problems such as an increased susceptibility to drought or other pathogens (Hull, 2009). Infection can also be systemic, where it has spread throughout the plant through the vascular system, or is restricted to the site of infection (Matthews, 1991).

There is great diversity between plant viruses, with viruses capable of infecting almost all crops grown for food around the globe (Hull, 2009). The genetic diversity between viruses is also extensive. The divergence of the nucleotide sequence of viruses in the same species is much greater than the divergence of any other known species on earth. These differences have accumulated through mutation recombination and re-assortment (Roossinck, 1997). Plant viruses can be divided into six categories based upon their genetic material. These categories cover double stranded DNA (dsDNA), single stranded DNA (ssDNA), double stranded RNA (dsRNA) and single stranded RNA (ssRNA), and can be seen in Figure 1.1.

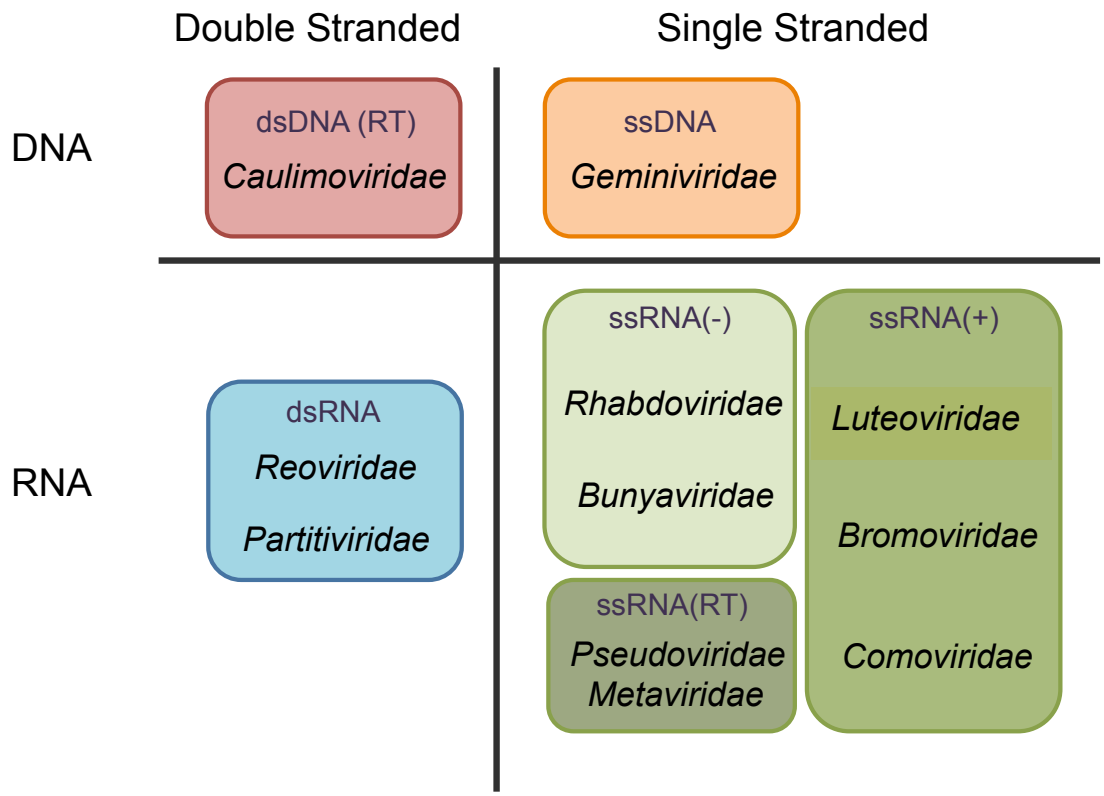


Figure 1.1 Classification of plant viruses based on their genetic information. The family names of infectious plant viruses have been grouped according to their genetic information. 'RT' refers to reverse transcribed, 'ds' double stranded, 'ss' single stranded, '-' negative sense, '+' positive sense. The size of the box approximately represents the number of genera within each family, with ssRNA(+) being the largest family, and ssRNA (RT) the smallest (Fauquat et.al, 2005)

Many mechanisms to prevent virus spread have been developed, for example, the use of pesticides to prevent the spread of virus through their insect vectors, but these have disadvantages in the indiscriminate way they kill insects. Virus resistant or tolerant crops have been bred by using different genotypes of plant species for many years, and this is still one of the most attractive options for creating virus-resistant crops. Plant breeders collect cultivars, and their wild relatives, to test them for susceptibility to diseases, and use those that are most tolerant or resistant to disease to continue breeding (van der Want and Dijkstra, 2006).

1.2.1 Virus diseases of sugar beet and oilseed rape

Sugar beet is susceptible to many viruses that can be transmitted by fungi, nematodes, physical contact, and most commonly aphids. From an economic point of view, all viruses reduce the quantity or the quality of the root. Because of this economic impact, the most devastating viruses infecting sugar beet have been investigated, and their genomes sequenced (Stevens et al., 2006). Table 1.1 lists some of the most important virus diseases identified to date.

Table 1.1 Viruses known to infect sugar beet, and their genus. Information adapted from Stevens et al., 2006.

Disease	Name	Abbreviation	Genus
Soil borne viruses			
Rhizomania	*Beet necrotic yellow vein virus*	BNYVV	Benyvirus
Beet soil-borne	*Beet soil-borne virus*	BSBV	Benyvirus
Beet soil-borne mosaic	*Beet soil-borne mosaic virus*	BSBMV	Benyvirus
Beet Virus Q	*Beet Virus Q*	BVQ	Pomovirus
Beet Oak Leaf	*Beet oak leaf virus*	BOLV	-
Beet distortion mosaic	*Beet distortion mosaic virus*	BDMV	Potyvirus
Yellows Viruses			
Yellows	*Beet mild yellowing virus*	BMYV	Polerovirus
	Beet western yellows virus	BWYV	Potyvirus
	Beet chlorosis virus	BChV	Polerovirus
	Turnip yellows virus	TuYV	Polerovirus
	Barley yellow dwarf virus	BYDV	Luteovirus
	Beet yellows virus	BYV	Closterovirus

Other viruses			
Beet Cryptic	*Beet cryptic virus*	BCV	Alphacryptovirus
Beet Curly Top	*Beet curly top virus*	BCTV	Curtovitus
Beet leaf curl	*Beet leaf curl virus*	BLCV	Nucleorhabdovirus
Beet Mosaic	*Beet mosaic virus*	BMV	Potyvirus
Beet Savoy	*Beet savoy virus*	BSV	-
Beet Yellow Net	*Beet yellow net virus*	BYNV	Luteovirus
Beet Yellow Stunt	*Beet yellow stunt virus*	BYSV	Closterovirus
Beet Yellow vein	*Beet yellow vein virus*	BYVV	Benyvirus
Cucumber Mosaic	*Cucumber mosaic virus*	CMV	Cucumovirus
Lettuce Yellow	*Lettuce infectious yellows virus*	LIYV	Crinivirus

In the UK oilseed rape is grown widely around the UK, and whilst many fungal and bacterial pathogens cause damage and restricted growth to plants, there are only three well known viruses that have a major effect on oilseed rape. These are listed in Table 1.2. The only one that causes major crop problems in the UK is *Turnip yellows virus* (TuYV) (Stevens et al., 2008).

Table 1.2 Viruses known to infect oilseed rape, and their genus. The most important viruses in the UK. Information adapted from Gladders et al., 2006.

Disease	Name	Abbreviation	Genus
Cauliflower Mosaic	*Cauliflower mosaic virus*	CamV	Caulimovieus
Turnip Mosaic	*Turnip mosaic virus*	TuMV	Potyvirus
Turnip Yellows	*Turnip yellows virus*	TuYV	Polerovirus

Many of these viruses are capable of infecting multiple hosts, and therefore replicate and survive throughout different growing cycles of plants. Understanding how these viruses interact with their hosts, through the

development of specific serological tests, has allowed the identification of more virus species and an understanding of how these viruses interact with their hosts, and other viruses (Stevens et al., 2006).

1.2.2 Virus Yellows

The yellowing of leaves of sugar beet has been observed since 1910, and was discovered to be caused by a virus, transmitted by aphids in 1936 (Stevens et al., 2006). At this time the yellowing of sugar beet leaves was thought to be caused by one virus, *Beet yellows virus* (BYV). In 1952, it was proposed that numerous viruses may cause the yellowing effect observed, as isolates from Ireland were not able to be precipitated with British BYV antiserum (Watson, 1940). Other viruses, *Beet mild yellowing virus* (BMYV) and *Beet western yellows virus* (BWYV) were identified in the UK and USA respectively, in the 1950s (Stevens et al., 2006).

There are now four recognised yellows viruses, BMYV, BWYV, *Beet chlorosis virus* (BChV), and *Turnip yellows virus* (TuYV). Only three of these can infect sugar beet; BMYV, BWYV and BChV. The fourth yellows virus was originally thought to be a UK isolate of the USA BWYV. This transpired not to be the case however, as TuYV cannot infect *Beta* species, it does however infect oilseed rape and was renamed TuYV in 2002 (Stevens et al., 2005, 2006). In oilseed rape TuYV does not lead to the yellowing of leaves, but instead leads to a reddening of the leaves, a symptom which is often confused with ‘stress’ symptoms, and therefore incorrectly treated (Gladders et al., 2006).

It was originally difficult to discern between BMYV and BWYV (syn. TuYV) as antiserum could not be generated that could distinguish between them. It wasn’t until antibodies were raised to *Barley yellow dwarf virus* that they could be serologically detected as different viruses (D’Arcy et al., 1989). The differences between the viruses were confirmed when sequencing of the 5’ termini showed significant differences, providing evidence that BMYV and BWYV (syn. TuYV) are independent species (Guilley et al., 1995; Lemaire et al., 1995). All the yellowing viruses identified belong to the family *Luteoviridae*,

and the two being studied in this investigation, BMYV and TuYV, are both form the genus Polerovirus.

Poleroviruses, like most of the Luteoviridae family, cannot be transmitted by surface contact with a virus. The viruses require wounding in order to be able to infect vascular tissues of the plant (they are phloem-restricted). Experiments involving BMYV found that virion infection was limited to the vascular bundles, and was found in all the cells of the petioles and stem of the infected plant (Stephan and Maiss, 2006).

1.2.3 *Beet mild yellowing virus*

Sugar beet infection with beet mild yellowing virus causes the leaves to become yellow and brittle, as shown in Figure 1.2. Secondary pathogens, such as the fungus *Alternaria alternata* are able to take advantage of the viral infection to enable colonisation, leading to the death of the leaf. The chlorosis of the leaf also disrupts photosynthesis and respiration, meaning that the plant cannot generate the energy required to grow to its full potential, often leading to reduced root yield. Disruption to metabolic processes also causes levels of nitrogen, sodium and potassium in the root to fluctuate, which causes the process of sugar extraction to become more complex (Stevens et al., 2006).

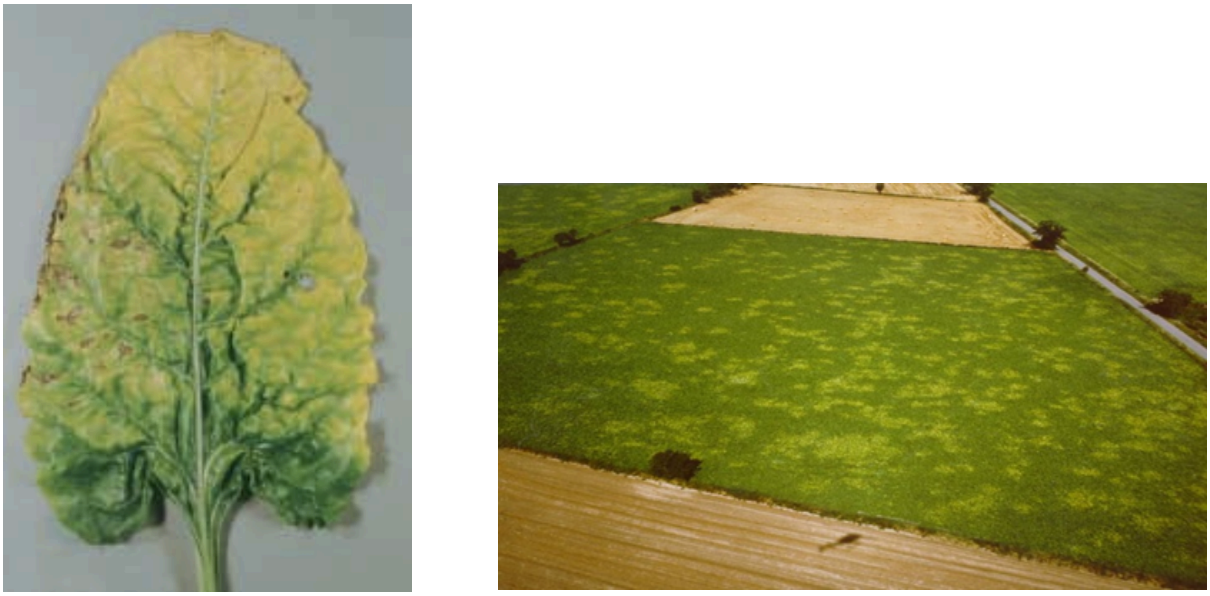


Figure 1.2 Sugar beet leaf infected with BMYV. Left: classic symptoms of BMYV infection can be seen with chlorosis and thickening of the leaf. Right: a field patchily infected with BMYV (Stevens et al., 2005).

BMYV infection in the early development of sugar beet can lead to a yield reduction of up to 30 % (Stevens et al., 2004). Infection at later stages of growth, where the plant has more than 20 leaves, does not occur as often, and therefore is thought to have only a small impact on sugar beet growth and quality. It has been estimated that 1.8 % of the total sugar beet produced in the UK is lost predominantly due to BMYV infection, equating to a financial loss of £5.5 million per year (Jaggard et al., 1998).

BMYV is capable of infecting a wide range of plants, including commercial crops such as fodder beet, red beet and spinach. Many weed species can also be infected with BMYV. This is especially as the virus is therefore able to survive in weed reservoirs around fields where crops are being grown, meaning that in the next growing cycle of crops, the viruses are still present in the local area (Russell, 1965; Stevens et al., 2006).

BMYV virus particles are icosahedral in shape, and around 26 nm in diameter. The genome is a single stranded positive sense RNA strand, contained within a protein shell. The BMYV genome is 5.7 kb, and was sequenced in 1995

(Guilley et al., 1995). It contains six open reading frames (ORFs); three 3' ORF's and three overlapping 5' ORFs (Stephan and Maiss, 2006). The virus is transmitted by *Myzus persicae* (green peach aphid) in a circulative non-propagative manner (Stevens et al., 2006).

1.2.4 *Turnip yellows virus*

Shortly after the identification of BMYV in the UK, a similar virus in the USA was identified, BWYV. These two viruses have similar hosts, capable of infecting important crops such as sugar beet, and both causing chlorosis of the leaves. In the UK a third virus was then identified, a BWYV-like virus which was capable of infecting crop plants, such as lettuce, that were resistant to BMYV, but susceptible to BWYV in the USA. It was assumed that this virus was a UK isolate of BWYV (Duffus and Russell, 1970; Stevens et al., 2008). The BWYV-like virus and BWYV were soon shown to be biologically and serologically similar, but the host ranges of the viruses were show to be more distinct. Unlike the USA BWYV, the UK BWYV-like virus was unable to infect sugar beet (Duffus and Russell, 1970). In 2002 the BWYV-like virus was recognised as an independent virus, and named *Turnip yellows virus* (TuYV) (Mayo, 2002). This difference was cemented when significant genomic differences were found between TuYV and BWYV in the 5' terminal region (Beuve et al., 2008; Hauser et al., 2000; Stevens et al., 2008).

TuYV is thought to be one of the most economically important disease affecting oilseed rape in the UK, causing plants to produce a decreased quantity and quality of product. Once infected with TuYV, oilseed rape leaves redden around the edges, and show signs of vascular reddening or yellowing as seen in Figure 1.3. Dwarfing of the plants is another indicator of infection with TuYV, common to many host plants infected with poleroviruses. These symptoms are often mistaken for symptoms of a nutrient deficient soil, and therefore the disease goes unnoticed, and untreated (Juergens et al., 2010; Stevens et al., 2008).

Figure 1.3 Oilseed rape leaves infected with TuYV. Symptoms of leaf infection include reddening and chlorosis of the leaf at the peripheries, and are often confused with nutrient deficiency symptoms (Stevens et al., 2008).

Oilseed rape infection with TuYV causes stunted plants, which produce fewer primary branches. This reduced size along with reduced leaf area and chlorosis of the leaves means the plant produces fewer seeds per pod, and of poorer quality than those of uninfected plants (Jay et al., 1999). Infection in these plants can lead to a 13.4 % reduction in yield of oil from the plant (Smith and Hinckes, 1985), although this Figure is debated worldwide, with some growers in Australia detecting a higher yield loss (Jones et al., 2007). In the UK, Stevens (2008) has estimated that TuYV infection in oilseed rape could cost £30-40 million per year.

As well as oilseed rape, TuYV is capable of infecting a broad range of host species from 13 plant families. Many of these are important agricultural plants, including lettuce spinach, pea and turnip. Similarly to BMYV, TuYV can also infect many weed species, meaning that reservoirs of viruses can persist throughout growing seasons (Stevens et al., 2008).

The TuYV genome consists of a single-stranded positive sense RNA encoding expression control and coat protein genes. The particles are typical for poleroviruses encoding three 3' open reading frames (ORFs) and the overlapping 5' ORFs. The virion is spherical and non-enveloped, with icosahedral symmetry of 180 coat proteins (Brault, et al. 2011; Stevens et al., 2005).

1.2.5 Genomic organisation of Poleroviruses

Poleroviruses (including TuYV and BMYV) have the same basic genome structure consisting of a linear, positive sense, single stranded RNA molecule. The genome is between 5.3 and 5.7 kb and encodes six ORFs, three 3' associated ORFs and three 5' ORFs. The three 5' ORFs are expressed from genomic RNA, and they are separated from the 3' ORFs by a 200 nucleotide non-encoding region. The three 3' ORFs are encoded by a subgenomic RNA. The six ORFs are the three 5' ORFs; ORF-0, ORF-1, and OFR-2, and the 3' ORFs; ORF-3, ORF-4 and ORF-5. The genomic and subgenomic ORFs, and their positioning are shown in Figure 1.4.

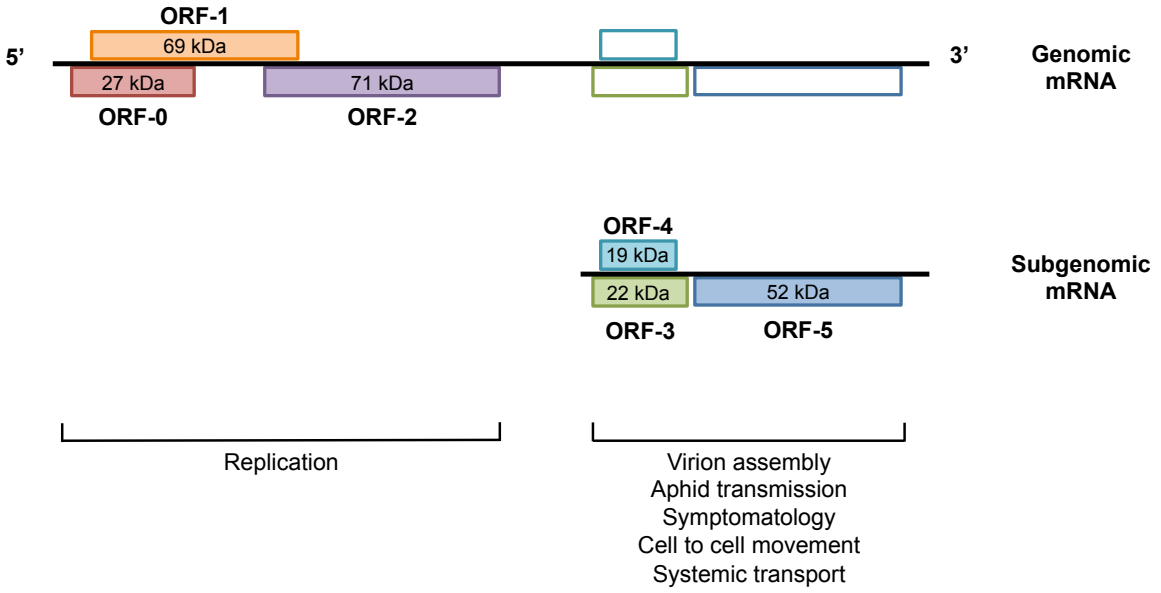


Figure 1.4 Organisation of and expression of Polerovirus genome. The genomic and subgenomic ORF expression is distinguished, along with the protein product activity for the 3' and 5' ORFs (adapted from Stevens et al., 2005).

Poleroviruses use multiple strategies for expression of the six ORFs including leaky scanning, ribosomal frame shift, protein fusions and read through domains, which will be discussed in more detail in the following sections (Martin et al., 1990; Stevens et al., 2005).

Poleroviruses have neither a 5' cap structure nor a polyA tail that host mRNA sequences contain. Instead, at the 5' end of the genome, poleroviruses have a viral genome linked protein (VPg). VPgs are small proteins that are thought to have many roles including acting as an RNA synthesis primer, signalling for RNA encapsulation, and a mediator of RNA translation and protein synthesis (Skaf et al., 2000; Wimmer, 1982; Wittmann et al., 1997). The VPg, a product of a fusion protein from the products of ORF1+2, has been shown to be vital for successful host infection through mutagenesis experiments in *Pea enation mosaic virus* (Skaf et al., 2000). In the Potyvirus family, it has been well documented that the VPg can interact with host translational machinery in order to initiate viral translation (Beauchemin et al., 2007).

1.2.5.1 5’ ORFs – Translation and function

The 5’ ORFs are translated directly from genomic RNA. Translation is initiated at an AUG start codon, which is preceded by a short untranslated leader sequence. The leader sequence often assists the RNA to form secondary structures. Translation of ORF-0 begins at the AUG start codon, to produce protein ‘P0’. The function of P0 is currently disputed, as the protein has not yet been described *in planta* (Stevens et al., 2005). The importance of the gene for virus accumulation is clear however, as mutations in ORF-1 of *Potato leafroll virus* (PLRV) completely eliminated the ability of the virus to accumulate in plant cells (Sadowy et al., 2001). The expression of PLRV P0 in potato plants leads to expression of disease symptoms in the leaves, suggesting that P0 may also play an important role in the appearance of symptoms (Brault et al., 2011). In TuYV (as well as other poleroviruses), P0 has been shown to have a role in the suppression of host post-transcriptional gene silencing (PTGS) mechanisms (Pfeffer et al., 2002). The host cell adopts PTGS to degrade RNA as a defence mechanism against viral infection. A region in the P0 protein has been identified as a potential F-box motif (important for protein-protein interactions), which is essential for P0 function. The P0 protein is thought to interact with key components in RNA-induced silencing complexes (RISC), making the complex inefficient by targeting individual proteins before the RISC is formed (Bortolamiol et al., 2007; Pazhouhandeh et al., 2006). P0 has no effect however on the fully formed RISC complex (Brault et al., 2011; Csorba et al., 2010).

ORF1 and ORF2 encode P1 and P2 respectively. These proteins are either expressed as P1 individually, or as a P1-P2 fusion protein. Both domains have been shown to be essential for viral infection and replication through mutagenesis analysis (Reutenauer et al., 1993). The P1 protein of PLRV encodes a protease closely followed by a domain thought to encode the genome-linked protein (VPg) in BMYV, TuYV, PLRV as well as other poleroviruses (van der Wilk et al., 1997). The proteinase encoded in P1 is thought to play an important role in the maturation of the VPg, also encoded by P1, by cleaving a site in the protein close to the N terminus of the VpG. In

experiments where the function of the proteinase was removed, a large protein was produced, and an un-cleaved and non-functioning VPg (Li et al., 2000). The P2 protein encodes an RNA-dependent RNA-polymerase (RdRp), which uses the VPg as a primer to replicate the viral genomic RNA (Martin et al., 1990).

The P1-P2 fusion protein is thought to be expressed by a ribosomal frame shift during the translation of the P1 protein. It is believed that the expression of P1 is more common than the expression of P1-P2, as ribosomal frame shifts are a fairly rare event (Stevens et al., 2005). A pseudo knot in the RNA structure causes the ribosome to pause, allowing time for anticodon:gRNA realignment, (Alam et al., 1999) and a -1 frame shift, which has been identified in the ORF-1 ORF-2 overlap region (Cornish et al., 2006; Nixon et al., 2002).

1.2.5.2 3' ORFs - Translation and function

Translation of the 3' proximal ORFs; ORF-3, ORF-4 and ORF-5, occurs after synthesis of the subgenomic RNA. The subgenomic RNA is thought to be dependent on the activation of the RdRp found in the P1-P2 fusion product at an internal promoter on the '-' (minus) strand produced during genomic RNA replication (Stephan and Maiss, 2006; Stevens et al., 2005).

The P3 protein encodes for the major capsid protein (CP) of the virus, and although this protein is not essential for RNA replication, when the protein is mutated the amount of RNA produced is greatly reduced (Reutenauer et al., 1993). The CP does play a vital role in the association of the virus with the aphid vector in order for transmission of the virus to take place (Gray and Gildow 2003; Torres et al., 2005). Further investigations into the importance of the coat protein suggest that it is essential for the infection of whole plants. A TuYV strain with a mutation in ORF-3, lead to an uninfected plant, except at the point of inoculation (Ziegler-Graff et al., 1996).

ORF4 is found within ORF3, but in a different reading frame. Similarly to ORF2, ORF4 relies upon a ribosomal frame shift for expression. It has been

shown that for most Poleroviruses, the ORF4 start codon is in a more favourable sequence context to that of ORF3, meaning that expression of the gene is not detrimental by being the second ORF. It is not clear about the exact ratios of ORF expression, but studies on PLRV have suggested that both ORFs are equally translated at a 1:1 ratio (Juszczuk et al., 2000; Stevens et al., 2005). The P4 protein is required for systemic spread of the virus throughout the plant, and has been shown to associate with the nuclear envelope of the host cell (Xia et al., 2007).

The P5 protein is produced as a fusion protein with P3. Expression of this protein relies upon a read-through of the P3 stop codon. This is caused by the suppression of the amber stop codon in ORF-3, and therefore the protein is known as the read through domain (RTD) protein (Brault, et al., 2011; Stevens et al., 2005). Experiments have shown that the RTD make up some of the 180 CP proteins, with the P3 coat protein making up the bulk of the CP, and the P3+P5 RTD making up a smaller amount (Peter et al., 2009). Whilst the virion is still able to be created with the RTD removed, they are not capable of being transmitted by aphids, and also have a reduced capability for systemic infection in their host (Brault et al., 2011; Reinbold et al., 2013; Reutenauer et al., 1993).

1.2.6 Transmission by Aphids

Like many Poleroviruses, both TuYV and BMYV are transmitted by *Myzus persicae*, commonly known as the green peach aphid, or the potato aphid. The viruses cannot be mechanically transmitted, or transmitted through seed. Aphid transmission is circulative, persistent, and non-propagative. The viruses cannot replicate inside the aphid, only in their host cell, and are able to be re-infected into other host plants (Stevens et al., 2008).

Figure 1.5 *Myzus persicae* or the green peach aphid. Taken from (Stevens et al., 2008)

The aphid acquires the virus whilst feeding on an infected plant. The aphids sharp stylet mouthpiece breaks the cell wall of the phloem sieve element or companion cell, and feeds within 15-30 minutes of initiation (Gray and Gildow, 2003). Virus can be ingested throughout this time from within 1 minute. The aphid epithelial gut cells actively transport virus particles through the cell and release them into the hemocoel. It then takes 12-16 hours for the virus particle to circulate through the aphid and be released into the aphid hemocoel (Garret et al., 1996). Once inside the aphid hemocoel, the virus is able to survive outside of a cell, little is known about this process, but it is thought to involve aphid associated factors (van den Heuvel et al., 1999). The virus is able to survive in the aphid hemolymph for several days, where it gradually makes its way to the salivary gland through passive transport. During feeding the aphid releases salivary enzymes to assist with tissue penetration and feeding. It is then that the viruses are also secreted into the plant, within 30 minutes of the aphid beginning to feed (Gray et al., 1994, 1991). A diagram showing the route of the virus through the aphid is shown in Figure 1.6.

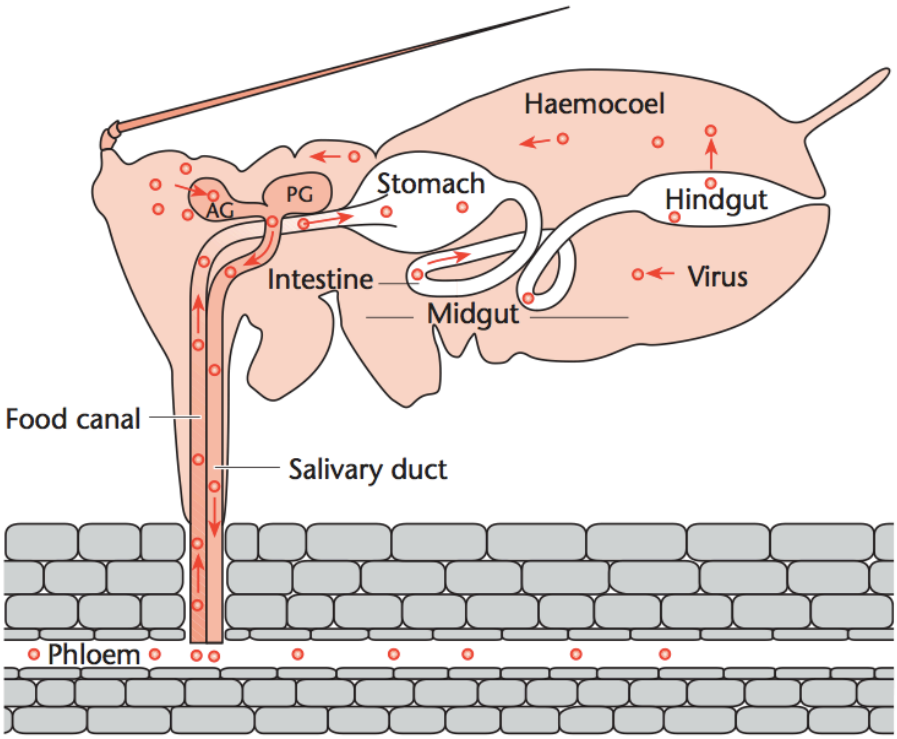


Figure 1.6 The route of the virus through the aphid. AG – accessory salivary gland; PG - principal salivary gland. Taken from (Brault et al., 2011)

Experiments with TuYV have indicated that once a virus is acquired by an aphid, after 24 hours the virus could be transmitted to a new host, but the virus can survive within the aphid for as many as four days. Many things can affect the transmission of the virus, including temperature and humidity, and the suitability of the plant to become a host (Stevens et al., 2008).

Transmission rates of TuYV by *M. persicae* have previously been reported at over 90% (Schliephake et al., 2000), with UK wide sampling of aphid populations suggesting that up to 72 % of aphids carry TuYV (Stevens et al., 1995). The weather also has a large impact on aphid numbers, with warmer winters, more aphids survive and are able to carry the disease to crops, such as oilseed rape, which are typically grown in winter. The aphids which survive in oilseed rape crops over the winter then cause extensive spread of the viruses throughout the following spring (Stevens et al., 2008). From late October through to April, aphids tend to survive around weeds, brassicas and

winter oilseed rape, beginning their migrations at the start of May to brassicas, sugar beet, and lettuce, before beginning migration back again in October. In this way, crops year-round can be infected with TuYV, and other viruses (Stevens et al., 2008).

1.2.7 Current control methods

Aphids are able to enter crop fields year round, presenting a high risk of infection with viruses. The transmission of BMYV and TuYV by aphids means that insecticides are commonly used to control virus spread. Seed treatment and foliar sprays are used to prevent the aphids spreading virus to further plants, but when faced with large numbers of aphids, and since transmission occurs so rapidly, the control of aphids becomes difficult (Stevens et al., 2006). The tight control over the use of pesticides, and the often-indiscriminate nature of the insecticides means that the use of these treatments are not regarded as an 'environmentally friendly' option. It has also been well documented that aphids are becoming increasingly resistant to commonly used insecticides, meaning other control methods must be sought (Stevens et al., 2008). Recently government regulations have tightened over the use of neonicotinoid insecticides. These neonicotinoids were developed to protect crops from aphids and other insects, and were used with some success in Australian oilseed rape crops to reduce TuYV infection (Stevens et al., 2008). The restricted use of these chemicals makes the search for other methods of control more pressing. Crop growers are therefore encouraged to not only use insecticides, but also adopt other prevention strategies.

Other strategies include the removal of infected material and reservoir plants from around fields in order to reduce the amount of virus in the local area that aphids could feed on. The large host range of these plants often makes that a difficult task. By sowing crops early, the amount of virus in the plant is reduced, as older plants tend not to become as infected, or as affected by the viruses, meaning a minimal reduction to yield (Brault et al., 2011; Stevens et al., 2008, 2006).

New technology has allowed advances in aphid forecasting. By being able to give preliminary forecasts of aphid movement around the UK, from Brooms Barn research station in Suffolk, crop growers can be alerted to any close-by aphid movement, allowing them to adapt their strategies to reduce the risk of infection (Stevens et al., 2008).

For many growers and scientists, the best method of virus control would be the use of genetically resistant crops. This would provide a more 'environmentally friendly' approach to disease resistance, without the heavy use of pesticides (Stevens et al., 2008). In crop varieties however, no major source of BMYV resistance has yet been found, but varieties of sugar beet with moderate resistance to BWYV have been developed in the USA, and will be used in areas with widespread BWYV infection (Stevens et al., 2006).

1.3 Host Resistance Mechanisms to Viruses

There are many viruses that have been characterised at a molecular level, which are capable of infecting a wide range of host plants, but it is true that most of these viruses are unable to infect most plants. When all the members of a species are resistant to a pathogen, it is defined as 'non-host' resistance. The reason for this resistance becomes clear when the diversity of plant viruses is considered. In order to infect a host plant, the virus must specialise to overcome host resistance mechanisms, and be able to utilise host replication machinery. As not every host plant uses the same mechanisms, viruses have become specialised to infect host plants with more similar mechanisms, or similar solutions to overcome defence mechanisms. When plants of a virus susceptible species include varieties/cultivars that confer genetic resistance to these pathogens, it is known as 'host' resistance. Studies of these resistance mechanisms has led to a greater understanding of both virus-host interactions, and also how to utilise these traits to create genetically resistant crops (Talbot, 2004).

There have been many examples of host resistance described including the use of resistance genes refered to as '*R*-genes', post transcriptional gene silencing (PTGS), and recessive resistance mechanisms. In every virus infection however, cells undergo similar responses such as alterations in host gene expression. The alteration of gene expression is usually a consequence of a defence response and typically leads to an accumulation of starch, increased respiration, decreased photosynthesis and an increase in the quantities of amino acids and organic acids within the cell. It has been frequently observed that as viruses move systemically around the plant, the older infected tissues change their gene expression patterns to a state where the cells can survive carrying out normal metabolic activity, whilst the virus is present. This is a key event to the virus, as it must be able to induce a state in the plant cell for it to be able to survive, whilst being able to overcome plant defences. It does this through a process called 'shut-off', where plant translation of genes is down-regulated, in favour of the up-regulation of viral genes. The shut off mechanism can occur in the plant cell merely because of the presence of the virus, non-specific shut-off, or because the virus acts upon the plant cell to decrease the amount of host translation, specific shut-off. Not all host genes are down-regulated during virus infection however. Many cell-to-cell movement protein expression levels are maintained throughout virus infection, and this is thought to aid virus movement throughout the plant. During infection with *Cucumber mosaic virus* (CMV) the heat shock chaperone protein Hsp90 is up-regulated, alongside an NADP+-dependent malic acid enzyme. The up-regulation of NADP+-dependent malic acid co-insides with an initial increase in photosynthesis within the cell. This is soon followed by a decrease in the amount of photosynthesis, as the amount of respiration taking place increases, with starch accumulating, and increased glycolysis activity. This leads to high levels of sucrose within the plant. It is thought that the sucrose, which would normally be exported around the plant through the phloem, is unable to be exported from the cells due to blockage with viral proteins, so instead the sucrose accumulates in the plant cell and is converted to hexose. The hexose inhibits photosynthesis, leading to chlorosis, and triggering downstream defence responses. A diagram of this response can be seen in Figure 1.7 (Talbot, 2004; Dickinson, 2005; Hull, 2009).

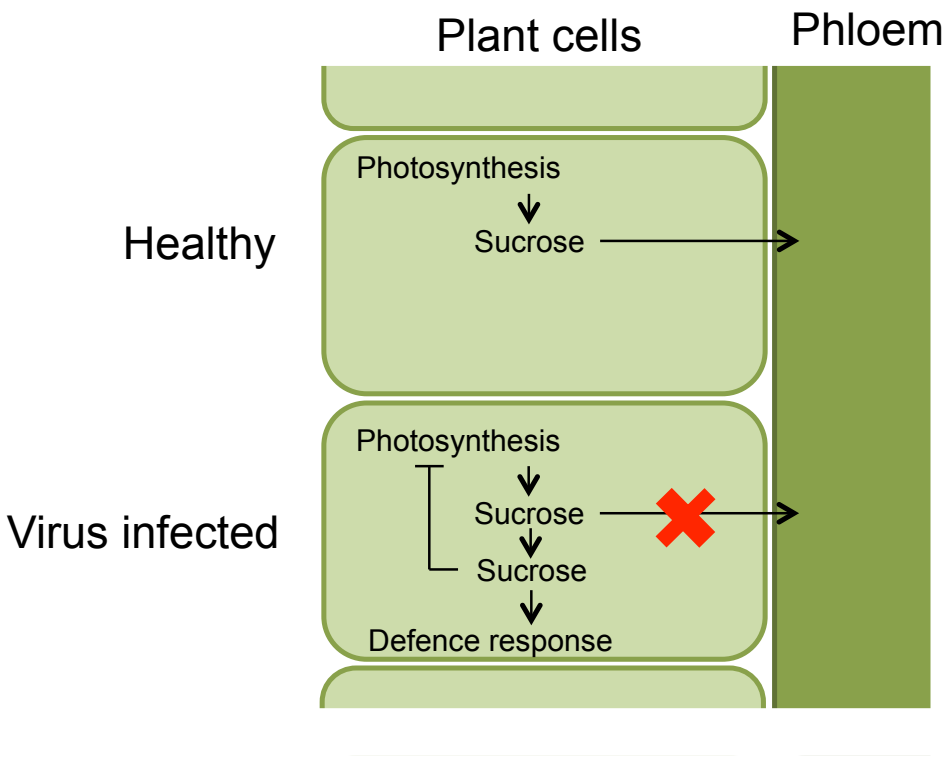


Figure 1.7. Genetic regulation within a cell in response to infection with CMV. CMV infection triggers an increase in NADP+ dependent malic acid enzyme, and an increase in photosynthesis. Sucrose is produced in the cell, and in a healthy plant is transported through by the phloem around the plant. In an infected plant the prescience of viral protein prevents transportation of the sucrose, which accumulates in the cell and is converted into hexose. Hexose inhibits photosynthesis, and triggers downstream host defence responses in the cell (Dickinson, 2005).

Natural barriers of the plant to viruses, for example a thick cuticle, and preformed chemical compounds, are an innate and physical resistance mechanism for the plant, but when breached, an active response to pathogen invasion is initiated (Dickinson, 2005).

1.3.1 The Plant Immune System

There are many types of pathogens that plants face on a daily basis including bacteria, fungi, oomycetes and viruses. The pathogens all have one thing in common; they rely on the plant host in order to complete their lifecycle. Unlike mammals, plants do not have immune defence cells that are able to travel around the plant, and instead rely on innate immune responses. These immune responses are triggered by effector molecules released by the pathogen. Effector molecules are often required by the pathogen in order to enter, colonise, break down walls or evade the host immune system (Jones and Dangl, 2006).

It is thought that there are two main divisions of the plant immune system. The first requires recognition of microbial pathogen associated molecular patterns (PAMPs) by pattern recognition receptors (PRR) in the plant cell membrane. These PRR are often able to recognise well conserved, important, slow evolving proteins such as flagellin (Jones and Dangl, 2006; Zipfel and Felix, 2005). These molecules induce PAMP triggered immunity (PTI) in plants, but not all plant species have the ability to recognise the same PAMPs, and response levels to PAMPS in different plants also varies (Zipfel and Felix, 2005). It is also thought that the recognition of these PAMPS causes plants to be on 'high alert', aiding in the recognition of other PAMSPs and increasing defence responses. These responses are mainly limited to bacteria, fungi and oomycetes however (Jones and Dangl, 2006).

The second division of the plant immune system is the use of polymorphic nucleotide binding leucine rich repeat (NB-LLR) domain proteins. These act mostly inside cells, and are usually encoded by *R* genes. *R* genes are specifically encoded disease resistance genes that recognise plant effector molecules, also known as avirulence effectors (*Avr*). This part of the immune system is often only triggered by the pathogen entering the cell, after it has overcome PTI. The response evoked by this second division of immunity is referred to as effector triggered immunity (ETI). The ETI response often leads

to a plant hypersensitive response, and most commonly cell death (Dangl and Jones, 2001; Jones and Dangl, 2006; Ting and Davis, 2005).

Jones and Dangl have developed a model, called the 'zigzag' model to demonstrate plant immune response levels, which can be seen in Figure 1.8. In this model PAMPs trigger a PTI response in cells. At this stage the pathogen is usually on the external side of the membrane, and so this is therefore an early response to pathogen infection. If the pathogen manages to evade the PTI response it enters the cell and uses effector molecules to increase the pathogen virulence level. Some of the effector molecules released by the pathogen can interfere with PTI signalling in order to allow the pathogen to colonise the host, and this is known as effector triggered susceptibility (ETS). Some of the effector molecules (*Avr*) can be recognised by NB-LLR proteins, encoded by *R* genes. This activates the ETI response that frequently leads to the HR, and cell death. Any surviving pathogens will go through the same ETS response upon the recognition of more *Avr* effector molecules released. In this way both effector molecules and *R* genes are gained and lost through natural selection (Jones and Dangl, 2006).

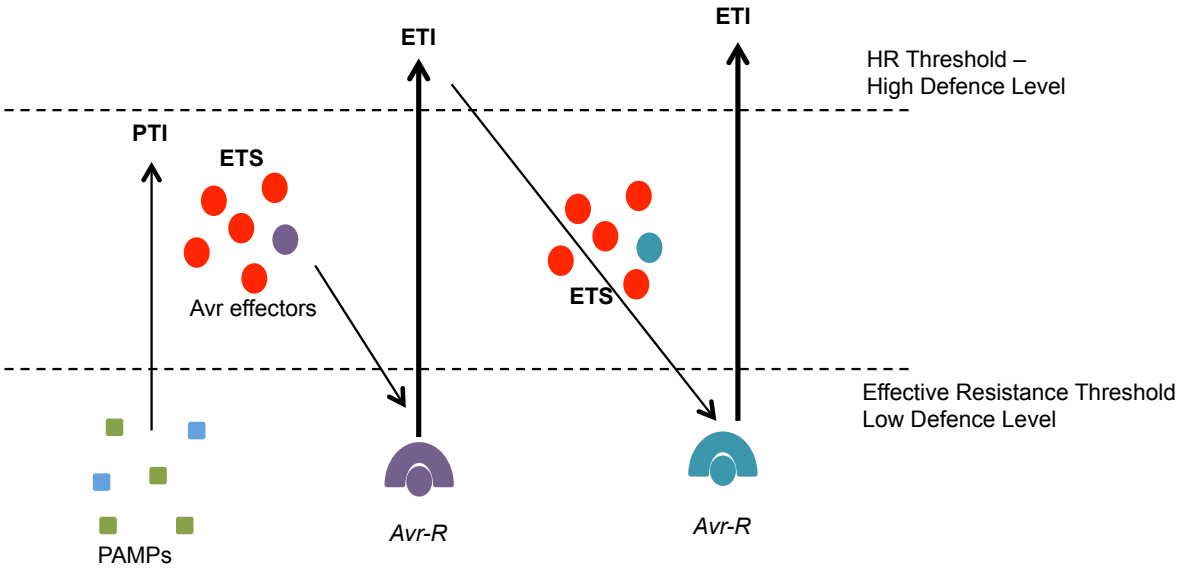


Figure 1.8 Plant immune response zigzag pathway. PAMPs (Pathogen associated molecular patterns) are recognised by PRRs (Pathogen recognition receptors) in the cell membrane. This causes a PTI (PAMP triggered immunity) response within the cell. Any pathogens that have successfully evaded PTI pass through the cell membrane where they release *Avr* effectors, which amongst other functions, can reduce the effect of PTI. These effectors may then be recognised by an NB-LLR molecule encoded for by *R* genes. This causes an ETI (effector triggered immunity) response which often leads to HR (hypersensitive response). Pathogens containing that effector (in the diagram purple) are then removed, and other *Avr* effectors are then recognised by NB-LLR in the plant. Adapted from Jones and Dangl, 2006)

Viruses have relatively small genomes that have evolved to contain genes required for pathogenicity. These genes could therefore all be classed as avirulence genes (*Avr*). *R* genes are encoded by the host to recognise pathogen *Avr* genes, in order to mount a defence response. If both the pathogen *Avr* and the host *R* are present, then the plant will be resistant to the pathogen (Dangl and Jones, 2001). Most *R* genes identified to date are monogenic and dominantly inherited (Maule et al., 2007). Arabidopsis genomes are known to encode more than 400 *R* genes, making up around 2 % of its genome (Dickinson, 2005). *R* genes are a good form of defence from viruses, and are a part of the natural host immune system. The recognition of *Avr* by *R* genes often results in a hypersensitive response (HR) of

programmed cell death (Flor, 1971; Fraser, 1990; Ritzenthaler, 2005). The active response can also lead to downstream signalling of defence mechanisms such as the thickening of cell walls to prevent virus cell-to-cell movement, and localisation of hydrolytic enzymes and other defence proteins (Dickinson, 2005).

R genes can be classified into eight distinct groups, depending upon their predicted protein structure (Martin et al., 2003). The largest group, which contains all *R* genes identified so far relating to plant virus resistance is the nucleotide binding – leucine-rich repeat (NB-LRR) class. These proteins contain a central nucleotide binding site, and a C-terminal , and can be sub-classified by their N-termini either having a coiled coil domain (CC) or a toll-interleukin receptor (TIR) region (Belkhadir et al., 2004; Ritzenthaler, 2005). Discussion continues regarding the exact mechanism of *R* gene and *Avr* interaction. Direct interaction has previously been demonstrated, where the *R* gene acts as a receptor for *Avr* (Jia et al., 2000). The 'guard-hypothesis' model proposes that in uninfected plants R proteins form complexes that act to inspect the plant for the presence of pathogens. When pathogens are recognised, defence responses would be initiated (Belkhadir et al., 2004; Dangl and Jones, 2001). Support for this theory was found through the activity of Nonrace-specific Disease resistance 1 (NDR1) in *A. thaliana*. The NDR1 protein was found to be anchored to the outside of the cell plasma membrane, an ideal location to detect the presence of pathogens and to induce defence signalling (Coppinger et al., 2004; Ritzenthaler, 2005). It has been suggested that these protein complexes instead of looking for interactions directly with *Avr* effectors, instead detect alterations in host proteins that are caused by pathogens (Glazebrook, 2005).

1.3.2 Recessive resistance genes

As well as the encoding *R-Avr* activated defence programs within a cell, recessive genes can provide effective resistance to plant viruses. Recessive resistance is brought about by homozygous recessive alleles encoding for a protein required by the virus to enter the cell, utilise the host replication

machinery or transport mechanisms. These resistance mechanisms are thought to convey effective and long lasting, durable resistance to plant pathogens (Dickinson, 2005; Maule et al., 2007; Ritzenthaler, 2005). This so-called 'passive' mechanism makes the virus particles incompatible with the host and is often referred to as a loss of susceptibility mutation. This type of mutation has been observed to have a significant effect on many viruses as they are unable to complete their lifecycle, but so far very few recessive resistance genes have been identified (Diaz-Pendon et al., 2004; Ritzenthaler, 2005).

Many of these mutations have been observed in the host eukaryotic translation initiation (eIF) complexes, and occur naturally (Diaz-Pendon, et al., 2004; Robaglia and Caranta, 2006). The importance of the eIF complex has further been demonstrated in Arabidopsis, where knock outs of these genes results in resistance to several viruses including *Turnip yellow vein virus*, *Turnip mosaic virus* and *Cucumber mosaic virus*. The eukaryotic initiation factor 4E (eIF4E) has been shown to play a major role in resistance to potyvirus infection (Gao et al., 2004; Piron et al., 2010; Ruffel et al., 2005). Plant cells have an isoform of the eIF4E gene known as eIF(iso)4E, and this gene has also been implicated as necessary for infection with other potyviruses in Arabidopsis. These viruses are specialised to utilise one of the two eIF complexes, but recent evidence has suggested that some potyviruses are capable of using either complex, or require the use of both (Ruffel et al., 2006; Sato et al., 2005) in order to initiate translation of their proteins. Some viruses have been shown to require the use of other proteins within the eIF complex, such as the scaffolding protein eIF(iso)4G (Albar et al., 2006).

The host cell eIF complex is formed with a cap-binding pocket, located within the eIF4E protein. The cap-binding pocket associates with the m^7G 5' cap structure of host cellular mRNAs in order to initiate translation. In all cases of potyvirus infection requiring eIF4E, the resistant form of the gene has been shown to contain mutations in the surface loop in close proximity to the cap-binding domain of eIF4E. In most cases the mutations consist of an amino acid substitution (Nieto et al., 2006). The *Avr* for this type of infection is

therefore assumed to be the VpG (Robaglia and Caranta, 2006). In *A. thaliana* the VpG is able to substitute for the host mRNA 5' cap structure in order to initiate translation (Maule et al., 2007).

Mutations in the EIF complex are not guaranteed to produce a resistant phenotype, as methods of cap-independent translation have been observed in several plant viruses. This method, instead of requiring a VpG, utilises the 5' untranslated region (UTR) interacting with eIF4G (and its isoforms) protein of the EIF complex to initiate translation (Gallie, 2001).

Other important recessive resistance genes encode membrane proteins including fatty acid conversion proteins, and transmembrane proteins in Arabidopsis and tomato plants (Tsujimoto et al., 2003; Yamanaka et al., 2000). It is thought that these proteins are important for viral localisation within the plant cell (Hagiwara et al., 2003).

A greater understanding of these recessive resistance genes will allow a clearer understanding of virus activity within the cell, and will reveal more gene targets that can be utilised by plant breeders to create durable genetic resistance in crop plants.

1.3.3. Post transcriptional gene silencing

As previously discussed, post-transcriptional gene silencing (PTGS) is an innate mechanism in plant cells, specifically to destroy foreign dsRNA. This directed mechanism means that many RNA viruses can be targeted for destruction, as even ssRNA viruses become double stranded during the replication of their genomes. This mechanism has the capacity to spread systemically throughout the plant, increasing whole plant immunity to the infecting virus (Maule et. al., 2007).

This process involving RNA silencing allows plants to recover from virus infection. Several important components are required for successful PTGS. These include the DICER-like (DICL1-4) enzymes that cleave dsRNA into

small RNA fragments and the RNA-induced silencing complex (RISC). The RISC complex uses the small fragments as a guide to degrade other homologous sequences throughout the plant (Maule et al., 2007; Ritzenthaler, 2005).

Viruses have however developed a resistance mechanism against RISC and PTGS. The 5' proximal ORF-0 of poleroviruses encodes a protein that interferes with PTGS. The expression of the P0 protein does not inhibit the expression of RISC proteins, but instead reduces their activity and function. The viral P0 protein contains an F-box like domain that has been shown to target the plant encoded PAZ motif of the ARGONAUTE protein (AGO1) in Arabidopsis (Baumberger et al., 2007). AGO1 is part of the RISC complex, and becomes destabilised by P0, reducing the activity of RISC (Bortolamiol et al., 2007). Importantly, the P0 protein is only able to destabilise AGO1 before it forms part of the RISC complex. Once the complex has been formed, P0 is not able to affect its function (Csorba et al., 2010). It achieves this by degrading AGO1, as the F-box like structure are commonly associated with E3 ubiquitin ligase complexes that marks proteins for destruction by the proteasome, although P0 must be using a different system as during proteasome inhibition the degradation of AGO1 still occurrs (Baumberger et al., 2007).

1.4 Aims and Objectives

This research project has been funded by a Norwich Research Park studentship, with Brooms Barn Applied Crop Sciences acting as CASE partners. The aim was to increase our understanding of poleroviruses infecting the common East Anglian crop plants, sugar beet and oilseed rape. The aims of the project were as follows;

Using a natural resistance screening method, identify durable recessive resistance genes to infection with BMYV and TuYV in *A. thaliana.* It is already known that *A. thaliana* Col-0 and Ler are susceptible to infection with both poleroviruses, and a screen of natural resistance of *A. thaliana* ecotypes from around the UK has identified one BMYV resistant ecotype, Sna-1. Genetic characterisation of this plant may illustrate possible recessive resistance gene targets.

The development of an infectious *Turnip yellows virus* (TuYV) clone, capable of being delivered by *A. tumefaciens* into plants in order to study infection. This has previously been carried using several poleroviruses, most recently by Percival-Alwyn (2010), with the virus *Beet mild yellowing virus* (BMYV). The use of this infectious clone would mean that aphid inoculations would no longer be required, and the study of virus infection mechanisms would become easier to perform.

Chapter 2 Materials and Methods

This chapter describes the experimental procedures used throughout this investigation. This includes the microorganisms, plasmids, buffers and media that were required, as well as plants and viruses used, and the methods of analysis.

2.1 Growth and maintenance of Bacterial strains

This study required the use of *Escherichia coli* and *Agrobacterium tumefaciens* bacterial strains for different cloning and infection experiments. Chemically competent *E. coli* strains were used for transformation with various recombinant plasmids. Electro-competent *A. tumefaciens* strains were transformed with recombinant plasmids, and in turn were transformed into *Arabidopsis thaliana* for stable transformation, or transient expression for delivery of recombinant virus

2.1.1 *E. coli* Bacterial Strains

Table 2.1 Strains of organisms used throughout this study

Strain	Genotype	Reference and use
DH5α™	Fφ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A *end*A1 *hsd*R19 (r_k-, m_k-) *pho*A *sup*E44 λ-*thi*-1 *gyr*A96 *rel*A1	Cloning of *TuYV* cDNA
XL2-Blue Ultracompetent Cells	*endA1 supE44 thi-1 hsdR17 recA1 gyrA96 relA1 lac* [F´ *proAB lacIqZΔM15 Tn10 (Tetr) Amy Camr].*	Stratagene, Agilent technologies. Used for cloning of *TuYV* cDNA and Col-O genes

2.1.1.1 Growth of *E. coli* cultures

E. coli were grown in liquid Luria and Bertani medium (LB medium), 1% w/v sodium chloride, 1% w/v tryptone, 0.5% w/v yeast extract (pH adjusted to 7.0 with sodium hydroxide), together with the relevant antibiotics. Liquid cultures were grown overnight in a shaking incubator at 250 rpm, 37 ºC. LB agar medium (LB medium, supplemented with the addition of 1.5% w/v micro agar) cultures were grown overnight at 37 ºC.

2.1.1.2 *E. coli* Transformation by Heat Shock

Transformations were performed using chemically competent *E. coli* strains. Plasmid DNA (50 ng) was added to 25-50µl of cells, and incubated on ice for 30 minutes. A negative control was performed using filtered and distilled water in place of plasmid DNA, to ensure that the cells were unable to survive selective antibiotics without being transformed. Samples were then incubated for 30 seconds at 42 ºC as a heat shock. Cells were placed on ice and 250 µl of LB medium was added. Cells were allowed to recover by incubating for 1 hour at 37 ºC (250 rpm). Cultures were then centrifuged at 4,000 x *g* for 3 minutes, the pellet re-suspended in 100 µl LB and plated onto LB agar plates, containing appropriate the appropriate antibiotics. The plates were grown overnight at 37 ºC.

2.1.1.3 *E.coli* Plasmid Isolation

Plasmids were purified from bacteria grown on LB agar plates, containing selective antibiotics. A single colony was picked and inoculated into 10 ml LB, containing the selective antibiotic. These were grown at 37 °C, 180 rpm, for 18 hours. Plasmids were extracted from the cultures following the Promega (Madison WI, USA) Wizard SV+ plasmid mini-preparation kit. The plasmids were eluted in filtered and sterilised distilled water, and analysed on a 1% (w/v) agarose gel.

2.1.2 *A. tumefaciens* Strains

A. tumefaciens strain AGL1 (Lazo et. al, 1991) containing the defective tumour-inducing plasmid pTiBo542 (Lazo et. al, 1991) was used throughout the study for cloning and plant transformation. The strain contains resistance to Rifampicin.

2.1.2.1 Growth of *A. tumefaciens* Cultures

A. tumefaciens was grown on modified LB medium, containing 0.5% w/v sodium chloride, 0.5% w/v yeast extract, 1% w/v tryptone (pH adjusted to 7.0 with sodium hydroxide), and the appropriate antibiotic. Cultures were grown at 28 °C for up to 48 hours on modified LB agar (as modified LB with 1.5% w/v micro agar added), and at 28 °C, 250 rpm in liquid modified LB.

2.1.2.2 *A. tumefaciens* AGL1 Electrically Competent Cells

A single *A. tumefaciens* AGLI colony was used to inoculate 10 ml of modified LB liquid medium, and grown in a shaking incubator at 28 °C, 250 rpm, overnight, containing 50 $mg.L^{-1}$ Rifampicin. The following day, 5 ml of this culture was used to inoculate 50 ml LB medium containing 50 $mg.L^{-1}$ Rifampicin, and incubated at 28 °C, 250 rpm, until the OD_{600} reached 0.6-0.7. The culture was then incubated on ice for 15 minutes, and centrifuged for 15 minutes at 2,000 x *g* at 4 °C. The pellet was then re-suspended in 35 ml of ice cold sterile water, and then centrifuged again for 15 minutes at 2,000 x *g* at 4 °C. The pellet was once more re-suspended in 35 ml of ice cold sterile water, and centrifuged for 15 minutes at 2,000 x *g*, 4 °C. The pellet was then re-suspended in 1 ml of ice cold 10% v/v glycerol, and divided into 50 µl aliquots. These were snap frozen in liquid nitrogen and stored at -80 °C.

2.1.1.2.3 *A. tumefaciens* Electro-Transformation

A. tumefaciens was transformed using electroporation. Aliquots of cells, 20 µl, were thawed and 1 µl of the plasmid preparation added. The mix was placed

in-between two electrodes of a pre-chilled 0.15 cm disposable electroporation chamber (BioRAD, Munich, Germany). The DNA was transformed into the cells using 330 µF Capacitance, 4000 Ω Resistance, and 400 V. Impendence was set to “low Ω” and Change Rate “fast”. Cells were then incubated at 28 °C 250 rpm, in 200 µl of modified LB for 2 hours, before being plated onto solid modified LB containing appropriate antibiotics and incubated at 28 °C for 48 hours.

2.1.2.4 *A. tumefaciens* Plasmid Isolation

Plasmids were purified from *A. tumefaciens* after being grown on modified LB agar plates, containing selective antibiotics. A single colony was picked and inoculated into 5 ml LB, also containing selective antibiotics. These were grown at 28 °C, 250 rpm, for 18 hours. Plasmids were extracted from the cultures following the Promega (Madison WI, USA) Wizard SV+ plasmid mini-preparation kit, with suggested adjustments from the manufacturer as follows; after re-suspending the cells, 100 µl of lysozyme solution (10 $mg.ml^{-1}$ in 10 mM Tris:HCl, pH 8.0) was added, and incubated for 5 minutes at room temperature. The plasmids were eluted in filtered and sterilised distilled water, and analysed on a 1% (w/v) agarose gel.

2.1.2.5 Selective Minimal Media

A. tumefaciens colonies were selected on minimal media in order to gain prototrophic colonies. The minimal medium used contains 21 $g.L^{-1}$ K_2HPO_4, 9 $g.L^{-1}$ KH_2PO_4, 2 $g.L^{-1}$ $(NH_4)_2SO_4$, and 1 $g.L^{-1}$ $Na_3C_6H_5O_7.2H_2O$, in distilled water. To this solution, 2 ml of filter sterilised 1M $MgSO_4$ and 20 ml 20 % w/v glucose was added. The media was heated to 55 °C and micro agar added to a final concentration of 1.5 % w/v.

2.1.3 Storage of Bacterial Cultures

Bacterial cultures were stored long term at -80 °C, by using 1.5 ml aliquots of bacterial culture, and adding 0.5 ml sterile 60% v/v glycerol.

2.1.4 Antibiotic Stocks

Table 2.2 Antibiotic concentrations. The stock solutions and the final concentrations of antibiotics used in this study.

Antibiotic	Solvent	Stock Solution ($mg.ml^{-1}$)	Final Concentration ($\mu g.ml^{-1}$)
Ampicillin	Water	100	100
Kanamycin	Water	50	50
Rifampicin	Methanol	20	50

2.2 Growth and Maintenance of *Arabidopsis thaliana*

The model organism *A. thaliana* was used in a variety of studies throughout this research. This section will describe methods used.

2.2.1 Growth of *A. thaliana* in compost

Around 50 seeds were sown into a 51 mm diameter plastic pot, containing compost. Seeds were vernalised for two nights in darkness at 4 °C in order to encourage the seeds to start germinating at the same time. Seeds were placed under a plastic propagator lid, and moved to short day rooms (22 °C, 8 hours light). When seedlings were approximately 1 cm in height they were transferred to individual pots, or individual compartments in 4 x 5 seed trays, containing compost. Plants grown for infection remained in the short day growth room for a further 5 weeks, whilst plants grown for seed and crossing were transferred into long day growth rooms (22 °C, 16 hours light).

2.2.2 Growth of *A. thaliana* on media

Seeds collected after flower dip analysis were sterilised, plated, and grown on plant medium. Sterilisation of seeds used the gas sterilisation method as follows. Seeds were sterilised in wax bags, 1g of seeds in each bag. Three ml of Hydrochloric acid was added to 100 ml of bleach in a glass dish. A metal grid was placed over the dish, seeds bags placed on the grid, covered with a large glass desiccator, and left overnight. Seeds were then removed from the desiccator and left to dry for 1 hour. Plant growth medium (4.3 $g.L^{-1}$ Murashige and Skoog Salts, 0.5 $g.L^{-1}$ MES (2-(n-morpholino)-ethanesulfonic acid), 0.1 $g.L^{-1}$ Myo-inositol, 1 $ml.L^{-1}$ G.M. vitamins, 8 $g.L^{-1}$ Bacto Agar, pH to 5.7 with KOH) was melted and 1 $ml.L^{-1}$ of the fungicide Nystatin (25 $mg.ml^{-1}$) added, along with Kanamycin (50 $mg.L^{-1}$) for selection.

2.2.3 Crossing of *A. thaliana* Ecotypes

Crosses were performed between different *A. thaliana* ecotypes under controlled conditions. Any siliques, and open buds were removed from the bud cluster until 4 immature flower buds remained, whilst any mature flowers were saved. The sepals, petals and anthers were removed from the immature flower bud leaving only the stigma. Mature flowers from another plant were then used to dab pollen from the anthers onto the exposed stigma. The pollinated stigma was then covered gently with Clingfilm, to maintain humidity, and labelled. After 3 days, the cling film was removed, the siliques left to develop, and seed collected when silques had dried.

2.2.4 *A. thaliana* DNA extraction

DNA was extracted from plant leaves using the Qiagen DNeasy® Plant Mini Kit. The emulsification of the tissue was carried out by snap freezing in liquid nitrogen, and grinding the leaf in liquid nitrogen using a blue plastic homogeniser. DNA was eluted in 100 µl of nuclease free water.

2.2.5 *A. thaliana* RNA extraction

Viral and plant RNA was extracted from leaf tissue using the Qiagen RNeasy® Mini Kit as described by the manufacturers instructions. The RNA clean up steps were also followed, and RNA was eluted into 30 µl of nuclease free water.

2.2.6 Amplified fragment length polymorphism (AFLP^{TM})

The DNA fingerprinting analysis tool AFLP^{TM} was used in a bulked segregant analysis of two different *A. thaliana* ecotypes. AFLP^{TM} allowed identification of DNA markers linked to a gene for BMYV (*Beet mild yellowing virus*) susceptibility. The F_2 population analysed was from a Col-O x Sna-1 cross, and individual progeny were analysed for resistance/susceptibility using TAS-ELISA (Percival-Alwyn, 2010). The 20 most infected plants were bulked, and the 20 least infected plants were bulked, and DNA extracted from leaf material for each bulk. DNA extraction was carried out by leaf samples being snap frozen in liquid nitrogen and ground using a pestle and mortar. To this, 7 ml of Extraction buffer (10mM Tris-HCl pH 8.0, 0.5 M EDTA pH 8.0, 5 M NaCl, 1.5% w/v SDS, 0.1% v/v β-mercaptoethanol) was added. Extracts were incubated at 65 ˚C for 12 minutes with occasional gentle agitation. Exactly 200 µl of 5 M KAc was then added and incubated on ice for 10 minutes. The two extracts were centrifuged at 8000 *x* g for 8 minutes at 5 ˚C and the supernatant transferred into clean tubes. From this point, the samples were phenol purified and ethanol extracted as described in section 2.4.9. DNA samples (usually 1 µg) were then digested with the restriction enzymes *Pst*I and *Mse*I. These enzymes were chosen because *Mse*I is a frequent cutter of DNA, and *Pst*I is a rare cutter. *Pst*I and *Mse*I adaptor molecules were ligated to the DNA using T4 DNA ligase (Thomas et al., 1995). PCR (polymerase chain reaction) was carried out with a range of primers with partial specificity to the adaptor molecules. Primers to *Mse*I contained 3 differing selective nucleotides after the complimentary adaptor sequence. *Pst*I primers contain 2 selective nucleotides and were radiolabeled with ^{33}P-γ-ATP. Twenty two different combinations of primers were used (primers differed in their selective

nucleotide sequence), and PCR products were run on a denaturing 4.5 % w/v polyacrylamide gel, as described by Thomas et al., (1995). Gels were dried and autoradiograpy carried out to visualise the PCR products that were amplified to include the radiolabelled *Pst*I primer.

2.2.7 Transformation of *A. thaliana* with *A. tumefaciens*

A. tumefaciens strain AglI (see section 2.1.2.2) was used to introduce DNA into *A. thaliana* genome using the floral dip method (Clough and Bent, 1998). *A. tumefaciens* was grown in 10 ml liquid LB medium containing 50 $mg.L^{-1}$ Kanamycin, at 28 ˚C overnight. This culture was used to inoculate 200 ml of LB medium, also containing Kanamycin, as well as 150 µM acetosyringone, a chemical known to be present in plant wounds that acts to attract *A. tumefaciens.* This was incubated for a further 24 hours at 28 ˚C. Cultures are then centrifuged at around 3,000 x *g* for 15 minutes. The pellet is re-suspended in an equal volume of infiltration media (2.164 $g.L^{-1}$ ½ MS (Murashige and Skoog) medium, 5% w/v sucrose, 3mM MES (2-(N-Morpholino)ethanesulfonic acid, 4-Morpholineethanesulphonic acid), 500 $µl.L^{-1}$ Silwet L-77, 300 mM Acetosyringone, pH 5.5). Plant bud ends that were about to flower, were dipped in the suspension for around 2 minutes with gentle agitation, and then left to recover at high humidity in a shaded area of a greenhouse for 24 hours. Plants were then grown in long day rooms (16 hours light under UV bulbs, 22 ˚C) and seed collected. Successful transformation was determined by the growth of seeds on selective medium as described in section 2.2.2.

2.3 *Myzus persicae* Infection Experiments

Myzus persicae, the green peach aphid, was used in infection experiments to transfer viruses from confirmed infected plants, to uninfected plants. The TuYV isolate used throughout the studies was UK-BB TuYV.

2.3.1 *M. persicae* cultures

M. persicae aphids were grown in Perspex cages at 20-22 ˚C, under continuous light. Their food source was *Brassica pekinensis*. Cultures were then sub-cultured to feed on infected material. Aphids to become BMYV infected were fed on BMYV infected sugar beet, and aphids to be infected with TuYV were fed on infected oil seed rape for at least 48 hours before use. Infection was confirmed by TAS-ELISA.

2.3.2 Inoculations of *A. thaliana* with *M. persicae*

Viruliferous *M. persicae*, was used to inoculate *A. thaliana,* grown in short day rooms (8 hours light, 22 ˚C). At least 10 Aphid nymphs were placed on each *A. thaliana* plant to increase the probability of virus transfer. Aphids were left for 1 week to feed on *A. thaliana* before being removed by the insecticide Admire® Pro Systematic Protectant (Bayer CropScience Ltd.). Plants were then grown for 6-8 weeks to allow systemic virus infection, before the leaf material was tested by TAS-ELISA.

2.3.3 TAS-ELISA

Triple antibody sandwich enzyme linked immunosorbent assay (TAS-ELISA) was carried out on plant sap, from leaf material infected with a virus strain, to determine whether the plant was susceptible, or resistant. Recipes for buffers used in TAS-ELISA can be found in Table 2.3. Ninety-six well plates were coated in polyclonal immunoglobulin G (polyclonal IgG) (1:1000), and mixed with coating buffer. The plate was incubated for 1 hour at 37 °C. The plate was washed three times for 3 minutes in 1 x PBS. Each well being used then had 200 μl blocking buffer added and was incubated at room temperature for 1 hour. Leaf material was chosen from the midrib of the plant, where there were typical infection signs in the leaf, such as thickening and reddening. The total leaf material from each plant weighed 0.2 g, and was taken 4-6 weeks after inoculation with virus. Each leaf was diluted 10x its weight, and so had

an appropriate volume of extraction buffer added (usually 1.8 ml). The leaf was crushed using a pestle and mortar, with the extraction buffer, releasing the sap from the leaf. The sap and extraction buffer mix (100 μl) was added to an individual well of the 96 well plate. Healthy (uninfected material) and confirmed infected controls were also carried out in the same way. The plates were covered and stored at 4 °C overnight. The following morning, plates were emptied and washed four times for 3 minutes, in 1 x PBS. Monoclonal antibody (MAF-24) was then added as a 1:1000 dilution with extract buffer, and 100 μl of this mix was added to each well. The plates were covered and incubated for 2 hours at 37 °C. Plates were once again washed four times for 3 minutes in 1 x PBS. Anti-mouse antibody (Sigma, Antimouse IgG) was then added in a 1:1000 dilution with extract buffer, and 100 μl added to each well, covered, and incubated at 37 °C for 2 hours. The plates were then washed four times for 3 minutes in 1 x PBS. A 5 mg substrate tablet (Sigma phosphate substrate) was dissolved in 10 ml substrate buffer and 100 μl added to each well. The phosphate substrate is 4-Nitrophenyl phosphate disodium salt hexahydrate, and is converted to 4 nitrophenol, and phosphate by alkaline phosphatase, which is a distinct yellow colour. The plate was left for 1 hour at room temperature for the yellow colour to develop. The plates were then placed in a plate reader (Athos 2001) and measured at 405 nm.

Table 2.3 Buffers required in TAS-ELISA as described in section 2.3.4

Buffer Name	Constituents
Coating buffer (pH 9.6)	1.59 g Na_2CO_3, 2.93 g $NaHCO_3$ dissolved in 1 L distilled water
Substrate buffer (pH 9.8)	97 ml diethanolamine, water to a total volume of 1 L
10 x PBS	480 g NaCl, 12 g KH_2PO_4, 12 g KCl, 174 g Na_2 HPO_4 $12H_2O$
Washing buffer (pH 7.4)	1 L 1 x PBS, 5 ml Tween 20
Blocking buffer	100 ml 10 x PBS, 1 g Milk Powder, 900 ml distilled water
Extract buffer	100 ml blocking buffer, 0.5 ml 10% tween 20.

2.4 Genetic Analysis Methods

The following methods were used to analyse DNA and RNA from plants, viruses and bacteria.

2.4.1 Polymerase Chain Reaction (PCR)

PCR was used as an analytical tool throughout the study. PCR was required for identification of different plant ecotypes during CAPS analysis, creation of cDNA libraries, verification of T-DNA and plasmid inserts, in overlap PCR and ligation independent PCR.

2.4.1.1 Oligonucleotide Primers

The software *Gene Runner* Version 3.01, (Hastings Software Inc., NY, USA www.generunner.com) was used to design primers. Wherever possible, primers were designed to have specific characteristics as follows; i) GC content of each primer to be at least 50 %, ii) melting temperature (Tm) is to be between 55-60 °C, iii) no loops or hairpins will form during temperatures

experienced during PCR. The *Gene Runner* software was used to analyse these characteristics, and primers were ordered from MWG Operon, Eurofins (Ebersberg, Germany). Primer stock solution was made to 100 µM, and working concentrations of 20 µM. A full list of primers used in this study will be found in the relevant results chapter. The lab stock numbers for each primer are given in brackets.

2.4.1.2 GoTaq®Flexi PCR

The Promega enzyme GoTaq®Flexi DNA Polymerase was used for all initial PCR to test primers, and also for the molecular analysis of plant ecotypes. Reactions were set up in accordance with the manufacturers recommendations.

2.4.1.3 Phusion® High Fidelity DNA Polymerase

Phusion® (Thermo Fisher Scientific) was used for cloning target sequences because of its high fidelity, as the Phusion® polymerase enzyme has an error rate >50-fold lower than that of *Taq* DNA polymerase. PCRs were set up according to the manufacturer's recommendations, and usually with a 30-cycle repeat.

2.4.2 Overlap PCR

Overlap PCR was used to join two fragments of DNA in equimolar concentrations. The polymerase enzyme used for this reaction was Phusion® (Thermo Fisher Scientific) due to its high fidelity. The reaction was set up using 0.2 mM dNTPs (Bioline), 1x HF PCR buffer (Thermo Fisher Scientific), 20 units.ml^{-1} Phusion® enzyme, and equimolar amounts of two different DNAs (determined by analytical gel electrophoresis) containing a complimentary overlap region of around 20 nucleotides. The cycling conditions typically used were 98 °C for 30 seconds, followed by 98 °C for 10 seconds, 55 °C for 30

seconds, 72 °C for 30 seconds Kb of sequence to be amplified, repeated twenty times, followed by a final step of 72 °C for 5 minutes.

2.4.3 cDNA Synthesis RT (reverse transcriptase) PCR

First strand synthesis of cDNA was carried out on RNA extracted from infected plant leaves to amplify viral RNA contained within the leaf extraction sample. The enzyme used in this reaction was Superscript® II Reverse Transcriptase (Invitrogen by Life Technologies). The protocol was followed as described by the manufacturer, using specifically designed reverse primers to the virus, or plant gene.

2.4.4 In-Fusion® HD Cloning

Cloning of some PCR products used the In-Fusion® HD Cloning Kit Clontech Laboratories, Inc. Primers were designed and reactions set up following the manufacturer's instructions in the handbook (October 2011) VI. Protocol I: In-Fusion Cloning Procedure w/Spin-Column Purification.

2.4.5 DNA Restriction Digestion

Restriction digests were carried out on purified plasmids and PCR products in order to confirm their identity, and on PCR fragments for cloning. Restriction digests were set up according to manufacturers instructions (Roche Diagnostics, New England Biolabs, and Invitrogen).

2.4.6 DNA Ligation

DNA ligations were carried out using T4 DNA Ligase (Invitrogen by Life Technologies). Ligations were set up according to manufacturers protocol. This typically consisted of a 3:1 molar ration of insert to vector. In each reaction no more than 5 Units.ml^{-1} of T4 DNA ligase were used, and reactions were incubated at room temperature for at least 17 hours.

2.4.7 Agarose Gel Electrophoresis

DNA (plasmid, genomic and PCR product) was visualised using 1% (w/v) agarose gels. Each 50 ml gel contained 50 ml of 1x TAE buffer (40 mM Tris acetate, 1 mM EDTA, pH to 8.0), and 0.5 μg/ml Ethidium Bromide. Loading buffer (20 % w/v Ficoll 400, 1 % w/v Orange G, 5mM EDTA) was added to DNA samples. Gels were run typically at 100 V and photographed using a UV transilluminator (BioRAD, Munich, Germany).

2.4.8 DNA Extraction from Agarose Gel

DNA fragments were extracted from agarose gel after digestion to remove any unwanted DNA, using QIAquick® Gel Extraction Kit (Qiagen). The manufacturers protocol was followed, and DNA eluted into 30 µl of nuclease free water.

2.4.9 Ethanol Precipitation of DNA

DNA that required either concentration, or purification following restriction enzyme digestion, was phenol extracted and ethanol precipitated. An equal volume of Phenol:Chloroform;Isoamyl alcohol was added to eluted DNA, and thoroughly mixed. The mix was then centrifuged at 16,000 x *g*, for 8 minutes. The upper aqueous phase was carefully removed and 1/20th volume of 3M NaAc (pH4.8) was added, and 2.5 x total volume of ice cold 100 % ethanol. After incubation at -20 °C for at least 2 hours the mix was centrifuged at 16,000 x *g* for 15 minutes and the supernatant was discarded. The pelleted DNA was washed with 1 ml of ice cold 100 % ethanol, centrifuged again at 16,000 x *g* for 1 minute, and the supernatant discarded. The tube was then vacuum centrifuged at 16,000 x g for 5 minutes, or until any remaining liquid was removed. The pellet of DNA was then re-suspended in 20 µl nuclease free water.

2.4.10 DNA sequencing

Template DNA and appropriate primers were sent to *The Genome Analysis Centre* (TGAC), at the Norwich Research Park for sequencing. Data files received from TGAC were assembled using the BioEdit© Biological Sequence Alignment Editor version 7.0.5.3 (Ibis Bioscience, Carlsbad, CA, USA). Sequences were compared using the local alignment tool EMBOSS (European Molecular Biology Open Software Suite) *Water*, through the EMBL-EBI (European molecular biology lab - European bioinformatics institute) website (Cambridge, UK). Both nucleic acid sequences, and predicted amino acid sequences were compared using this software.

Chapter 3 Identification of a *Beet mild yellowing virus* (BMYV) resistance gene by exploiting natural variation in *A. thaliana*

3.1 Introduction

Arabidopsis thaliana, sometimes known as mouse ear cress, is a member of the Cruciferae, and is a flowering plant native to Europe and Asia. *A. thaliana* is used commonly as a model organism in the study of genetics, evolution and plant development. Characteristics such as it's diploid genetics, a relatively small genome and short growth cycle make it an ideal model organism for genetic studies. *A. thaliana* is an autogamous species, which means that many wild plants are inbred, and pure breeding for many genetic variants. These are usually referred to as ecotypes, meaning a distinct type of plant that has adapted to its surroundings, although in the literature its is becoming increasingly common to refer to these ecotypes as accessions (Koornneef et al., 2004).

A. thaliana first came to prominence in 1907 with the publication by Friedrich Laibach that identified 5 homologous pairs of chromosomes. Laibach later championed *A. thaliana* as a model organism in 1943. He noticed natural variation in the phenotypes of *A. thaliana* ecotypes and assisted in the creation of the first mutants of *A. thaliana* with the use of X-rays (Koornneef and Meinke, 2010). In 1975 *A. thaliana* re-emerged as a candidate for a plant model organism when Rédei published a review outlining it's suitability for mutation, evolution, and physiological studies amongst others (Rédei, 1975). It was in 1985 that *A. thaliana* finally was described as a model organism for plant genetics, and the following year it was transformed using *Agrobacterium tumefaciens*. Sequencing of the genome was initiated in 1990, with individual research groups from across Europe, Japan, and the United States coming together to publish the complete *A. thaliana* genome in 2000 (AGI, 2000). Following the publication of the genome sequence, a project to understand the

function of *A. thaliana* genes was initiated (Chory et al., 2000). The Arabidopsis Information Resource (TAIR) now holds all annotated sequence information currently known about Arabidopsis genes. As well as sequence and annotation information, TAIR also integrates the Arabidopsis Biological Resource Center (ABRC), and the European Arabidopsis stock centre (NASC) seed stock databases, allowing users to link genes of interest to available seed stocks (Swarbreck et al., 2008). This highly integrated, easily accessible genome information, along with the vast availability of ecotypes and mutants available make it an ideal molecular and genetic model for studying cellular processes in all plants, including crop plants.

Natural variation in plants has long been of interest to evolutionary biologists, geneticists and crop breeders. These traits are investigated to understand how plants to adapt to their surroundings to maximise their chances for growth and reproduction (Trontin, et al., 2011). *A. thaliana* lends itself particularly well for studying natural variation because of the range of environments it is found in, showcasing its potential genetic variation within the species.

There are two main methods for studying natural variation in *A. thaliana*, i) investigating different accessions, and ii) by using recombinant inbred lines (RIL). Using the first method, *A. thaliana* accessions are used to identify polymorphisms directly linked to the adaptation being studied. This often requires the use of linkage disequilibrium mapping, which shows the relationship between two or more polymorphisms to their phenotypes, and leads to the identification of potential candidate genes. The second type, RIL, allows genetic identification of phenotypic traits by crossing parent plants to gain homozygous plants at various alleles (Trontin et al., 2011).

There are just over 750 naturally occurring ecotypes of *A. thaliana* from around the world, stored by the seed stock centres ABRC and NASC (TAIR, 2013). The availability of a wide range of ecotypes, which grow in different geographical locations, under a variety of environmental pressures will provide a source of genetic diversity. This might also provide a range of degrees of resistance to potential pathogens that have be tested for, and can continue to

be used as a tool enabling the identification of important host resistance factors.

3.2 UK Accession Screening and Analysis of Segregating Population

Genetic screening of *A. thaliana* was carried out using 80 ecotypes from around the UK. Each individual geographical isolate was collected by Professor Eric Holub and was previously described by Percival-Alwyn (2010). The ecotypes tested previously showed a range of resistance to *Hyaloperonospora arabidopsis* (Hall et al., 2009). Screening of infected plants to determine the presence of viruses historically has used visual assessment, as well as transmission studies. Techniques now commonly use triple antibody sandwich enzyme-linked immunosorbent assay (TAS-ELISA) (D'Arcy, et al., 1989).

TAS-ELISA requires the use of monoclonal antibodies with the ability to discern between viruses such as BMYV and BChV. The viruses show a high level of sequence homology at the 3' end, but a the 5' end can show as little as 30% homology, (Stevens, et al., 2004) allowing two monoclonal antibodies to be used in order to identify viruses infecting plants. The antibody MAF-24 identifies both BMYV and BChV (Smith et al., 1996) whilst the antibody BYDV-PAV-IL-1 has the ability to distinguish between the two viruses by only associating with BMYV (D'Arcy et.al., 1989). MAF-24 was used throughout the infection studies in order to identify the presence of these viruses. There are, however, limitations to TAS-ELISA in this study. *A. thaliana* leaves that were sampled varied in size and developmental stage. Also, due to their small size, a core borer was unable to be used for leaf sampling. Another problem is that as poleroviruses are phloem limited viruses, the titre of virus is often low in *A. thaliana*, and virus particles are unstable meaning that it is not possible to compare results between plates, and especially between different screenings.
To investigate resistance to BMYV and *Beet chlorosis virus* (BChV) in *A. thaliana* ecotypes, Percival-Alwyn (2010) infected plants and screened for

resistance using TAS-ELISA. Infections were carried out using *M. persicae* as described in methods section 2.3, using the ecotypes Columbia and Landsberg as BMYV susceptible and BChV resistant controls respectively.

3.2.1 Ecotype Sna-1

Previous screenings carried out on the 80 UK ecotypes found one ecotype (Sna-1) gave consistently low TAS-ELISA readings, similar to uninfected plants suggesting that low levels of virus were present in leaf tissue. It was also found that all ecotypes infected with BChV contained high levels of virus particles, comparable to infected control plants. It was established that BMYV resistance in Sna-1 plants was not due to aphid resistance as this ecotype is still susceptible to *Turnip yellows virus* (TuYV), which is also aphid-transmitted (Percival-Alwyn, 2010).

To investigate if resistance was caused by a passive or active resistance mechanism, crosses were performed between Sna-1 and Col-0. The F_1 progeny of these crosses was infected with BMYV in order to establish if the Sna-1 resistance is a dominant or recessive trait. If the trait is dominant the resistance mechanism is likely to be an active response, and if recessive, likely to be a passive response. The F_1 progeny were found to give TAS-ELISA readings higher than that of the Sna-1, but lower than the Col-0 controls. This could indicate that susceptibility to BMYV in Sna-1 plants was influenced by a single gene, but this would need to be confirmed by genetic analysis of the F_2 generation.

To further test this idea F_2 progeny of the Sna-1 x Col-0 crosses were screened. Percival-Alwyn (2010) proposed that if a single gene was the cause of resistance to BMYV a ratio of 1:2:1 for highly resistant:intermediate susceptibility:highly susceptible infected plants would be seen. A x^2 analysis of the results was performed on the TAS-ELISA results that confirmed the1:2:1 segregation ratio. It was concluded that resistance to BMYV in Sna-1 was caused by a single recessive resistance gene. The project discussed in this

thesis began with the analysis of the genomes of selected Sna-1 x Col-0 crosses using Amplified Restriction Fragment Polymorphism (AFLP ^{TM}).

3.3 Mapping the chromosomal region of the BMYV resistance gene through Amplified Restriction Fragment Polymorphism^{TM} (AFLP^{TM}) analysis.

Bulked segregant analysis is a technique that allows the screening of two pools of DNA for differences at multiple genomic locations. A cross between two genetic variants (different *A. thaliana* ecotypes) is used to generate F_2 segregants of the two classes (either BMYV resistant or susceptible). Individual F_2s in each segregant class are bulked together to create two pools from which DNA is prepared for molecular analysis. In this strategy, as result of independent assortment, the genome-wide molecular variation in the two parents is equally represented in both pools, and the two pools will only differ with respect to markers that are linked to the selected trait (Michelmore et al., 1991).

The linked markers can be identified using a variety of different DNA fingerprinting strategies, including AFLP^{TM} (Thomas et al., 1995). AFLP^{TM} enables the construction of high density linkage maps. This approach easily allows the analysis of large amounts of DNA, whilst still being able to detect small differences in genome sequence (Thomas et al., 1995). By carrying out the DNA fingerprinting technique AFLP^{TM} on both pools, markers can be identified and ultimately mapped to a genomic region of *A. thaliana* linked to the trait of interest (Thomas et al., 1995).

Two pools of *A. thaliana* were produced using the Sna-1 x Col-0 F_2 population created by Percival-Alwyn (2010). The 20 most highly infected F_2 plants were collected into the susceptible '*S*' bulk, and the 20 least susceptible were collected into the resistant *'R'* bulk. If the single susceptibility gene hypothesis is correct, then the resistant pool will be homozygous recessive for the Sna-1 resistance gene, whilst the susceptible pool will be heterozygous, containing

one dominant, and one recessive copy of the gene. Genomic DNA from the two pools was extracted as described in Section 2.2.4 and used for AFLP™ analysis as described in Section 2.2.6. The AFLP™ analysis was carried out using 24 *Mse* primers and one *Pst* primer, seen in Table 3.1, the results can be seen in Figure 3.1. It is expected that much of the DNA seen in the AFLP profile will be identical for the resistant and the susceptible pools, as the bulked populations are segregating for suspected single gene, therefore any differences in the DNA pattern may be due to a susceptibility gene in the susceptible pool.

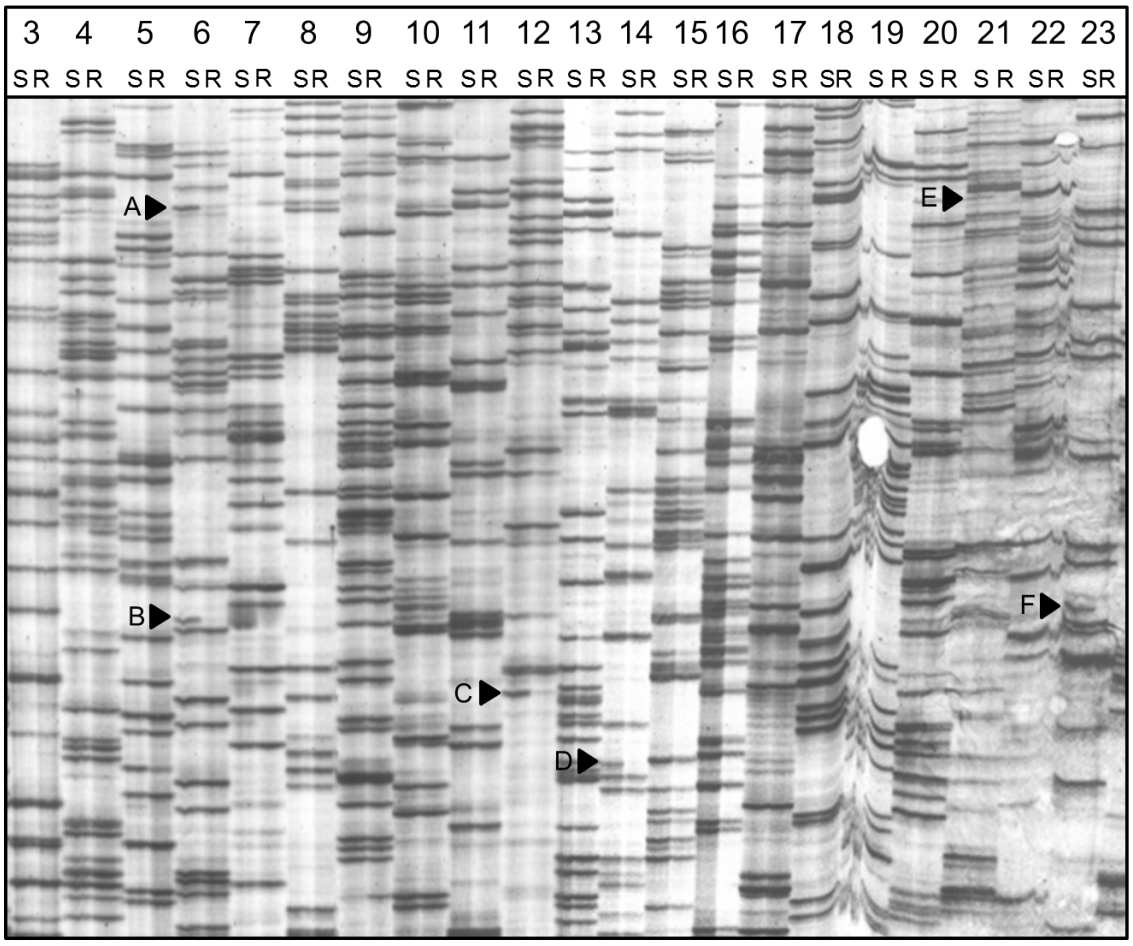


Figure 3.1 AFLP analysis of susceptible *'S'* and resistant *'R'* bulked segregating populations. (The complete gel is not shown) The gel shows the *Pst* primer results in combination with 20 *Mse* primers (details of primers can be found in Table 3.1) numbered 3-23. Reactions were carried out with both the susceptible *'S'* and the resistant *'R'* bulked DNA samples, and run adjacently for each primer combination. DNA fragments highlighted with black arrows and labelled A-F are putative AFLP™ markers linked to the Col-0 susceptibility gene.

Table 3.1 Primer combinations used for AFLP™. The Pst Primer was used in conjunction will all Mse primers listed, and used in combinations for which the results are shown in Figure 3.1. Adapter sequences are shown in bold.

Primer Set Number	Primer ID	Sequence
All	P14	GACTGCGTACATGCA**GAT**
1	M31	GATGAGTCCTGAGTAA**AAA**
2	M32	GATGAGTCCTGAGTAA**AAC**
3	M33	GATGAGTCCTGAGTAA**AAG**
4	M34	GATGAGTCCTGAGTAA**AAT**
5	M35	GATGAGTCCTGAGTAA**ACA**
6	M36	GATGAGTCCTGAGTAA**AAC**
7	M37	GATGAGTCCTGAGTAA**ACG**
8	M38	GATGAGTCCTGAGTAA**ACT**
9	M39	GATGAGTCCTGAGTAA**AGA**
10	M40	GATGAGTCCTGAGTAA**AGC**
11	M41	GATGAGTCCTGAGTAA**AGG**
12	M43	GATGAGTCCTGAGTAA**ATA**
13	M44	GATGAGTCCTGAGTAA**ATC**
14	M45	GATGAGTCCTGAGTAA**ATG**
15	M46	GATGAGTCCTGAGTAA**ATT**
16	M47	GATGAGTCCTGAGTAA**CAA**
17	M48	GATGAGTCCTGAGTAA**CAC**
18	M49	GATGAGTCCTGAGTAA**CAG**
19	M50	GATGAGTCCTGAGTAA**CAT**
20	M51	GATGAGTCCTGAGTAA**CCA**
21	M52	GATGAGTCCTGAGTAA**CCC**
22	M53	GATGAGTCCTGAGTAA**CCG**
23	M54	GATGAGTCCTGAGTAA**CCT**
24	M81	GATGAGTCCTGAGTAA**TAG**

The AFLP gel in Figure 3.1 shows that the AFLP profiles of the resistant and susceptible bulks are similar, as expected. However, some markers only appear in the susceptible DNA pool, and not in the resistant pools. This is consistent with the hypothesis that susceptibility to BMYV infection in Col-0 is conferred by a single dominant gene. These AFLP markers were excised from the gel and re-hydrated. In total six markers were isolated for further investigation.

3.4 Sequence analysis of AFLP™ identified DNA fragments

The AFLP™ fragments labelled A-F in Figure 3.1 were analysed by DNA sequencing. The sequences of four fragments (A, B, C and F) were successfully obtained and are shown in Table 3.2. Analysis using BLAST (http://www.arabidopsis.org/Blast) indicated that all four fragments mapped to *A. thaliana* chromosome 4, and a map showing the approximate locations of the fragments can be seen in Figure 3.2.

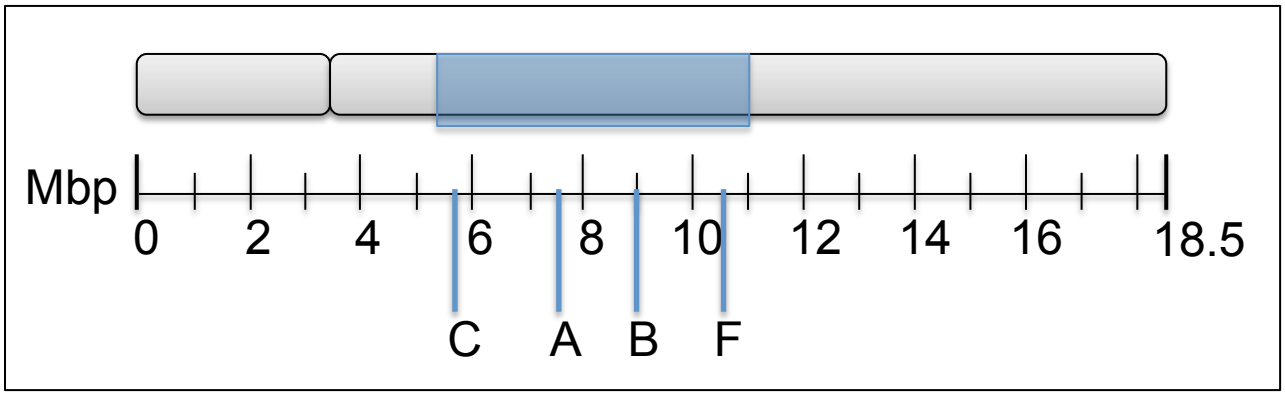


Figure 3.2 Map of *A. thaliana* chromosome 4 indicating the approximate genomic location of four DNA fragments identified through AFLP™. The chromosome map shows the locations of fragments A, B, C and F that were identified in Figure 3.1. Fragments have been mapped to the Col-0 *A. thaliana* genome.

Table 3.2 DNA sequences of fragments identified using AFLP™. Results show the sequence letter that relates to the labeled sequence in figure 3.1. Chromosomal location is on chromosome 4 of *A. thaliana* Col-0, and the percentage sequence identity to the published Col-0 genome.

Fragment	Chromosomal location (bp)	Contig length (bp)	Sequence (5'-3')	% identity to Col-0 published sequence
A	7533846 - 7534207	362	TGCAGCATGCGGAGTCTTGGTCTTATCGAGCTTCCCTGAATC CGTCTGCTTATACACAGTAACTATAATTACTTCCTAAATAAAAC TCATGTAGACTAGTATGAAGAAATGACAAGAGAACCAGTTTC AAACCTTGTTTTCAGCTGAACCTCCTCCTAATGTAGGAAGAC CCATATGAAGTATATACTCGGAATCAACAATACCGATATTTTTG GTTCGATCACCCTATGCCAAGCAACAACAACAATA CTAAGA	100
B	9019869 - 9020048	180	TGCAGCAGGCTCTGCTGGCAGCTGAAGGGAAACATTGGAGA TTGTTTGCAAAAGACAAACTCGATCGCCAGCCCATGTAGAAG AATTACACTCGACCCCCTGTTCATCACCTTGTTCATCTAGCCC TTCAGAAATAGCCAAATTATCATACATACATTCATGAAAATTGG TTTACTCAGG	96
C	5781417 - 5781520	102	TATTTCCGGTGGAAAGATAACAACTTGAACTCCGATTTGGCTT TCGATCGTAACAACGTGAAATGCTTGATTCATTTCATTGTAAA AATTTGTTGCCTTTCT	86
F	10618118 - 10618291	174	TGCAGCAAATATACGGGGATCCATTTCATCTAACGGGAATACC GGCGAAGAAGTAAGTCTCATCCGGTACTGAATAGAAGATCCT AAACTGTGAGAAGTGGTAGATTATCAGAAAAAGAACTGAAAC AACGACCATCATGACTCTGAATATTGGGCTTACGCCTTTGCA GGTTA	100

All four AFLP markers mapped to *A. thaliana* chromosome 4, in a region between 5.7 and 10.6 Mbp. The results give a strong indication that the phenotype being investigated is encoded within this region. Further bioinformatic analysis of this region on chromosome 4 was carried out, identifying a large number of gene sequences, including one for the eukaryotic translation initiation factor 4E (*eIF4E*) (located at 10016567-10018228 Mbp). The protein product of this gene has been described previously as an important factor for RNA virus infection, specifically in the Potyvirus family (Nieto et al., 2006; Piron et al., 2010; Robaglia and Caranta, 2006; Ruffel et al., 2004).

3.5 Sequencing Sna-1 *eIF4E* and *eIF4G*

Bioinformatics analysis revealed a potential gene controlling resistance and susceptibility to BMYV in Sna-1 and Col-0 *A. thaliana* ecotypes respectively. To investigate this further the Sna-1 *eIF4E* allele was sequenced in order to determine if there were any sequence differences when compared to the published Col-0 sequence. Alongside the analysis of the *eIF4E* genes, eukaryotic translation initiation factor 4G (*eIF4G*) was also sequenced. The eIF4G protein interacts with the eIF4E protein to form translational machinery in the cell. Although *eIF4G* is encoded for on *A. thaliana* chromosome 3, this gene was sequenced as a control.

DNA was extracted from leaf material as described in section 2.2.4. Primers were designed to amplify fragments of around 600 bp, with overlapping regions of at least 100 bp that covered the complete protein coding regions of these genes. The primers designed for both *eIF4E* and *eIF4G* sequencing reactions are shown in Table 3.3.

Table 3.3 Primers for sequencing the *eIF4E* and *eIF4G* genes of the *A. thaliana* ecotype Sna-1. A) Shows sequences of primers for *eIF4E* amplification, and B) shows sequences of primers for *eIF4G* amplification.

A

Primer Set	Primer name	Sequence 5'-3'
1	eIF4E R1 (H62)	GGGCTCTTGTCAAGTAAACAT
	eIF4E F1 (H63)	ATTAGAGGAAAGCAAGAAAGG
2	eIF4E R2 (H64)	CACAGAGAGACTGTTTGATGAG
	eIF4E F2 (H65)	TGTGAAGTAAAGTAGAAGAGAC
3	eIF4E R3 (H66)	CATCAACCCTCAGTCATC
	eIF4E F3 (H67)	ACTCCCAAATCTGTTCTAAC
4	eIF4E R4 (H68)	ATTCGTCAACGTTTCCGTCT
	eIF4E F4 (H69)	TTGTAGCTGCTAGCGATCAAC

B

Primer Set	Primer name	Sequence 5'-3'
1	eIF4G F1 (H34)	CCGTCCAATAAAACCCTAAT
	eIF4G R1 (H35)	CGGTTCTCCTGTAACTAACAGT
2	eIF4G F2 (H36)	GGTATGATCGTTATTCACGG
	eIF4G R2 (H37)	GAATTATTTATGTAGGCAAC
3	eIF4G F3 (H38)	TTCCTGCTCGAACTACCTCA
	eIF4G R3 (H39)	CGGGAGTCTGCATATGCATA
4	eIF4G F4 (H40)	CAAGTCAAACGCAGAAGTCT
	eIF4G R4 (H41)	GACATTGGACCATTATTTAACG
5	eIF4G F5 (H42)	ACCACATTCTAACCCACCTC
	eIF4G R1 (H43)	GATTTCAGTCGTTGTAACTGG
6	eIF4G F2 (H44)	GTCTCTGGAGTACCAAATTC
	eIF4G R2 (H45)	CCTGACAGTGTTTCGTGTTT
7	eIF4G F3 (H46)	AAATCTACGGAAGGTTCAA
	eIF4G R3 (H47)	GCATTCACAGAATTATCTGC
8	eIF4G F4 (H48)	GAACCTGTTACCTGCCATAC
	eIF4G R4 (H49)	GTTGGAAATTTGAACCACGA
9	eIF4G F1 (H50)	GAGGAAACTCGGGAGTTTA
	eIF4G R1 (H51)	CTTCGCACCTGAAGTCTTTT
10	eIF4G F2 (H52)	GCTTCTCAATAAATGTCAGG
	eIF4G R2 (H53)	AGCTGGAGGTGATAGCATTC
11	eIF4G F3 (H54)	CAGAGATGCTGCACAAGAAC
	eIF4G R3 (H55)	CTAACTGTTCTTCAGACAAAGC
12	eIF4G F4 (H56)	CAGCTTATGTTCATAGGGAA
	eIF4G R4 (H57)	CTTTCATGCACATACCAATC
13	eIF4G F1 (H58)	GTGCGTTTTATGTATGATGT
	eIF4G R1 (H59)	ATCTCAACCAACGTTTCTTC
14	eIF4G F2 (I35)	CCCTTTAACCTCATGTGGTAA
	eIF4G R2 (I36)	GGGAAAAGTGTGACAGAGAAA

The sequenced fragments were then overlapped and constructed into a full length sequence using the *BioEdit* sequence alignment editor (Hall, 1999).

3.5.1 Sequencing Sna-1 *eIF4E*

The full sequence of Sna-1 *eIF4E* had a length of 1441 nt. This sequence was directly compared with that of the Col-0 *eIF4E* using the NBI EMBOSS Water (http://www.ebi.ac.uk/Tools/psa/emboss_water/) alignment tool. Results of the alignment showed 98.7 % sequence identity between Col-0 and Sna-1 *eIF4E*. This also included a 12 nt insert in the Sna-1 *eIF4E* gene, which is a directly duplicated 12 nt sequence. Figure 3.3 shows the 12 nt insert, and the duplicated sequence. The full nucleotide sequence and alignment showing introns and exons can be found in Appendix A and B.

```
Sna-1 EIF4E  1351 TGTTTGGTTTGATTTC-TTTTCTTTCAGGAGGATGCGAAGAAGCTCGACA   1399
                  ||||||||||.|||| |||||.||||||||||||||||
Col-0 EIF4E  1350 TGTTTGGTTTGGTTTCTTTTTTTTTCAGGAGGATGCGAA-----------   1388

Sna-1 EIF4E  1400 GGAAGCTCGACAGGAATGCAAAGAACGCTTACACCGCTTGA   1440
                   |||||||||||||||||||||||||||||||||||||||
Col-0 EIF4E  1389 -GAAGCTCGACAGGAATGCAAAGAACGCTTACACCGCTTGA   1428
```

Figure 3.3 Fragment of EMBOSS WATER alignment of the Sna-1 and Col-0 *eIF4E*. The 3' end of the sequence alignment, with the 12 nucleotide insert of Sna-1 sequence highlighted in blue. The duplicated 12 base pair sequence is underlined in red. Numbers denote nucleotide position in the full gene sequence.

The 12 nucleotide insert found in the Sna-1 *eIF4E* sequence, highlighted in Figure 3.3 is found in the 3' end of the gene. This insert was found to be an in-frame addition after analysis of the predicted amino acid sequence, seen in Figure 3.4 (full Sna-1 predicted amino acid sequence can be found in appendix C).

```
Sna-1eIF4E          1 MAVEDTPKSVVTEEAKPNSIENPIDRYHEEGDDAEEGEIAGGEGDGNVDE     50
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4E          1 MAVEDTPKSVVTEEAKPNSIENPIDRYHEEGDDAEEGEIAGGEGDGNVDE     50

Sna-1eIF4E         51 SSKSGVPESHPLEHSWTFWFDNPAVKSKQTSWGSSLRPVFTFSTVEEFWS    100
                      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4E         51 SSKSGVPESHPLEHSWTFWFDNPAVKSKQTSWGSSLRPVFTFSTVEEFWS    100

Sna-1eIF4E        101 LYNNMKHPSKLAHGADFYCFKHIIEPKWEDPICANGGKWTMTFPKEKSDK    150
                      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4E        101 LYNNMKHPSKLAHGADFYCFKHIIEPKWEDPICANGGKWTMTFPKEKSDK    150

Sna-1eIF4E        151 SWLYTLLALIGEQFDHGDEICGAVVNIRGKQERISIWTKNASNEAAQVSI    200
                      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4E        151 SWLYTLLALIGEQFDHGDEICGAVVNIRGKQERISIWTKNASNEAAQVSI    200

Sna-1eIF4E        201 GKQWKEFLDYNNSIGFIIHEDAKKLDRKLDRNAKNAYTA    239
                      |||||||||||||||||||||||    ||||||||||||
Col-0eIF4E        201 GKQWKEFLDYNNSIGFIIHEDAK----KLDRNAKNAYTA    235
```

Figure 3.4 Full predicted amino acid sequence alignment (EMBOSS WATER) of Sna-1 and Col-0 eIF4E. The blue box highlights the four amino acid insert in the Sna-1 sequence, and the red line indicates the duplicated four amino acid sequence.

The 12 base pair insert in the Sna-1 allele is in frame and would result in four amino acid sequence duplication in the protein. This insert could cause a change in the overall shape of the protein. Conceivably this could have an effect on virus-protein interactions within the cell, and might affect translation of virus RNA. Figure 3.5 shows the protein structure of eIF4E, and the estimated position of the amino acid insert (modelled by Dr. Andrew Hemmings *University of East Anglia*).

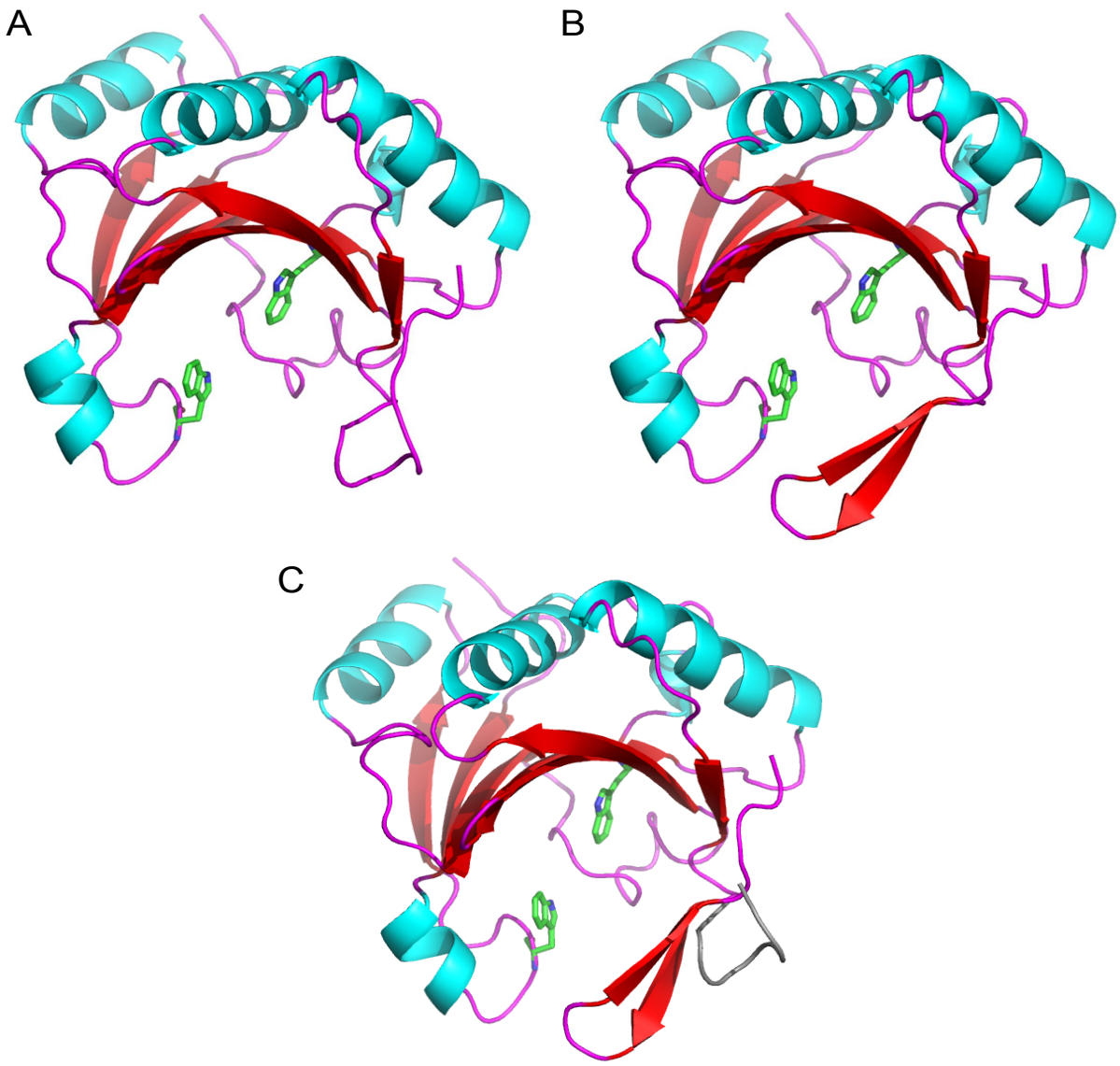


Figure 3.5 eIF4E protein structure. The predicted protein structure of eIF4E from Col-0 (A), Sna-1 (B) and an overlapping model of Col-0 and Sna-1 eIF4E (C). The green structure shows the predicted mRNA cap-binding domain. The four amino acid insert found in Sna-1 eIF4E can be seen as a larger red loop around the mRNA cap binding domain. In model C, the difference between the Col-0 and Sna-1 eIF4E proteins can be seen with the loop found in Col-0 shown in grey, and the Sna-1 four amino acid insert in red. The model was made using the program SwissModel™, and based on the crystal structure protein data bank 2WMCH *Pisum sativum* (Modelled by Dr. A. Hemmings, *University of East Anglia*).

Previous known sequence mutations causing recessive resistance to other RNA viruses have all been associated with the loops around the beta sheets (Robaglia and Caranta, 2006). This insert appears to be positioned near to the CAP binding domain of the eIF4E protein. The consequences of an alteration to the cap binding domain could be that the VPg of the RNA virus can no longer recognize the site, and is therefore unable to initiate translation, meaning the virus can no longer replicate inside the cell.

3.5.2 Sequencing Sna-1 *eIF4G*

The Sna-1 *eIF4G* was sequenced and assembled in the same way as *eIF4E*. The full sequence of Sna-1 *eIF4G* can be seen in Appendix D. The total fragment length of Sna-1 *eIF4G* determined was 7483 nt. Alignment with the Col-0 *eIF4G* sequence using the EMBOSS WATER tool showed 99.9% identity, with four single-base pair substitutions, all pyrimidine for purine or *vice versa* (sequence alignment is shown in Appendix E). The amino acid sequence (Appendix F) aligned with the Col-0 eIF4G (Appendix G) revealed an identity score of 99.9 %. The dissimilarity arises from two amino acid substitutions. Amino acid number 1487 in Col-0 is proline, and is changed to phenylalanine in Sna-1. These are both non-polar amino acids so the substitution might not have too great an effect on protein structure and function. The other amino acid substitution is amino acid 284, where a threonine in the Col-0 protein is substituted for proline in the Sna-1 protein. This is a bigger change as threonine is a polar molecule, but as threonine is not a charged molecule the substitution is likely to have little effect on the structure or function of the protein.

3.5.3 Cleaved Amplified Polymorphic Sequence (CAPS) analysis to identify Sna-1 and Col-0 *eIF4E*

In order to be able to identify plants containing the Sna-1 or Col-0 *eIF4E* gene in future crosses, a CAPS analysis was developed to easily distinguish the two alleles of this gene. This exploited a single nucleotide mutation in the Sna-1 *eIF4E* gene, at nucleotide position 1310, where an adenine base is switched

to a guanine base, creating the restriction sequence for the enzyme *BspHI*. This nucleotide switch is shown in Figure 3.6.

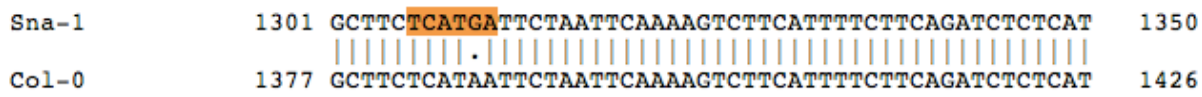


Figure 3.6 Alignment of the region of *eIF4E* containing the *BspHI* site in the Sna-1 sequence. The region highlighted in orange is the *BspHI* restriction enzyme site, only found in the Sna-1 version of *eIF4E* due to a single base switch of an 'A' to a 'G' at base pair number 1310.

PCR fragment of a section of the *A. thaliana eIF4E* gene with *eIF4E* primer set number 1 (described in Table 3.3A) generated an amplification of 640 bp. Digestion with the restriction enzyme *BspH1* of the Sna-1 *eIF4E* product creates two fragments of 406 bp and 244 bp. The enzyme is unable to digest the Col-0 fragment, and therefore the fragment remains at 640 bp. This allows easy genotyping of the *eIF4E* locus in all F_2 plants form a Sna-1 x Col-0 crosses. Examples of homozygous and heterozygous plants, as revealed by CAPS analysis, can be seen in Figure 3.7.

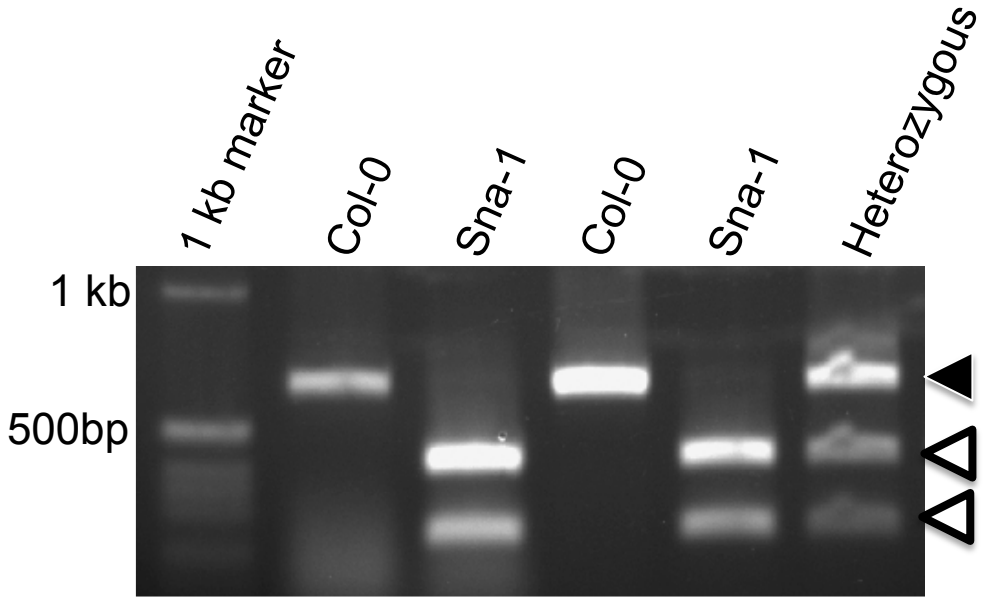


Figure 3.7 PCR amplified and *BspHI* digested *eIF4E* DNA fragments of *A. thaliana* plants. The DNA profiles of homozygous Sna-1 and Col-0 genotypes are shown, together with the profile of a Sna-1/Col-0 heterozygote. The black arrow indicates the 640 bp fragment produced by amplification of the Col-0 *eIF4E* gene, whilst the white arrows with black borders indicate the Sna-1 *eIF4E* gene fragments after digestion with *BspHI.*

Figure 3.7 shows a clear distinction between all three genotypes and can be carried out quickly and easily. This CAPS assay was used in subsequent analyses to determine the *eIF4E* genotypes of Sna-1 x Col-0 F_2 progeny.

3.6 Discussion

The results here have built on the previous observation by Percival-Alwyn (2010) that the *A. thaliana* Sna-1 ecotype is resistant to BMYV infection. This resulted from a study of natural variation in *A. thaliana*. The advantage of natural variation studies is that many of the ecotypes have already been collected, and in some cases have been examined for other resistance traits (Hall, et. al., 2009). The natural resistance exhibited in ecotypes of *A. thaliana* gives a good starting point for the investigation of resistance to BMYV without the need for time-consuming, repetitive creation of RILs, or the generation of

induced variants for susceptibility or resistance using costly mutagenesis strategies. Percival-Alwyn (2010) proposed that the geographical position of the resistant ecotype might reflect adaptive variation to selection pressure for resistance to BMYV as a result of high levels of the sugar beet crop, and the potential increased incidence of BMYV, in that area. East Anglia is well known for growing large quantities of sugar beet, and is home sugar beet processing plants. The Sna-1 ecotype, resistant to BMYV, was found in Suffolk and therefore may have gained resistance as a result of being in close proximity to sugar beet plants. Being resistant to the virus carried by aphids might provide a selective advantage over other *A. thaliana* plants, allowing the Sna-1 ecotype to thrive in that area. This theory however does not account for the other three ecotypes tested from the same region, Dun-1, Far-1, and Lew-0 (all isolates from Suffolk), which were found to be highly susceptible to BMYV. It is therefore not clear if this is a true adaptive mutation in the Sna-1 ecotype.

Percival-Alwyn (2010) also reported the resistance trait to be caused by a single recessive gene, after genetic analysis of the F_2 progeny from a Sna-1 x Col-0 cross. The pattern of resistance and susceptibility in these progeny showed that resistance to BMYV was caused by the lack of a dominant susceptibility gene i.e. the resistance observed is recessive and therefore is consistent with a passive rather than activated form of resistance. Passive resistance is thought to be much more durable than the alternative active resistance response (Jones and Dangl, 2006; Maule et al., 2007; Uma et al., 2011). This is because active, or hypersensitive resistance, occurs when the plants natural immune system is triggered by recognition of highly conserved pathogen associated molecular patterns (PAMPs) (Uma, et al., 2011), or the recognition of pathotype-specific effector (avirulence) proteins recognized plant R proteins (Jones and Dangl, 2006; Maule et al., 2007). This often leads to a response such as thickening of the cell walls, creation of reactive oxygen intermediates (ROI) or programmed cell death, to target and destroy the virus, or prevent its spread around the plant (Morel and Dangl, 1997; Uma et al., 2011). The alternative method is recessive, or passive resistance. In this type of resistance, the virus cannot infect the cell due to the inadequacy of the host cell machinery. If the virus is unable to use the host cell components to

replicate, then it cannot infect the cell. It is for this reason that these mutations in plants are referred to as a 'loss of susceptibility' (Ritzenthaler, 2005).

AFLP analysis, the sequence analysis of the linked markers and bioinformatics of the DNA sequences led to the identification of a possible source of this resistance, in the form of the *eIF4E* gene. The gene was found to be located in the area of chromosome 4 where fragments of DNA were identified to differ between resistant and susceptible pools of *A. thaliana*. The *eIF4E* gene (At4g18040) plays an important role in the recruitment of a protein complex in order to initiate translation of host cell mRNAs. The eIF4E protein contains a cap-binding domain as shown in Figure 3.5. It has been proposed that the cap binding region of eIF4E associates with the 5' end of mRNA, where a 5'-^{7m}GpppN-cap promotes translation and prevents degradation of the molecule (Kawaguchi and Bailey-Serres, 2002). The eIF4E protein is therefore induced to recruit the protein eIF4G, which acts as a scaffolding protein to recruit other proteins to form the translational machinery, as shown in Figure 3.8.

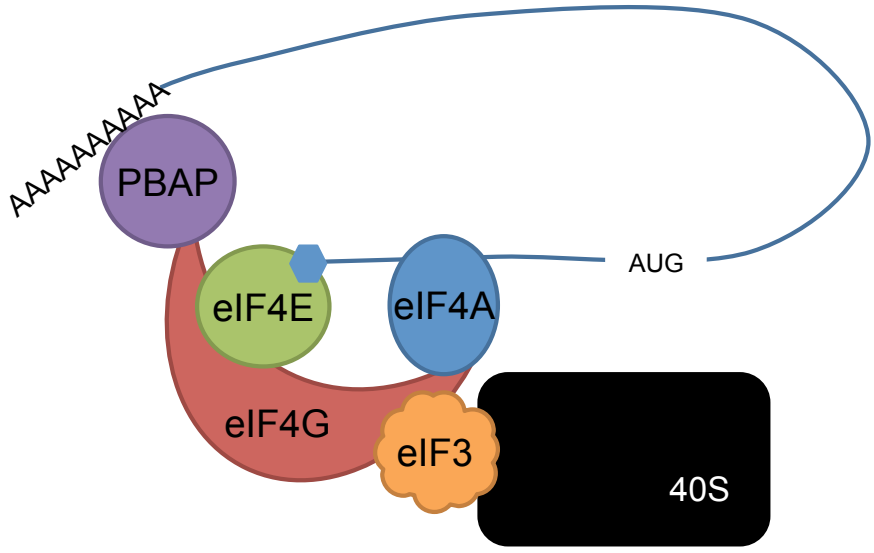


Figure 3.8. The recruitment of eukaryotic translation initiation factors and other proteins after initiation. Binding of the 5' mRNA cap structure (shown by the blue hexagon) to the eIF4E protein recruits the eIF4G scaffolding protein, forming the complex eIF4F. The poly-A binding protein (PABP) is then recruited to associate with

the 3' poly-A tail of the mRNA causing a looped mRNA structure. The eIF4A protein, a DEAD box helicase protein responsible for unravelling any secondary structures, is also recruited alongside eIF3 proteins. Multiple eIF3 proteins are recruited to form a complex, which in turn recruits the 40S ribosome subunit, alongside other associated proteins (not shown here). Following recruitment of proteins the mRNA is scanned for a suitable initiation codon (AUG). This then triggers the binding of the 60S ribosomal subunit and translation begins. Image adapted from Robaglia and Caranta, 2006.

Recessive resistance has been heavily linked to eIF4E previously in investigations of Potyvirus infections. Resistance mechanisms have been shown in a wide variety of crops, between which there are relatively few amino acid differences between eIF4E proteins (Robaglia and Caranta, 2006). This sequence has therefore been highly conserved in plants and other eukaryotes, and shows that it is an essential cellular component.

Potyviruses are single stranded, positive sense RNA viruses of around 10kb. Their genome contains a virus encoded protein (VPg) at the 5' end, as well as a 3' polyadenylated tail (Nicaise et al., 2007). The VPg is linked to the viral RNA by a tyrosine residue, and has been shown to play an important role in replication and translation of the viral genome (Eskelin et al., 2011; Murphy et al., 1991). During examination of the VPg in a yeast-two hybrid system, it was shown that the VPg has the ability to interact with the cap binding pocket of eIF4E (Léonard et al., 2000; Schaad, et al., 2000), initiating translation of the viral RNA. It has also been shown that a mutation in the VPg of the Potyvirus *Turnip mosaic virus* (TuMV) leads to the inability of the virus to interact with translation initiation factors, thus leaving the virus unable to infect its host (Léonard et al., 2000).

The relationship between single stranded positive sense RNA viruses and eIF proteins has been demonstrated in crops, including lettuce and pea, as well as in the model organism *A. thaliana.* Resistance to *Lettuce mosaic virus* was seen from point mutations, and also the deletion of sequences in the *eIF4E* gene (Nicaise et al., 2003). *Pea seed borne mosaic virus* is also unable to infect pea plants after amino acid substitutions in eIF4E were shown to inhibit

cell-to-cell movement of the virus (Gao et al., 2004). In pepper plants, an amino acid substitution in eIF4E homologues also induced resistance to *Potato virus Y* and *Tobacco etch virus* (Ruffel et al., 2005). All these naturally occurring mutations are inherited as recessive traits, meaning changes in amino acid sequence of eIF4E compromise virus infection. These papers also showed that the region in which these amino acid substitutions are occurring in eIF4E is in the cap-binding pocket. This supports the hypothesis that these Potyvirus infections are at least partially determined by the ability of the VPg to interact with the cap-binding domain of eIF4E.

Currently published data has only described the substitution of amino acids in eIF4e to cause resistance to Potyviruses. The sequencing analysis described here showed a 12 bp sequence duplication in the Sna-1 allele of *eIF4E* that results in a four amino acid duplication, with respect to the Col-0 allele (Figure 3.3). From this we must deduce that the insert in this case is inducing resistance in the Sna-1 plant, rather than an amino acid substitution event. The positioning of the insert is important. It too is around the cap-binding domain (Figure 3.5) as has been described with the substitutions in other plants inducing resistance. It is possible that the sequence duplication compromises interaction with BMYV RNA in a similar way.

All the viruses that have been discussed so far are from the Potyvirus family. This is the largest family of virus, and have so far made up the majority of discoveries for recessive resistance genes with over 60% of all known recessive resistance mechanisms relating to *Potyviruses* (Diaz-Pendon, et. al., 2004). This is not surprising as the viral genome, with its 5' VPg and 3' polyadenylated tail, is similar in structure to most host mRNAs in plants, with a predicted functional similarity of the viral VPg and the mRNA 5' cap. Other viruses however, like BMYV a Luteovirus, do not have polyadenylated (polyA) tails, and some viruses lack both a polyA tail and a VPg (Kneller, et al., 2006). Poleroviruses do contain a VPg at the 5' end, which substitutes for the normal cap structure of mRNAs, but do not possess a 3' polyadenylated tail (Brault, et al., 2011). Infection by other viruses not of the Potyvirus family has been seen. The Carmovirus, *Melon necrotic spot virus*, lack both a 5' VPg and 3'

polyadenylated tail. Resistance to this virus in melon was caused by a single amino acid substitution in the melon homologue of eIF4E. It was later revealed that the 3' untranslated region (3'-UTR) was interacting with eIF4E, although further details have not yet been elucidated (Nieto et al., 2006). This shows that recessive resistance genes relating to *eIF4E* are not limited to Potyviruses.

There is very little information about natural recessive resistance mechanisms in Luteoviridae family relative to the Potyvirus family. Recently a study by Reinbold et al. (2013) has shown through mutagenesis that eIF4E is an important factor for BMYV infection, which this study supports. Further investigation of the Luteovirus family has shown that other components of the eIF4F and isoforms of the protein are important for infection with *Turnip yellows virus* (TuYV), another Luteovirus (Reinbold et al., 2013).

The eIF4G protein has been widely reported as another important recessive resistance factor to infection with viruses, including many Potyviruses. This cap-independent method of translation seen in *Tobacco etch virus* makes eIF4G a target for viral infection also (Gallie, 2001).

Isoforms of the eIF4F complex also exist (these are shown with their known interactions in Figure 6.13). The proteins eIF4E and eIF(iso)4E share around 40-45 % sequence homology, whilst the eIF4G and eIF(iso)4G proteins share 35% sequence identity (Kawaguchi, and Bailey-Serres, 2002). In *A. thaliana* the isoforms of eIF4F (eIF(iso)4F) complex are encoded for on separate chromosomes to the *eIF4F* genes, with *eIF(iso)4E* and *eIF(iso)4G* on chromosome five. It is thought that eIF4F and eIF(iso)4F are not utilised for the same function, as their mRNAs are differentially expressed around the plant and during different phases of growth. In *A. thaliana* eIF(iso)4F was found to be expressed predominantly in flowers and the developing tissues, whilst eIF4F was expressed throughout the plant, except in the roots (Rodriguez et al., 1998). It is in fact the isoforms of the eIF4F complex that many viruses are affected by when mutations take place. It is well documented in viruses such as *Turnip mosaic virus* and *Lettuce mosaic virus*

that infection is dependent upon the eIF(iso)4E components, found through mutagenesis of *eIF(iso)4E* (Duprat et al., 2002). The Luteovirus TuYV, closely related to BMYV, is shown to be reliant on eIF(iso)4G for infection in *A. thaliana,* as similar mutagenesis of the gene induced resistance to the virus. This eIF(iso)4G-induced susceptibility to TuYV indicates that even closely related viruses such as BMYV and TuYV can have significantly different infection strategies (Reinbold et al., 2013).

In some cases of virus infection, the virus is not limited to the use of one of the isoforms. *Pepper veinal mottle virus* (PVMV) has been shown to be able to use both the equivalent eIF4E and eIF(iso)4E in pepper plants, resulting in infection. The mutation of both of these genes (one of either *eIF4E* or *eIF(iso)4E* is mutagenised, whilst the other exhibits the natural mutant resistance form of the gene) creates a PVMV resistant plant, whilst the mutant form of just one gene allows the plant to be infected (Ruffel et al., 2006).

Previous studies of naturally occurring recessive resistance genes suggest that whilst the virus can no longer use the eIF machinery, the plant can still function as normal (Duprat et al., 2002; Lellis, et al., 2002; Léonard et al., 2000; Ruffel et al., 2002). Further investigation into Sna-1 eIF4E functionality would be interesting to see if this is also the case for this natural mutation, or if the plant is relying on the isoform of the gene for translation of its mRNAs. As the plant is able to grow normally whilst containing the Sna-1 eIF4E, it is likely that the gene is still functional.

Further investigation into the function of this gene in relation to infection with BMYV is now required. Through investigation of knock-out mutations, and point mutations in the *eIF4E* gene, infection studies and complementation experiments, the importance of this gene for BMYV infection will be further explored.

Chapter 4 Investigating the role of *A. thaliana* *eIF4E* and *eIF4G* in BMYV infection.

4.1 Introduction

Previous work (Chapter 3) identified the *eIF4E* gene as a candidate BMYV susceptibility gene in *A. thaliana*. Natural variation in this gene was shown in resistant and susceptible *A. thaliana*, and as discussed in Chapter 3. The gene has previously been shown to be an important susceptibility factor in plants, for various RNA viruses. One way to investigate the importance of this gene in BMYV infection is to exploit the considerable genetic resources available in Arabidopsis in the form of well characterised mutations in any gene of interest (Swarbreck et al., 2008). These include T-DNA knock out mutants of *eIF4E* and previously characterised EMS mutants (McElver et al., 2001; Yoshii et al., 2004; Alonso et al., 2003).

In recent years a wealth of genome sequencing information, especially in *A. thaliana*, has been gathered. The next big challenge is to identify genes and their functions. Many gene functions can be deciphered with the use of assays and sequence homology studies, by comparing sequence information with genes in other organisms whose functions are known (Wesley et al., 2001). When these options are not available however, mutagenesis by chemicals, radiation, transposons, T-DNA and RNAi silencing can be utilised to allow functional analysis of target genes (Alonso et al., 2003; Ashrafi et al., 2003; Azpiroz-Leehan and Feldmann, 1997; Martienssen, 1998).

The use of chemicals and radiation in mutagenesis has a disadvantage compared to other methods as mutagenesis is essentially random and difficult to trace. Other methods such as RNAi, allow a gene-targeted approach. RNAi however, does not normally allow stable inheritance of the silenced gene, and may not produce a full silencing effect (Alonso et al., 2003). Insertion

mutagenesis offers a solution to this problem, as gene disruption and inactivation is usually heritable, and traceable in plants (Alonso et al., 2003).

Insertional mutagenesis results from random insertion of a known fragment of DNA into the genome being studied, potentially resulting in a loss of gene function. This technique was first carried out using transposable elements (Azpiroz-Leehan and Feldmann, 1997). Transposons are short sequences of genetic material that are able to move from one site to another within genomes. Whilst transposons allow the easy generation of a large number of insertion populations, T-DNA insertions create fewer insertions, but those that are created have a more stable nature, and show less bias to where they are inserted (Azpiroz-Leehan and Feldmann, 1997; Bouchez and Hofte, 1998).

T-DNA (transfer DNA) is a section of genetic information from the tumour-inducing (Ti) plasmid of *Agrobacterium tumefaciens*, surrounded by defined border sequences. In nature Agrobacterium uses the Ti plasmid to insert genetic information into a plant cell, where it is expressed in order to create nutrients required for the bacteria by creating plant tumours. T-DNA technology has modified this process to remove the tumour creating genes, and instead allows the insertion of a specific section of DNA, alongside a reporter gene, such as antibiotic resistance (Azpiroz-Leehan and Feldmann, 1997). A diagram of T-DNA insertion can be seen in Figure 4.1.

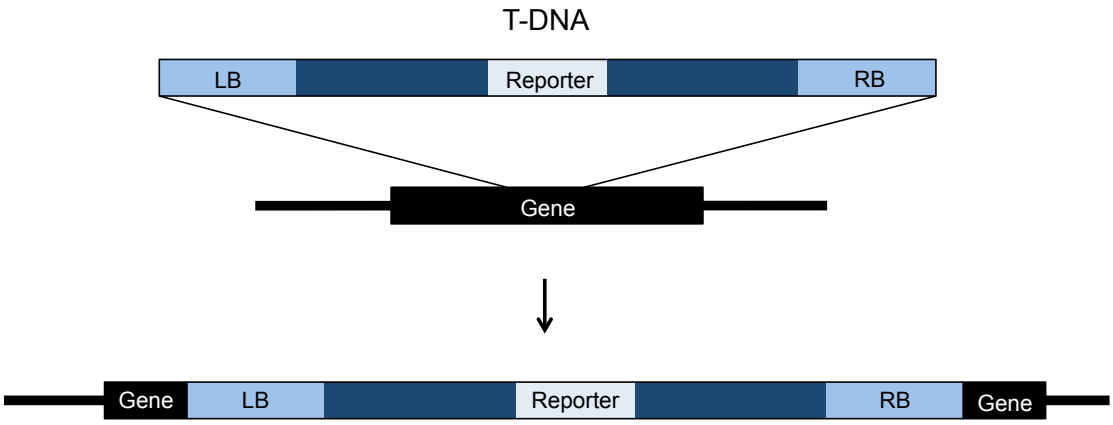


Figure 4.1 T-DNA insertions into a random gene. T-DNA can be seen containing a left border (LB) sequence and a right border (RB) sequence. The "reporter" section is often antibiotic resistance. T-DNA insertion normally results in a disrupted plant gene and consequent inactivation of the encoded protein. (Modified from Azpiroz-Leehan and Feldmann, 1997).

Many large scale insertion mutagenesis studies have taken place in *A. thaliana*, including Alonso et al. (2003), where T-DNA insertions were reported in more than 21,700 genes. Technologies such as Inverse PCR can then be used to determine the T-DNA flanking regions, and as a result the precise genome location of the insert can be deduced. The process involves restriction digestion of genomic DNA from plants with T-DNA inserts, followed by ligation of the resulting DNA fragments. PCR is then carried out on the ligated fragments using primers complementary to the T-DNA/gene junctions. Sequencing of the fragments produced from this PCR allows identification of the sites where the T-DNA has integrated (Alonso et al., 2003; Azpiroz-Leehan and Feldmann, 1997).

With so many individual insertion sites, in some cases multiple insertion sites within a gene occurs. It is increasingly likely to find an Arabidopsis line available with a T-DNA insert in any gene of interest. As there are now so many lines available, TAIR and NASC list *A. thaliana* ecotypes available to order, T-DNA lines, and maps of the insert locations. T-DNA insert lines were therefore ordered containing loss of function mutations in both *eIF4E* and *eIF4G*.

T-DNA insert lines in *eIF4E* and *eIF4G* were then used in infection studies with BMYV. So far, T-DNA insertions have been discussed with relevance to understanding gene function within the plant, but this technique also allows the plants fitness to be tested under different conditions, and compared to wild type plants (Bouchez and Hofte, 1998). In this study the susceptibility and resistance of the plants to BMYV were tested. This was done by comparing infection levels of plants containing T-DNA inserts in *eIF4E,* or *eIF4G,* to wild type susceptible Col-0, and resistant Sna-1. If BMYV infection requires the use of eIF4E, by taking a naturally susceptible plant, and inducing loss of function (i.e. with a T-DNA insert) in this gene, the plant should become more resistant to the virus. In addition to T-DNA insertion lines in *eIF* genes, two characterised *eIF* mutants identified in forward genetic screens for resistance to *Cucumber mosaic virus* (CMV) in Arabidopsis were also utilised (Yoshii et al., 2004).

4.2 *eIF4E* Knock-out mutations

The *eIF4E* gene is located on chromosome 4 between 10016567 - 10018228 bp (TAIR gene model AT4G18040.1). TAIR identified a T-DNA insertion in this gene in a line designated SALK_145583C. This T-DNA mutation created by Alonso et al. (2003) occurs within an intron, as shown in Figure 4.2.

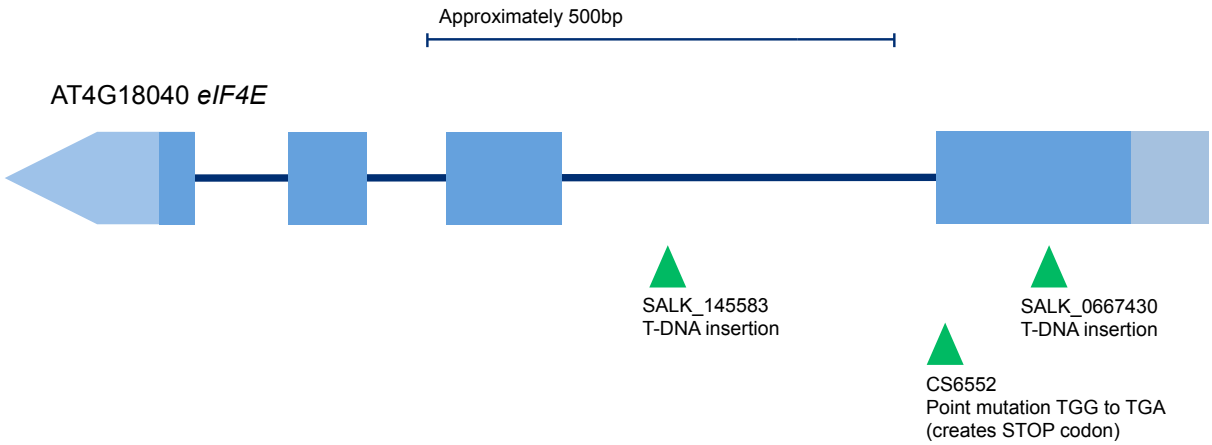


Figure 4.2 *eIF4E* gene schematic showing T-DNA insertion and point mutation sites. The mutation locations in *eIF4E* mutant lines SALK_145583, SALK_0667430 and CS6552. The T-DNA insertion in SALK_0067430 occurs in exon 1, as does the EMS induced point mutation in the CS6552 line. The T-DNA insertion in SALK_145583C occurs within an intron. All mutations are marked with a green triangle in the diagram. The mutations are all in the background Col-0. Further details can be found in Table 4.1.

A BMYV infection study using Col-0, Sna-1 and SALK_145583C was set up to investigate the effect of *eIF4E* on BMYV infection. Results are shown in Figure 4.3. Infection studies and analysis were carried out as described in section 2.3. Throughout this study, all error bars were calculated as follows:

$$\text{Standard Error} = \frac{\text{Standard Deviation}}{\sqrt{n}}$$

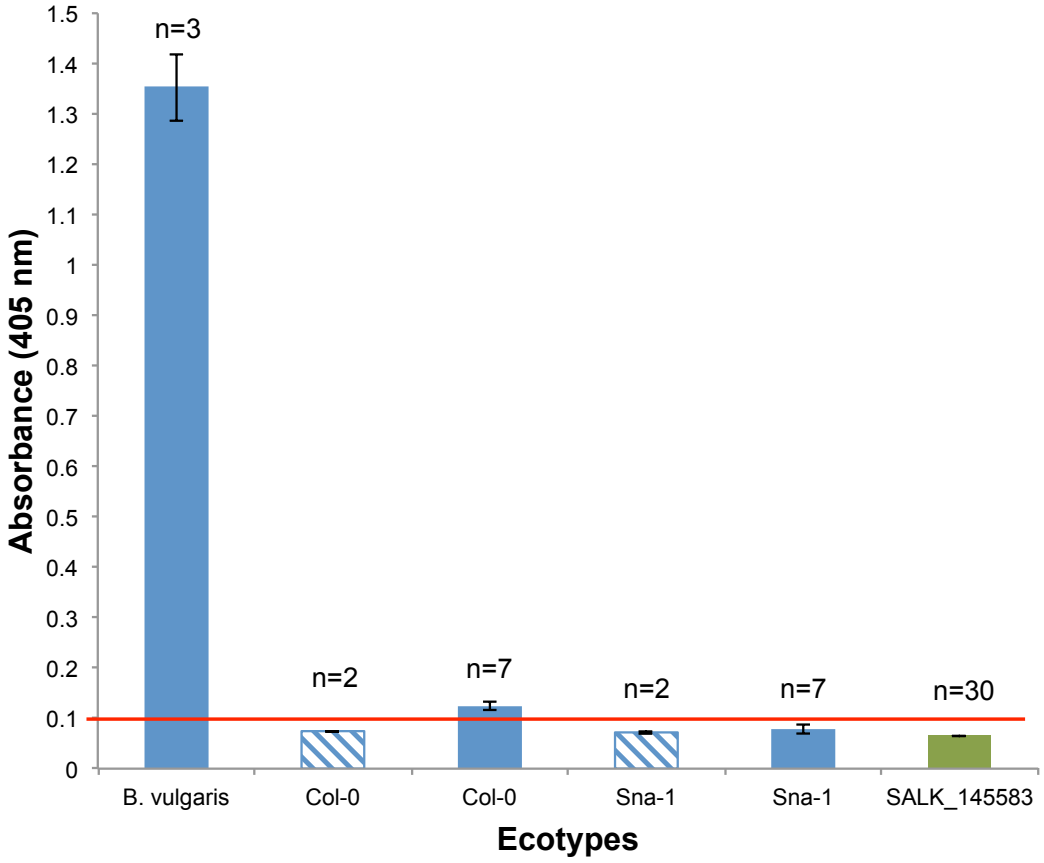


Figure 4.3 *eIF4E* T-DNA insertion BMYV infection study TAS-ELISA results. The results show *B. vulgaris*, the positive control for BMYV infection; viruliferous aphid insects were allowed to feed on, *Col-0*, *Sna-1*, and the *Col-0 eIF4E* KO line SALK_145583C (T-DNA insertion, shown in green) plants. Striped bars indicate plants that were not exposed to aphids, whilst solid bars show plants that were exposed to viruliferous aphids for 1 week. "n" refers to the total number of plants tested, and error bars show the standard error. The red line shows the computer-derived threshold of resistance to BMYV (above indicates susceptible plants and below resistant).

The TAS-ELISA results shown in Figure 4.3 indicate, as previously shown by Percival-Alwyn (2010), that Col-0 is susceptible to BMYV, and Sna-1 is resistant. The red line seen in Figure 4.3, at an absorbance of 0.1, is representative of a threshold value, above which a plant can be considered infected with BMYV. Uninfected controls are all seen to be below the threshold, whilst infected *B. vulgaris* and Col-0 are seen at values above 0.1. The T-DNA insertion mutation in the *eIF4E* gene was also shown to be

resistant. This suggests that mutation of this gene has caused the plant to become resistant to BMYV. More, and varied, mutations in this gene are also needed to be tested to substantiate this hypothesis.

4.3 Analysis of *eIF4G* knock-out lines

As previously discussed in Chapter 3, eIF4E is a protein that forms a complex with other proteins in order to form translation initiation machinery. Another important component in this complex, which has also been shown to be a virus susceptibility factor is eIF4G. Although the sequenced Sna-1 *eIF4G* showed good sequence homology to Col-0 *eIF4G* (section 3.5.2), this does not indicate that *eIF4G* does not have an impact on BMYV infection. In order to test this, a knock-out mutation in *eIF4G*, SAIL_87_A01 (Sessions et al., 2002), that contains a T-DNA insertion in the first exon of the gene (Figure 4.4) was tested in BMYV infection studies.

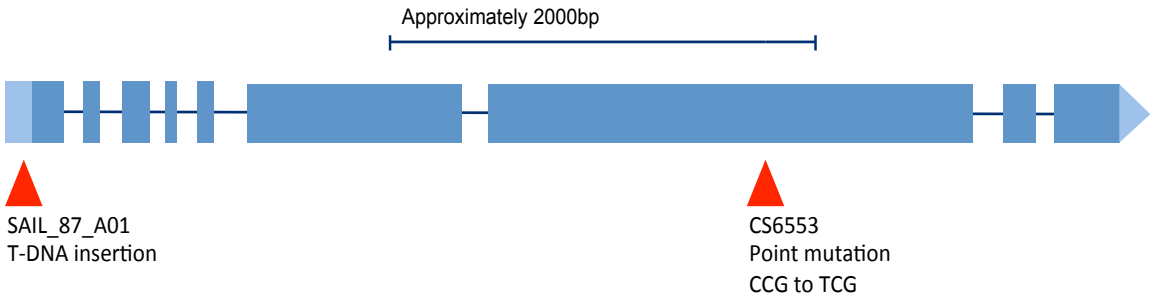


Figure 4.4 *eIF4G* gene schematic showing T-DNA insertion and point mutation sites. The mutations in the mutant lines SAIL_87_A01 and C56553 occur in exons marked with a red triangle in the diagram. The SAIL_97_A01 mutant line contains a T-DNA insertion in exon 1, whilst the C56553 contains an EMS induced point mutation in exon 7. Both mutations are in the background Col-0. Further details are found in Table 4.1.

The SAIL_87_A01 *eIF4E* T-DNA insertion line was exposed to BMYV, and tested for infection using TAS-ELISA. If *eIF4G* were important in the process of BMYV infection, it would be expected that the *eIF4G* knock out would be more resistant than the Col-0 ecotype. Results for these experiments are shown in Figure 4.5

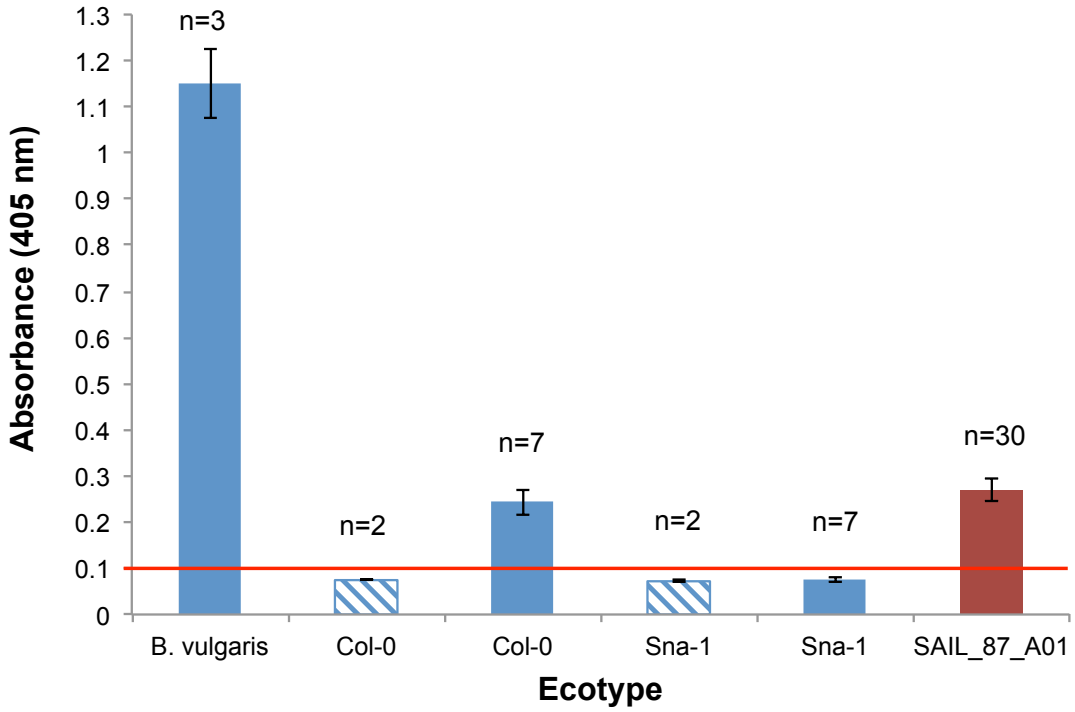


Figure 4.5 *eIF4G* T-DNA insertion BMYV infection study TAS-ELISA results. The results for four plants are shown; *B. vulgaris*, the positive control BMYV viruliferous insects were allowed to feed on, *Col-0, Sna-1*, and the *Col-0 eIF4G* KO SAIL_87_A01 (T-DNA insertion, shown in red) plants. Striped bars indicate plants that were not exposed to infection, whilst solid bars show plants that were exposed to viruliferous aphids grown on them for 1 week. "n" refers to the total number of plants tested, and error bars show standard error. The red line shows the computer-derived threshold of resistance to BMYV (above susceptible, below resistant).

The TAS-ELISA results shown in Figure 4.5 indicate that T-DNA insertion in *eIF4G* has not compromised susceptibility of Arabidopsis to BMYV. This suggests that BMYV is not dependent on eIF4G in the translation initiation complex in order to successfully infect *A. thaliana*.

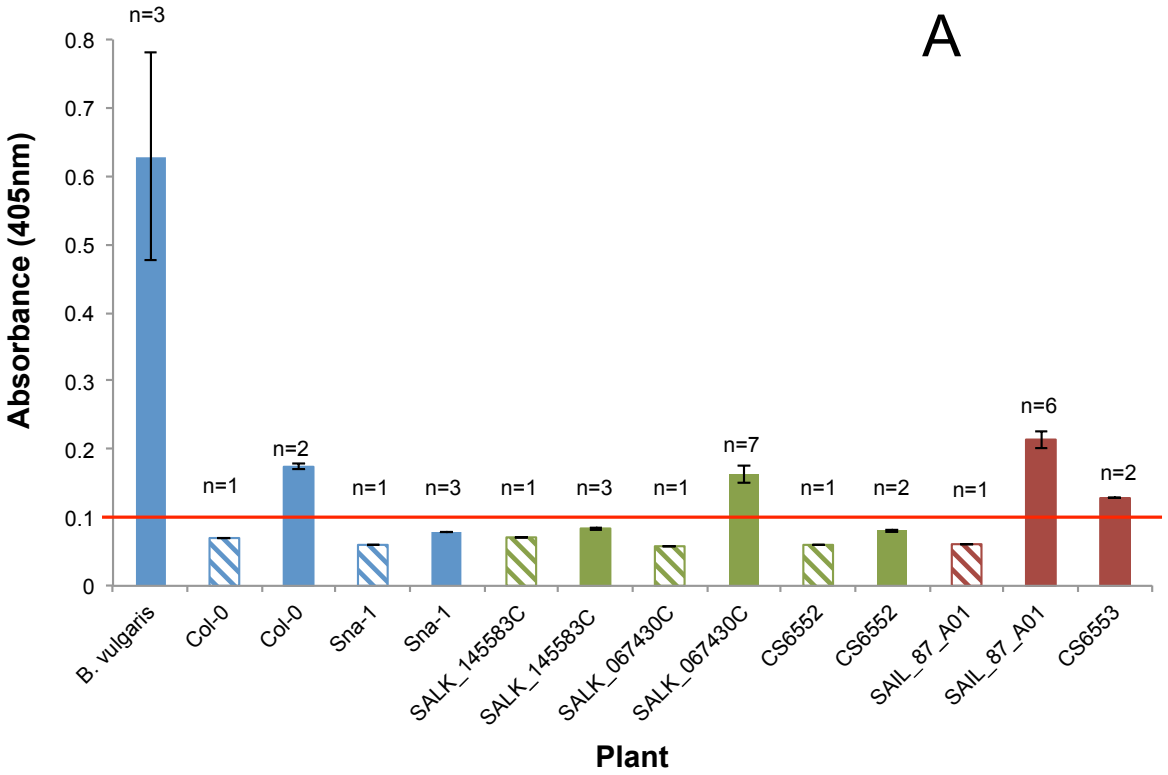
4.4 Further investigation into *eIF4E and eIF4G* T-DNA insertions.

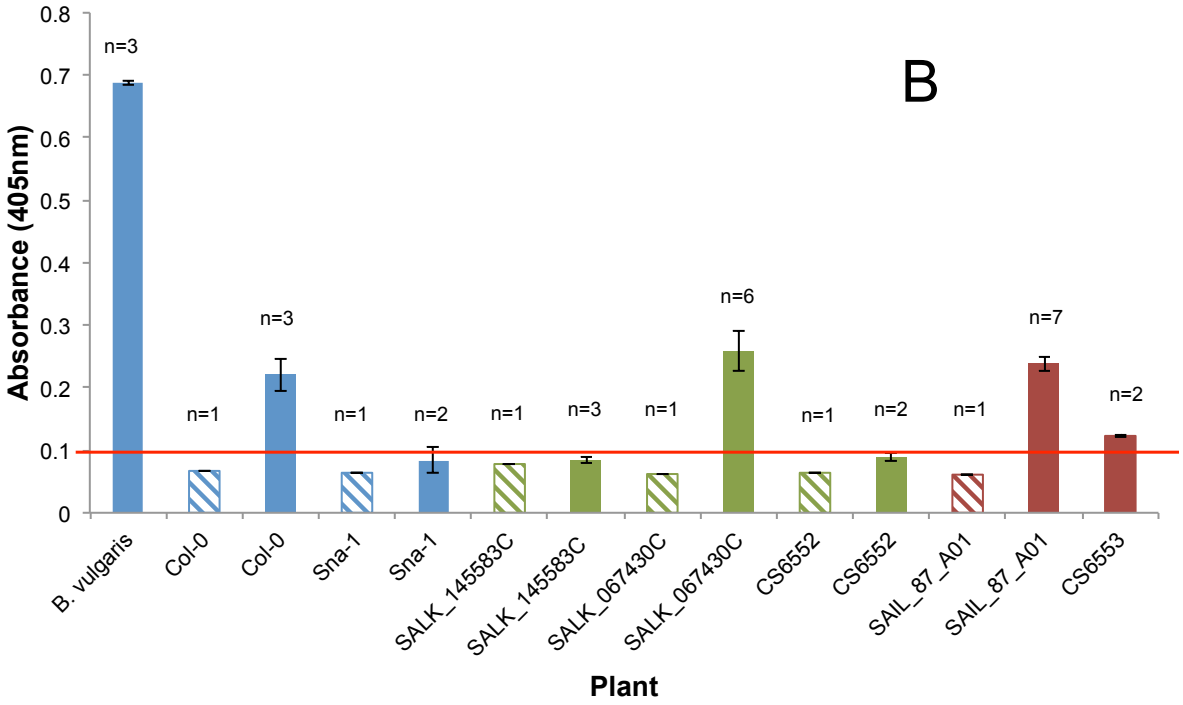
Following results presented in sections 4.2. and 4.3, larger numbers of T-DNA insertion mutations, as well as point mutations in *eIF4E* and *eIF4G* were investigated. Figure 4.2 and 4.4, and Table 4.1 list the plant lines tested.

Table 4.1 Details of ecotypes used in this infection study. Plant lines used for infection with BMYV, the genes in which they occur and the type of mutations they are.

Plant line	Gene	Type of Mutation	Intron/ Exon	Reference
SALK_145583C	*eIF4E*	T-DNA	Intron	Alonso et al., 2003
SALK_067430C	*eIF4E*	T-DNA	Exon	Alonso et al., 2003
CS6552	*eIF4E*	Point mutation	Exon	Yoshii et al., 2004
SAIL_87_A01	*eIF4G*	T-DNA	Exon	McElver et al., 2001
CS6553	*eIF4G*	Point Mutation	Exon	Yoshii et al., 2004

BMYV infection was performed as described (section 2.3) and TAS-ELISA was then carried out, with three repeats of each experiment. The results are shown in Figure 4.6.


A
Absorbance (405nm)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
n=3
n=1
n=2
n=1
n=3
n=1
n=3
n=1
n=7
n=1
n=2
n=1
n=6
n=2
B. vulgaris
Col-0
Col-0
Sna-1
Sna-1
SALK_145583C
SALK_145583C
SALK_067430C
SALK_067430C
CS6552
CS6552
SAIL_87_A01
SAIL_87_A01
CS6553
Plant


B
Absorbance (405nm)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
n=3
n=1
n=3
n=1
n=2
n=1
n=3
n=1
n=6
n=1
n=2
n=1
n=7
n=2
B. vulgaris
Col-0
Col-0
Sna-1
Sna-1
SALK_145583C
SALK_145583C
SALK_067430C
SALK_067430C
CS6552
CS6552
SAIL_87_A01
SAIL_87_A01
CS6553
Plant

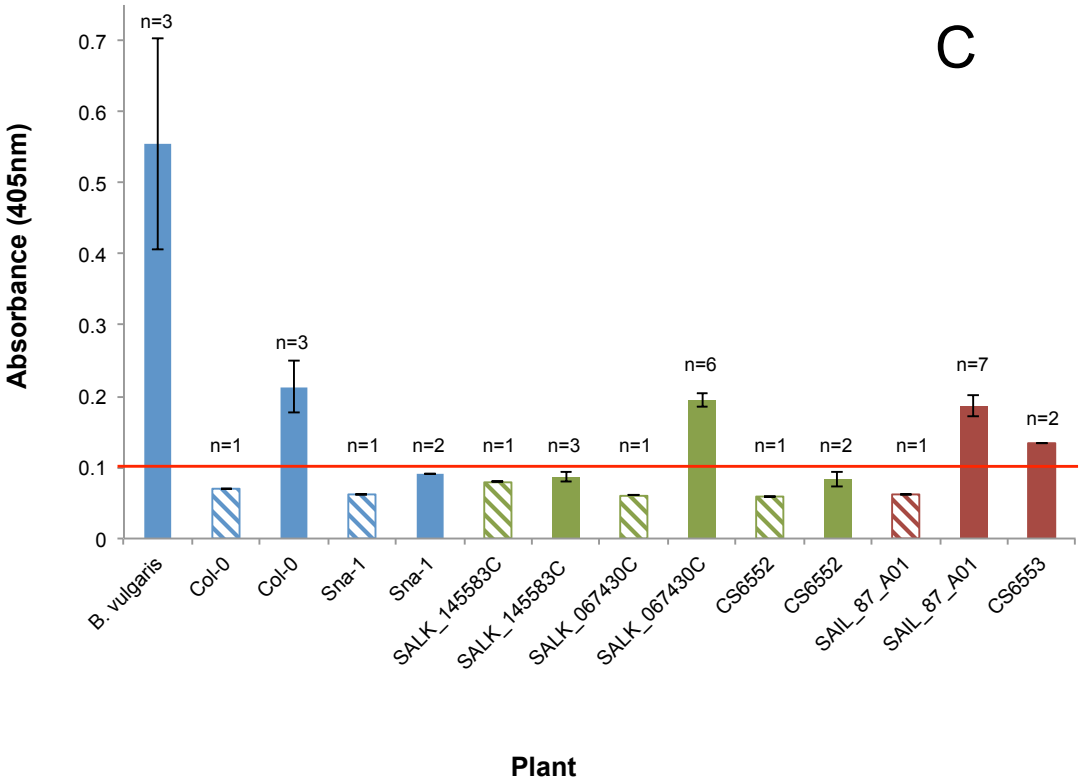


Figure 4.6 *eIF4E and eIF4G* T-DNA insertion and point mutation BMYV infection study TAS-ELISA results. A, B, and C are the three repeats of each experiment. Striped bars indicate uninfected controls, whilst solid bars indicate plants exposed to BMYV viruliferous aphids. The T-DNA and point mutations used can be seen in Table 4.1. The Blue bar denotes control of BMYV infections in *B. vulgaris,* Col-0 and Sna-1. Green bars represent plant lines containing a mutation in *eIF4E,* whilst red shows mutation in *eIF4G.* "n" refers to the total number of plants tested in each replicated experiment, and error bars show SE. The red line shows the computer-derived threshold for resistance to BMYV (above susceptible, below resistant).

Figure 4.6 shows the results of three repeats that are consistent. *B. vulgaris* was used as a positive control in each case, and was the source of viruliferous aphids used to infect the *A. thaliana* lines. Col-0 wild type plants, as shown in previous experiments, were susceptible to BMYV, whilst the Sna-1 ecotype was resistant to the virus. The *eIF4E* T-DNA insert line SALK_145583C was seen in each repeat to have low absorbance readings, and therefore more resistance to the virus, as does the *eIF4E* point mutation line CS6552. These results support the hypothesis that disrupting *eIF4E* function results in greater resistance to BMYV infection. The SALK_067430C line does not fit this same

pattern. In each infection study seen in Figure 4.6, the absorbance readings are similar to that of Col-0, indicating susceptibility to BMYV. Closer inspection of this line showed it has been reported at TAIR that the T-DNA insertion does not eliminate eIF4E function and *eIF4E* mRNA can still be detected (noted by Karen Browning, Caranta Lab, France (2007), www.arabidopsis.org). If this is the case, the line SALK_067430C still contains a functional *eIF4E* and should show similar infection patterns to the background Col-0.

The *eIF4G* T-DNA insertions and point mutation lines also show a consistent pattern across the three repeats. All repeats show a higher absorbance reading than the resistance threshold level, suggesting the plants are susceptible to BMYV. These three mutations give stronger evidence that *eIF4G* is not as important a factor for BMYV infection compared to *eIF4E.* However the *eIF4G* mutant SAIL_87_A01 shows consistently higher absorbance readings than, CS6553 (Figure 4.6).

4.5 Molecular characterisation of *eIF4E* T-DNA inserts using PCR

The results presented in Figure 4.6 suggest that the SALK_067430C line is susceptible to BMYV infection. As the T-DNA insert in this line is in *eIF4E*, this is not the result that was expected, and is not consistent with other data on *eIF4E* mutants. Therefore characterisation of this mutant was performed using PCR.

The lines tested for the presence of a T-DNA insert were the two *eIF4E* T-DNA insertion lines, SALK_145583C, and SALK_067430C. Both T-DNA insertions are reported to contain simple inserts from the pROK2 plasmid. Two sites in *eIF4E* were chosen around the T-DNA sites, and PCR primers were designed. Left and right border T-DNA primers were designed according to the published pROK2 sequence. Figure 4.7 shows a diagram of the PCR amplifications performed. Table 4.2, shows details of the primers used.

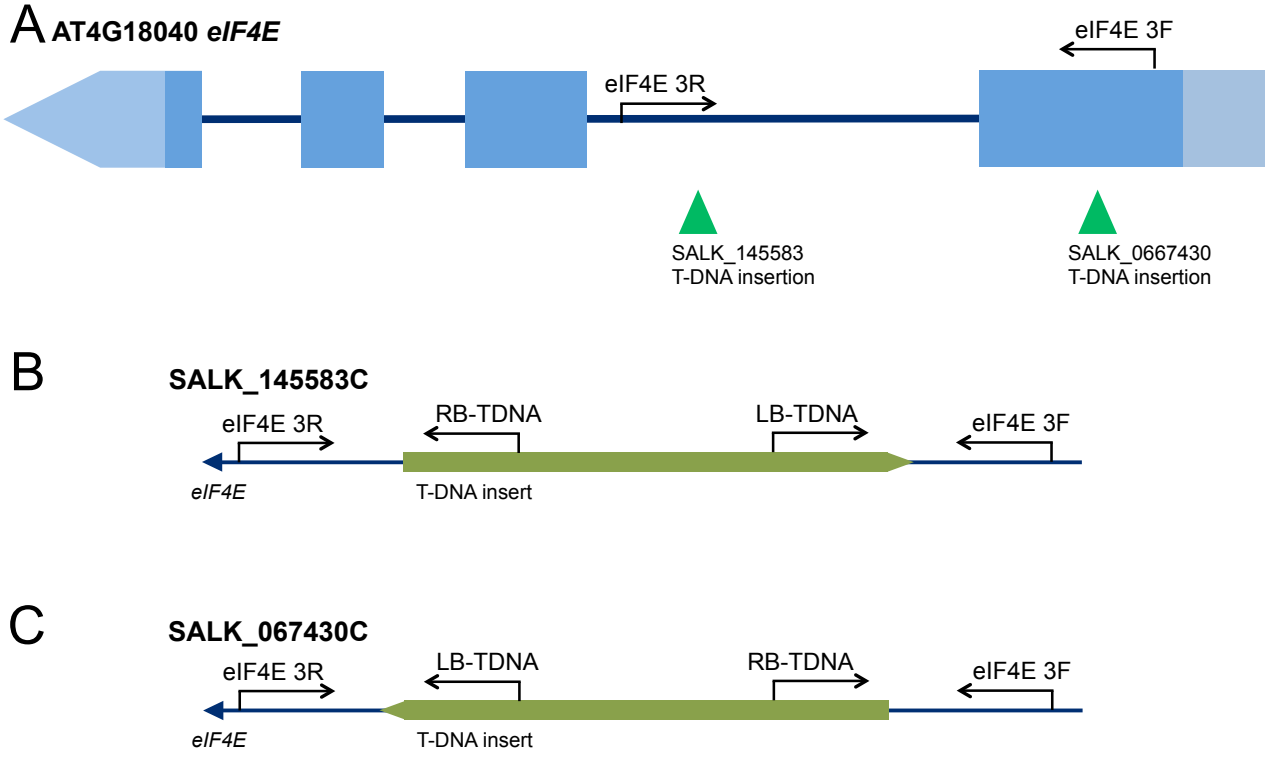


Figure 4.7 *eIF4E* gene T-DNA insertions for PCR analysis. A) *eIF4E* gene diagrams showing introns and exons, with the T-DNA insertion points marked as green arrows. Long arrows in black show the 5'-3' orientation of the primers, their target sites, and the primer name. B) The T-DNA insertion left border (LB) to right border (RB) orientation in the SALK_145583C line. The primers and their target sites are marked on the diagram. C) The T-DNA LB to RB insertion orientation in the SALK_067430C line.

Table 4.2 Details of Primers used in DNA PCR of T-DNA insertions. Diagrams of where primers anneal *eIF4E* and T-DNA sequences can be found in Figure 4.7.

Name of Primer	Sequence of Primer (5'-3')
eIF4E 3R (H66)	CATCAACCCTCAGTCATC
eIF4E 3F (H67)	ACTCCCAAATCTGTTGTAAC
LB-TDNA (H06)	TCCTTTCGCTTTCTTCCCTTCCTTTCTC
RB-TDNA (H07)	GGTTTCTGACGTATGTGCTTAGC

Primers were used in combination as follows, right border (SALK_145583C: eIF4E 3R + RB TDNA, SALK_067430C + Col-0: eIF4E 3F + RB TDNA), left border (SALK_145583C: eIF4E 3F + LB TDNA, SALK_067430C + Col-0: eIF4E 3R + LB TDNA) and the *eIF4E* fragment (All: eIF4E 3R + eIF4E 3F). PCRs were carried out using GoTaq®Flexi (section 2.4.1.2), and the PCR products were run on a 1 % w/v agarose gel (section 2.4.7) If homozygous T-DNA inserts are present, then right border and left border fragments should be visible, but the native *eIF4E* fragment would be absent. If there were no T-DNA inserts present, the only fragment produced would be the *eIF4E* fragment, at around 600 bp. If the T-DNA insert were heterozygous, all three fragments would be visible. Typical results from this PCR are shown in Figure 4.8.

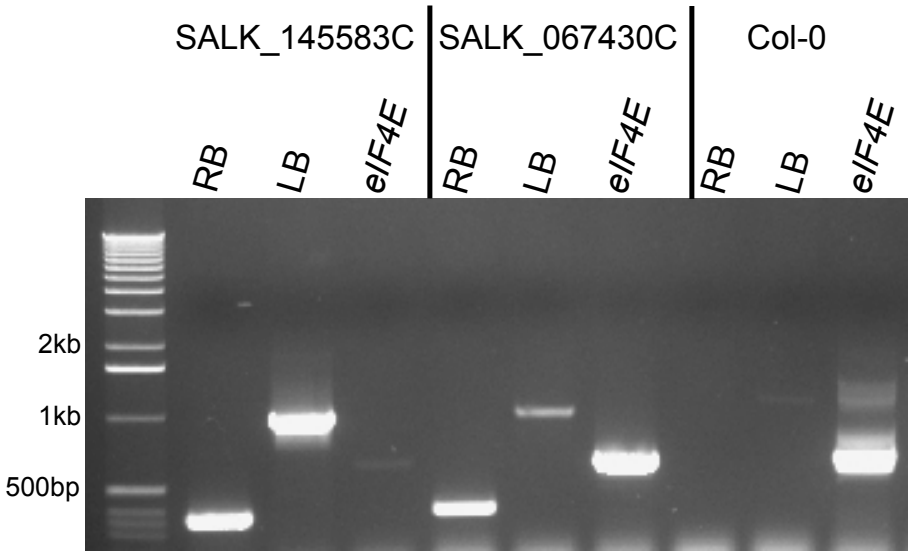


Figure 4.8 PCR results of T-DNA confirmation. A representation of individual plants extracted DNA. PCR was carried out using template DNA from

SALK_145583C, SALK_067430C and Col-0 individuals. Each template DNA was used for three reactions, the right border (RB), the left border (LB) and *eIF4E* fragment (*eIF4E*).

Figure 4.8 shows that when using SALK_145583C template DNA, typically DNA fragments were able to be amplified when using the right and left border primers, but no native *eIF4E* fragment was produced. This suggests that the T-DNA insert is present this line as predicted. The SALK_067430C line shows amplification products from all three sets of PCR indicating that although a T-DNA insert is present, it is heterozygous and not a pure breeding line as reported at TAIR, so *eIF4E* mRNA and protein can still be made. Col-0 shows no PCR products with right and left border sets, only the wild type *eIF4E* fragment. This is as expected, as this indicates no T-DNA insert is present.

4.6 Discussion

The aim of these experiments was to substantiate the claim that eIF4E is an important susceptibility factor for BMYV infection in Arabidopsis. A range of *A. thaliana* T-DNA insertion and mutant lines were obtained from NASC to test their susceptibility to BMYV infection. The background of all of these plants was Colombia, an ecotype previously shown to be susceptible to BMYV. Virus infection was monitored using a TAS-ELISA procedure. This process relies upon the use of antibodies specific to the coat protein of the virus. This technique has limitations, as the results are purely qualitative, indicating resistance or susceptibility through absorbance readings.

Throughout the study it could be seen that the amount of virus present in *B. vulgaris* positive control plants is two- to three-fold higher than is found in susceptible *A. thaliana* plants. This may be due to the constant infection cycle in *B. vulgaris,* as the plant used as a control is the one used as the source of viruliferous aphids used for infection studies. It may also be due to the fact that a low virus titre is required in order to cause infection in *A. thaliana.*
Multiple repeats of the infections were carried out, but because of the instability of virus particles, we cannot compare the results from individual

ELISA plates, and can only compare results from within the same plate. The leaf size of *A. thaliana* is relatively small, and differed between mutant lines. This meant that a core borer could not be used to control the weight and mass of leaf tissue tested. Instead, the leaves harvested were measured by weight, and as close to the same developmental stage as possible. However, Infection studies using TAS-ELISA remain the standard method to investigate virus accumulation in plant tissue (Reinbold et al., 2013). A more satisfactory procedure might be to attempt qPCR analysis of virus RNA.

Infection studies on *eIF4E* T-DNA insertion and point mutation lines showed consistent results. Importantly, the Arabidopsis lines containing a T-DNA insert in *eIF4E* have the same phenotype as their parental background Col-0. Figure 4.6 indicates that the T-DNA insertion line SALK_145583C showed infection levels below the resistance threshold, indicating these plants are now resistant to BMYV. The point mutation line CS6552, where the *eIF4E* Tryp99 is mutated to a translation stop codon, also showed consistently low infection levels suggesting resistance to BMYV. This point mutation in exon 1 of the gene leads to a truncated protein, meaning it is unable to produce functional eIF4E. These results are consistent with other data, from Reinbold et al. (2012).

One other mutant studied however, SALK_067430C, that contains a T-DNA insertion in exon 1 of *eIF4E* did not fit this pattern. The infection levels seen in this mutant line were comparable to that of wild type Col-0, and showed the plant to be infected with BMYV. Further investigation into this line showed that SALK_067430C had previously been reported to be producing *eIF4E* mRNA, and was therefore not a true gene knock out (Karen Browning, Caranta Lab, France 2007). The T-DNA insertion line was then investigated in order to determine whether the T-DNA insert was present. Primers were designed to complement *eIF4E* as well as the T-DNA insert, as shown in Figure 4.7. Along with SALK_067430C, the PCR was carried out on SALK_145583C and Col-0 plants. As the orientation of the LB to RB T-DNA insert differed in each of the mutant lines, slightly different primer combinations were used for the PCR (shown in Figure 4.7). The PCR results indicated that the resistant line

SALK_145583C was homozygous for the T-DNA insertion in *eIF4E.* PCR analysis of SALK_067430C however revealed that both the T-DNA, and the full-length *eIF4E* fragments were produced showing this stock is heterozygous for the T-DNA insertion. Consequently, these plants are still able to produce a functional eIF4E product.

In conclusion, this series of experiments further supports the hypothesis that eIF4E is an important susceptibility factor for BMYV infection of *A. thaliana.* To study the effect of *eIF4G* on infection with BMYV, a T-DNA insertion line, and a point mutation line were tested. Both mutant lines showed the same pattern of susceptibility, and in the case of the T-DNA line, showed similar susceptibility to Col-0 in replicated experiments. The point mutation in line CS6553 showed the least susceptible phenotype. It is not clear why this is the case, and any difference may be due to the fact that there were dissimilar numbers of plants tested to the T-DNA insertion lines. Another possibility is that whilst the T-DNA line disrupts the whole eIF4G protein, the CS6553 line contained a point mutation, which altered amino acid number 1327 from proline to serine. The complete removal of eIF4G gives three possibilities; i) the translation initiation factor complex could not form at all; ii) an isoform of the eIF4G (eIF(iso)4G1 and eIF(iso)4G2) protein was able to substitute functionally; iii) the complex is not wholly dependent on the scaffolding protein eIF4G. There is little evidence to support any of these possibilities, but the fact that BMYV is still able to infect the plants would suggest that the translation initiation complex is still able to form in order to have a fully functioning eIF4E.

Previous studies by Reinbold et al. (2013) have supported the suggestion that BMYV was mildly affected by disruption to *eIF4G,* but similarly to this study, infection levels were above the resistance threshold, meaning the plants are still considered to be infected. The study also stated that the only mutant to significantly reduce the susceptibility of Col-0 to BMYV was a T-DNA insertion or a point mutation in *eIF4E*. This suggests that BMYV is using eIF4E in order to infect the plant, and a knock-out of eIF4G function has little or no effect on eIF4E. The point mutation in *eIF4G* is slightly different however. This mutation causes a change in the amino acid content of the protein, which could have an

influence on protein-protein interactions in the translation initiation complex. If this is the case, then the slightly reduced susceptibility phenotype seen in Figure 4.6 for CS6553 compared to the other *eIF4G* T-DNA mutants, could be due to inefficient complex formation leading to the disruption of eIF4E and the reduced ability for BMYV to infect the plant. Further investigation into both of these infections is required to fully determine the biochemistry of the translation initiation complex and BMYV interaction.

Results from these infection studies suggest that BMYV is using eIF4E, part of the host translation initiation machinery, in order to infect the cell. To prove this, complementation experiments would need to be performed in transgenic plants. This would involve inserting the susceptible Col-0 *eIF4E* allele into resistant plants (containing the Sna-1 allele of *eIF4E*), in order to make them susceptible to BMYV

Chapter 5 *eIF4E* functional complementation

5.1 Introduction

It was shown in Chapter 4 that *eIF4E* is an important factor contributing to BMYV susceptibility in *A. thaliana.* This was shown by infection experiments of *A. thaliana* containing a T-DNA insert in the *eIF4E* gene, which resulted in resistance to BMYV infection. In order to gain further evidence that *eIF4E* plays a role in BMYV infection a complementation test was performed. This functional complementation test aimed to take a BMYV-resistant *A. thaliana* ecotype, (such as Sna-1), and transform it with the *eIF4E* susceptibility allele from Col-0. The hypothesis was that the transgenic plants will be susceptible to BMYV.

A. tumefaciens was first used in experiments to produce genetically engineered tobacco plants in 1983 (Barton et al. 1983; Herrera-Estrella et al. 1983). The technique makes use of Agrobacterium's natural lifecycle. In nature Agrobacterium carries a segment of transfer DNA (T-DNA) on a tumour inducing (Ti) plasmid containing oncogenes and opine catabolism genes. When the T-DNA is transferred into the plant cell, uncontrolled proliferation of the plant cell occurs, alongside production of opines that the bacteria use as a source of nitrogen and carbon (Tzfira and Citovsky, 2006). Agrobacterium infects mostly dicotyledonous plants, and the resulting disease is known as Crown Gall (Chilton et al., 1978; Watson, et al., 1975). The transfer of single stranded T-DNA from the Ti plasmid, followed by stable integration of this gene into the host cell genome is an example of horizontal gene transfer (Lacroix and Citovsky, 2013). Genetic modification of the Ti plasmid has allowed the oncogenes to be replaced with genes of interest, including a selectable marker gene. The process allows specific DNA sequences to be incorporated into the target plant's genome, creating a transgenic plant (Tzfira and Citovsky, 2006).

In order for transformation to occur, specific genes must be present in the Agrobacterium. Genes encoding proteins required for T-DNA production and infection of plant cells are found in the bacterial chromosomal (*chv*) and Ti plasmid virulence (*vir*) genes. As well as proteins encoded for from within the Agrobacterium, infection also requires several host cell proteins for efficient integration of T-DNA into the host cell genome, although this process is not fully understood (Tzfira and Citovsky, 2006).

Initial experiments utilising Agrobacterium DNA transfer encountered problems when working with Ti plasmids, as they were difficult to clone and manipulate using recombinant DNA technology and they generally had a low copy number. In addition, they were not able to replicate in *E. coli*. In order to get around this problem, a binary vector system was developed (Hoekema et al., 1984; Lee and Gelvin, 2008). It was realised that the T-DNA gene fragment did not need to be on the same plasmid as the *vir* genes, as long as both were present within the same bacterial cell. This allowed the creation of two smaller plasmids, one the T-DNA binary vector containing selectable markers, antibiotic resistance, the gene of interest and T-DNA border repeats. The other plasmid, a *vir* helper plasmid, contained the essential *vir* genes required for T-DNA synthesis and transfer to plant cells. The smaller T-DNA binary vector could therefore be modified to aid cloning. This involved the integration of a poly-cloning site and the ability to construct recombinant plasmids in *E. coli*. As well as this, oncogenes and opine synthase genes were removed. Many T-DNA binary vectors have now been created with various origins of replication (*ori*) that can be used in both *E. coli* and *A. tumefaciens*, and can contain a variety of bacterial and plant selectable markers (Lee and Gelvin, 2008).

Transformation begins with bacterium-plant attachment. Plant wound signals, such as acetosyringone trigger the expression of *vir* genes (Godwin et al., 1991). The acetosyringone interacts with the membrane bound VirA protein, which in turn activates VirG. VirG activates transcription of other important proteins such as VirD1 and VirD2 (Turk et al., 1994). VirD1 and VirD2 produce a single stranded T-DNA molecule with the VirD2 molecule covalently

attached to the 5' end (Filichkin and Gelvin, 1993; Tzfira and Citovsky, 2006). This molecule is then exported along with other Vir proteins through a VirB/D4 type IV secretion system into the cytoplasm of the host cell (Christie, 2004). Upon arrival in the host cell, the T-DNA is coated in VirE2 for protection, and to enable its translocation to the host nucleus (Abu-Arish et al., 2004; Citovsky et al., 1989). Traveling through the host cell cytoplasm, nuclear import, T-DNA un-coating and finally T-DNA integration into the host genome are all carried out utilising the host cell machinery (Tzfira and Citovsky, 2006). This process is shown in Figure 5.1.

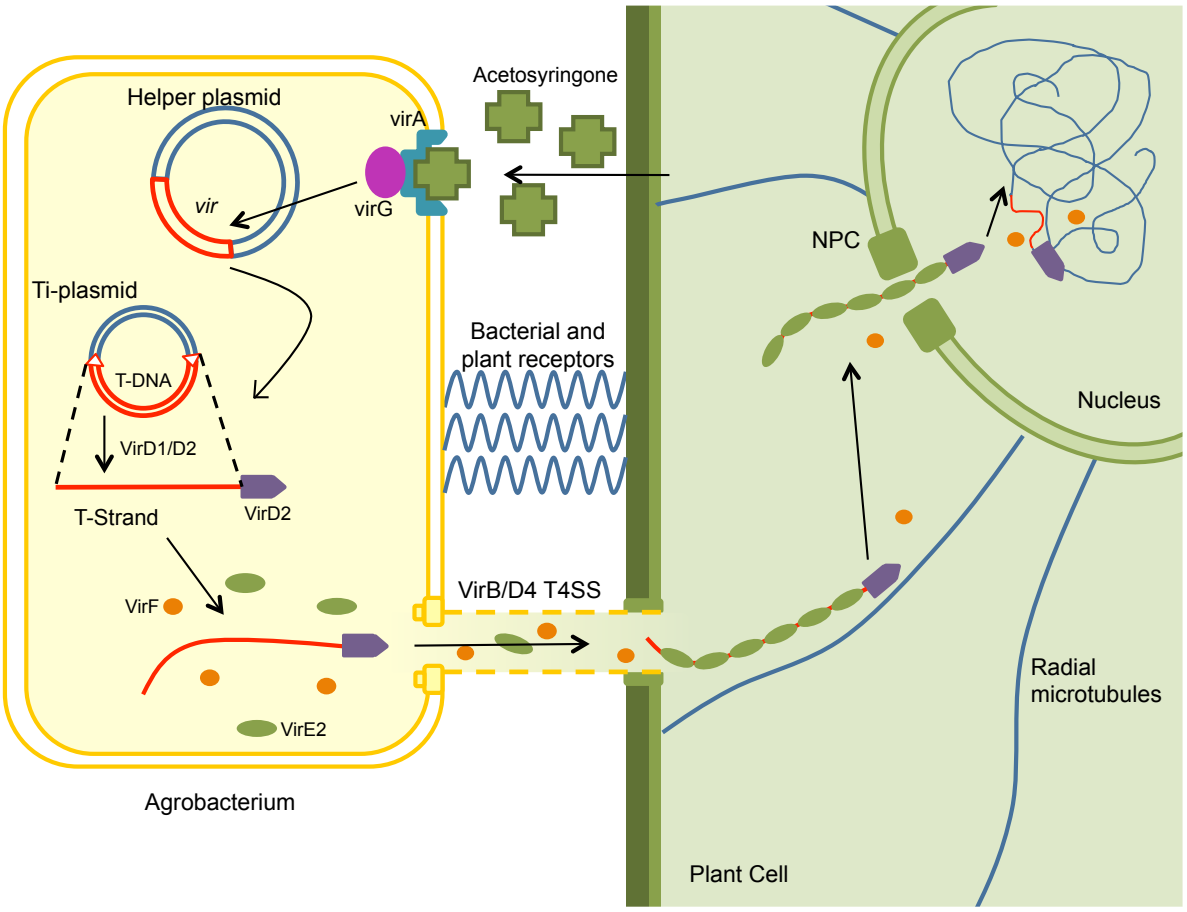


Figure 5.1 Transfer of Agrobacterium T-DNA into a plant cell host. The Agrobacterum cell recognises a plant cell and *vice versa* through membrane bound receptors. The plant cell releases wound signals such as acetosyringone, which are recognised by the Agrobacterium VirA membrane spanning protein. VirA activates VirG, which in turn initiates the translation of the other *vir* genes. VirD1 and VirD2 replicate the T-DNA region of the Ti-plasmid and create a single stranded DNA molecule (the T-DNA) and VirD2 attaches to the 5' end to guide the T-DNA to the VirB/VirD4 type 4 secretion system (T4SS). VirE2 and VirF accompany the T-DNA through the T4SS, where upon arrival in the plant cell, VirE2 coats the T-DNA in order to protect it. The T-DNA makes its way to the nucleus using radial microtubules for guidance, and enters the nucleus through nuclear pore channels (NPC). The T-DNA is then stripped of the VirE2 proteins and integrated into the host cell genome (Adapted from Tzfira and Citovsky, 2006).

Initial transformations of tobacco plants with Agrobacterium (Barton et al 1983; Herrera-Estrella, et al 1983) was found to work well, but limitations of host species greatly held back its use in other crop plants. Other methods of plant

transformation soon developed, many requiring complicated preparation procedures of cells and tissues, and using particle bombardment to transform them. These methods required a wealth of technical abilities and expensive laboratory equipment for successful transformation of cell protoplasts, and consequent regeneration of plants from these cells (Christou, 1996; Clough and Bent, 1998). These techniques can also lead to extensive somaclonal variation (undesired DNA modifications in the clonal progeny of single parent plant clones) often induced by stress (Labra et al., 2004; Larkin and Scowcroft, 1981). The development of a simple method for plant transformation was required for the progression of plant molecular sciences. A technique that could widely be used by both plant cell biologists and molecular biologists, requiring less expertise and expensive equipment would greatly advance the field of study.

The “Agrobacterium vacuum infiltration” approach was established by Betchtold (1993). This simpler and much more reliable technique for plant transformation built upon previous methods of root tissue culture followed by plant regeneration. The method used uprooted and flowering *A. thaliana*, which were vacuum infiltrated with Agrobacterium. This whole-plant transformation method resulted in progeny being grown on selective media (usually antibiotic media, whose resistance was carried on the *A. tumefaceiens* Ti plasmid) to identify successful transformants (Clough and Bent, 1998). Other benefits to the Agrobacterium vacuum infiltration method include the high transformation efficiency of plants. Large numbers of transgenic progeny are able to be collected, and the amount of somaclonal variation is minimal (Clough and Bent, 1998; Labra et al., 2004). The main drawback to this method was that it only seemed successful when transforming *A. thaliana.*

The “Floral dip” method was developed as a simplified modification of the “Agrobacterium vacuum infiltration” method. Previous methods had required vacuum infiltration of the entire *A. thaliana* plant, whereas the “Floral dip” method limits the tissues being transformed, to the flowers (Bent, 2006; Clough and Bent, 1998). The method involves immersing the flowering parts

of the plant into Agrobacteria suspended in media. The media contains sucrose or glucose, and Silwet L-77, which allows the surface tension of the liquid to be reduced, permitting more efficient transformaiton. Transformations using this method occur at a relatively high rate with between 0.5-3% of progeny seed being reportedly transformed (Bent, 2006; Clough and Bent, 1998). Further investigation into this method found that the female reproductive tissues were the targets for flower dip transformations (Bechtold et al., 2000; Desfeux et al., 2000; Ye et al., 1999). Many of these experiments were confirmed to be successful transformations of *A. thaliana* with the use of *gusA* (found on the Ti plasmid), encoding β-glucuronidase (GUS). GUS staining identified that the female parts of the plant were the most important for inheritable transformations, as the areas of the plant that were GUS stained were the ovules, with very little staining of the pollen (Bechtold et al., 2000; Desfeux et al., 2000; Ye et al., 1999). The timing of infection with Agrobacterium during flower development was also found to be important. During Arabidopsis flower growth, an open vase-like structure called the gynoecium forms, containing the developing ovules. Three days prior to anthesis (flowering of the plant) a stigmatic cap forms over the gynoecium, effectively sealing it. Inoculation with Agrobacterium five days prior to anthesis was found to give the highest transformation rates. This was further supported by the Arabidopsis CRABS-CLAW mutant which maintains an open gynoecium throughout flower development giving a six times greater transformation rate over other Arabidopsis ecotypes (Desfeux et al., 2000). Understanding how, and where, the Agrobacterium are transforming the flower tissues gives significant insight into how this technology can be modified to allow other plants, such as crop plants, to be able to be genetically modified in a similar, simple and effective way.

The ability to modify plant DNA has provided molecular biologists with many more opportunities to understand the genetics of plants. As previously discussed in Chapter 3, large T-DNA libraries have been created. These libraries allow the clarification of many gene functions by knock out studies. Whilst this is a useful technique, the addition of genes to plants is also useful. Functional complementation experiments also allow an understanding of

specific gene function through gain-of-function rather that loss-of-function analyses. This transformation, often using Agrobacterium as a vector can convert a mutated plant to a wild type one through the integration of a wild type gene.

In these experiments the BMYV susceptible *eIF4E* allele from Col-0 was cloned into a Ti plasmid and transformed into Agrobacterium. Plants containing the resistant Sna-1 allele (plant line JIC62) were transformed using Agrobacterium. The progeny of transformed plants were tested for susceptibility to BMYV by TAS-ELISA. Plants that were successfully transformed with the Col-0 *eIF4E* were shown to be susceptible to infection with BMYV.

5.2 Cloning of Col-0 *eIF4E*

Cloning of Col-0 *eIF4E* was carried out in several stages. The gene was amplified from Col-0 genomic DNA and initially cloned into the versatile vector pUC19. This decision was made because the plasmid has a high copy number in *E. coli*, allowing a more efficient method to identify successful clones. Col-0 *eIF4E* was cloned using the In-Fusion® HD cloning kit (Clontech, section 2.2.4). This method uses a PCR based approach to cloning, instead of relying upon ligation reactions, and was adopted because conventional ligation of restriction enzyme digested DNAs proved unsuccessful. Primers to carry out In-Fusion® cloning were designed to clone the *eIF4E* in two halves, named 5' and 3', according to their position in the gene. Primers were also designed to overlap the pUC19 vector in the multi-cloning region, which also contained approximately 15 bp overlap with the relevant *eIF4E* gene fragment. A diagram of primer design can be seen in Figure 5.2, and the corresponding primer sequences in Table 5.1.

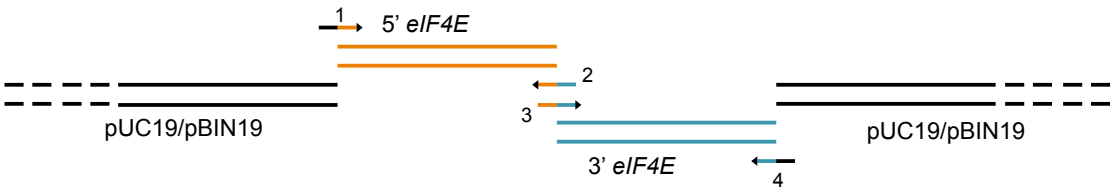


Figure 5.2 In-Fusion ® cloning primer design. Primers were designed for each overlapping fragment, labelled 1-4. Primers contained around 15bp of overlap to the corresponding *eIF4E* DNA sequence. The dotted lines of pUC19 indicate that the plasmid joins in a loop. The pUC19 and pBIN19 multi-cloning regions have the same sequence, so primers can be used with either plasmid. Sequences of primers can be seen in Table 5.1.

Table 5.1 Primer sequences for In-Fusion cloning of Col-0 *eIF4E* into pUc19. Primer sequences are colour coordinated in the same pattern as Figure 5.1. This is to allow identification of the overlap sequences to either *eIF4E* fragments, or plasmid sequence.

Number reference	Primer Name	Sequence 5'-3'
1	5' *eIF4E* IF F (J17)	TTCGAGCTCGGTACCAGTGGTCCTTTCAGACAGTT
2	5' *eIF4E* IF R (J19)	AAGTGGGAGGATCCTATTTGT
3	3' *eIF4E* IF F (J20)	ACAAATAGGATCCTCCCACTT
4	3' *eIF4E* IF R (J18)	GGATCCCGGGTACCAGTTACTAGTGAGTAGTGATGACA

PCR products were amplified using Phusion® High Fidelity DNA polymerase (ThermoFisher Scientific, methods section 2.4.1.3) and run on a 1% w/v agarose gel. Amplified *eIF4E* fragments can be seen in Figure 5.3.

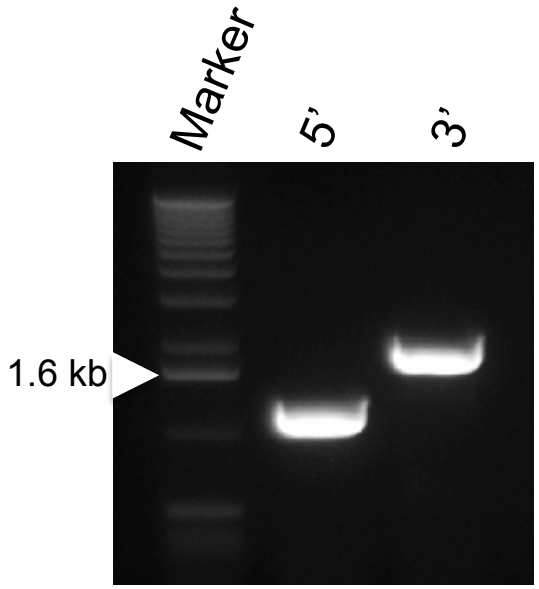


Figure 5.3 PCR amplified 5' and 3' Col-0 *eIF4E* DNA fragments. DNA was amplified using Phusion® high fidelity DNA polymerase, and primers shown in Table 5.1.

Figure 5.2 shows successful amplification of the *eIF4E* fragments, allowing cloning into pUC19 to take place (Section 2.2.4). The pUC19 vector was linearised using the *KpnI* restriction enzyme, and In-Fusion® cloning carried out with *eIF4E* as shown in Figure 5.1. The reaction mixture was then transformed directly into XL2-blue ultracompetent cells (Stratagene) as described in section 2.1.1.2. Transformations were plated on LB agar containing 50 µg.ml of the selective antibiotic Kanamycin. Six transformed colonies were randomly chosen and grown in liquid LB, overnight, and the plasmids were purified from these cultures (section 2.1.1.3).

In order to determine if the plasmids contained *eIF4E*, a series of restriction digestions were performed. The restriction enzyme *KpnI* was used to remove *eIF4E* from pUC19. As both these products have predicted sizes of 3kb, another restriction digest was performed with *KpnI* and *BamHI*. Col-0 *eIF4E* contains one *BamHI* site where the 5' and 3' amplified products join creating a 1.1 kb 5' *eIF4E* fragment, and a 1.9 kb 3' *eIF4E* fragment. As the remaining pUC19 does not contain a *BamHI* site, this fragment should remain at 3 kb.

The six digested fragments were analysed on a 1 % w/v agarose gel (Figure 5.4).

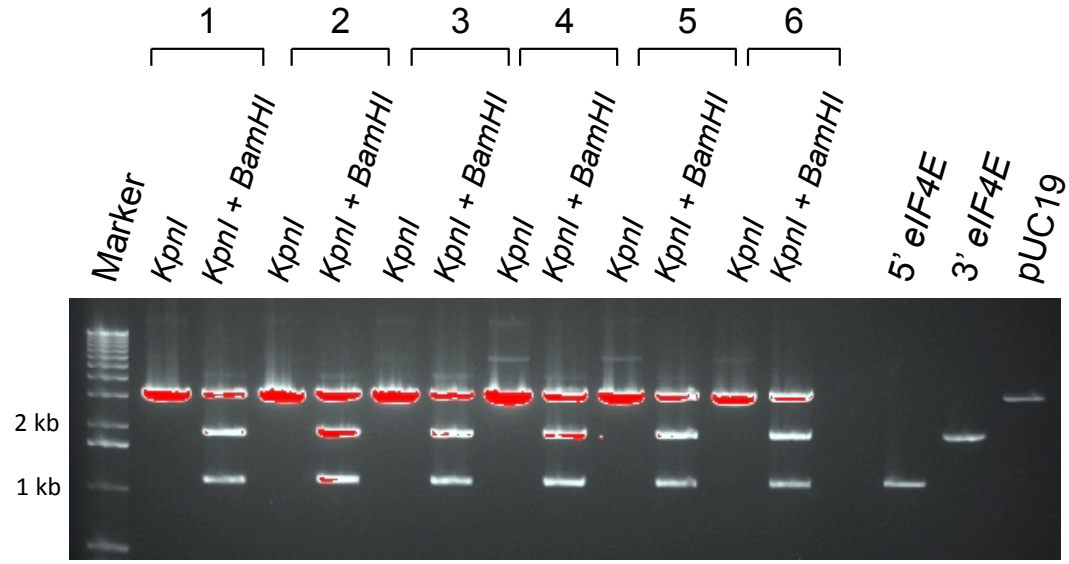


Figure 5.4 Restriction digest of pUC19 containing Col-0 *eIF4E.* The figure shows six transformed colony plasmid preparations digested with either *KpnI* or *KpnI* and *BamHI* restriction enzymes. Predicted sizes for fragments are as follows; puC19 3kb, full *eIF4E* 3kb, 5' *eIF4E* 1.1 kb, 3' *eIF4E* 1.9 kb. Fragments of *eIF4E* (amplified from Col-0 DNA) and pUC19 digested with *BamHI* and *KpnI* are also shown.

Figure 5.4 indicates that the cloning has been successful, and the pUC19 plasmid contains the predicted insert. Restriction digestion patterns suggest that the inserted fragment is *eIF4E,* but to be sure that the sequence is identical to that of Col-0 *eIF4E*, plasmid number 1 was sequenced. Sequencing returned a 100 % identity and no gaps when compared with the published Col-0 *eIF4E* sequence (NCBI).

5.3 Construction of *eIF4E* in pBIN19

In order to carry out complementation analysis, Col-0 *eIF4E* must be contained on a Ti plasmid. As well as containing an antibiotic reporter gene, the T-DNA also contains a polylinker region to clone in target genes for plant expression. A commonly used Ti plasmid is pBIN19 (Bevan and Lane, 1984; Lee and Gelvin, 2008) This plasmid is extremely useful as it can be

maintained in both *E. coli* and *A. tumefaciens*. In-Fusion® cloning of *eIF4E* into pBIN19 was not successful, so instead a ligation method was used.

Using the sequenced pUC19+*eIF4E* as a template, universal forward and reverse primers (M13 and M14, sequence shown in Table 5.2) were used to amplify Col-0 eIF4E from a pUC19 clone using the Phusion® High Fidelity DNA polymerase (ThermoFisher Scientific). Amplification products are shown in Figure 5.5. The expected size is 3 Kb.

Table 5.2 PCR primers used in amplification of *eIF4E*. Forward and reverse primers used to amplify cloned Col-0 *eIF4E* from the sequenced template pUC19+*eIF4E*.

Primer name	Primer Sequence (5'-3')
M13 F (A11)	GTAAAACGACGGCCAGT
M13 R (A12)	GGAAACAGCTATGACCATG

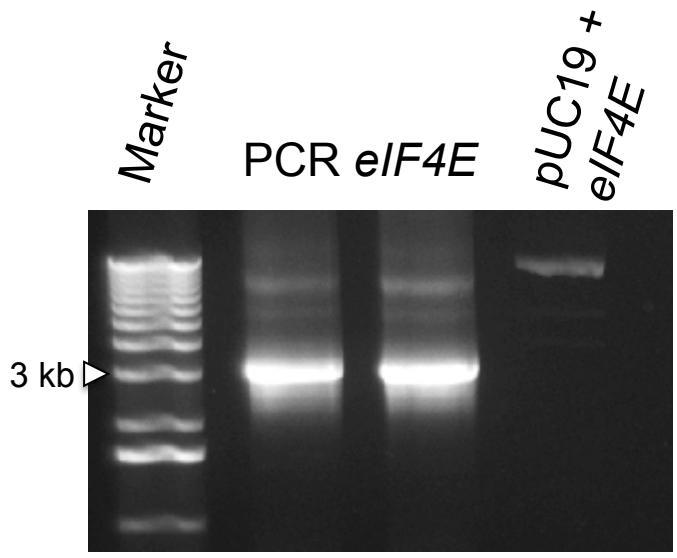


Figure 5.5 PCR amplification of *eIF4E* from pUC19+*eIF4E* template (also shown in lane 4). The predicted Col-0 *eIF4E* fragment size is 3 kb, as expected.

The PCR product was then digested with the restriction enzymes *SstI* and *XbaI,* as was the plasmid pBIN19 (*SstI* and *XbaI* sites were incorporated into

the *eIF4E* fragment during amplification with universal primers from pUC19). Digested DNA products can be seen in Figure 5.6.

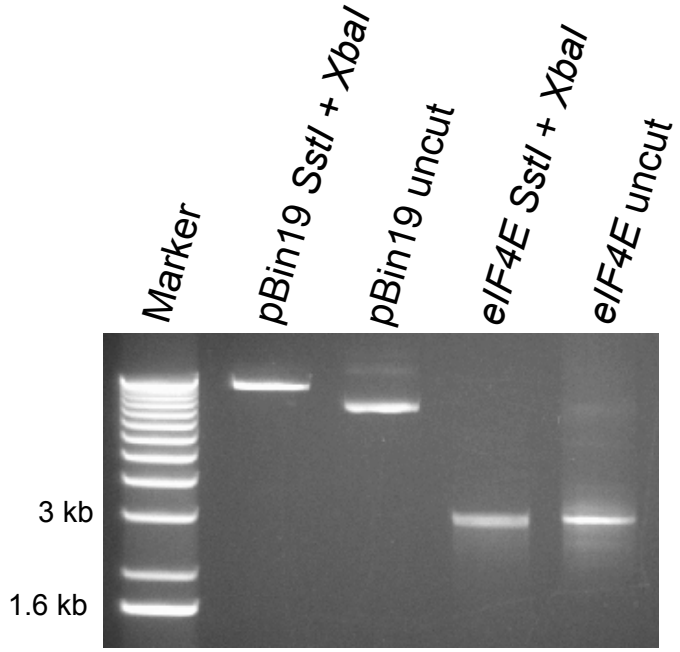


Figure 5.6 Digested pBin19 and *eIF4E* DNA products for ligation. DNA products can be seen after digestion with *SstI* and *XbaI,* and also without digestion. The pBin19 plasmid appears to linearise as shown by the predicted fragment of 11.7 kb. No change in size of the Col-0 *eIF4E* fragment was observed, confirming that there are no internal restriction sites.

Digested DNA products (pBin19 and *eIF4E*) were ligated as described in section 2.4.6, and transformed into XL-2 Blue ultracompetent *E. coli* cells (section 2.1.1.2). Transformed cells were grown on LB agar containing 50 µg/ml kanamycin. Colonies were picked from these plates and grown in liquid LB medium, and the plasmids purified. Restriction digests were performed using *KpnI* and *KpnI*+*BamHI* in order to identify the pBIN19 plasmid, and the 3' and 5' *eIF4E* fragments as previously described as seen in Figure 5.4. As pBIN19 is a low copy number plasmid PCR was also carried out to confirm the presence of *eIF4E* using the *eIF4E* sequencing primers *reverse 2*, and *forward 3* (listed in Table 3.3A), giving an expected product size of 1.1 kb. Figure 5.7 shows both the restriction digest (A) and the PCR amplification using GoTaq® DNA polymerase (B) of pBin19+*eIF4E.*

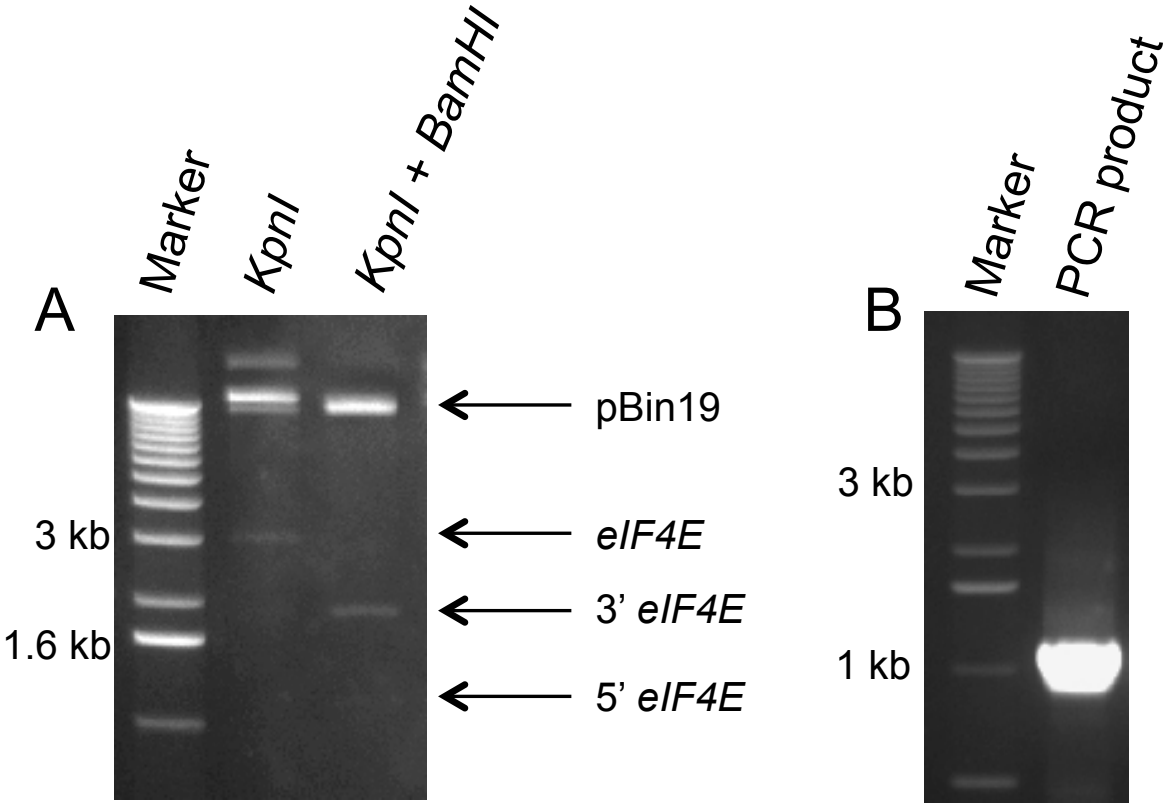


Figure 5.7 Analysis of pBIN19+*eIF4E*. A) Restriction digest of pBin19+*eIF4E* with *KpnI,* and *KpnI+BamHI*. The low copy number of the plasmid means that the DNA is more difficult to see. pBin19 (11.7 kb) can be seen in both lanes, whilst digestion with *KpnI* alone releases a fragment of 3 kb, consistent with the size of *eIF4E.* Digestion with *KpnI* and *BamHI* releases two smaller fragments of around 1.1kb and 1.9 kb, consistent with the sizes of 5' and 3' *eIF4E* respectively. B) PCR amplification product at around 1 kb, consistent with the predicted amplification size fragment when using sequencing primers *forward 2* and *reverse 3* (Table 3.3A).

These results strongly suggest that pBIN19 now contains the cloned Col-0 *eIF4E* gene, and therefore this plasmid could be electro-transformed into AgII *A. tumefaciens* cells (see sections 2.1.2.2 and 2.1.2.3).

5.4 pBIN19+*eIF4E* electro-transformation into *A. tumefaciens*

Transformed *A. tumefaciens* AgII were plated onto selective minimal media (section 2.1.2.5). The media contained 50 µg/ml Kanamycin to select for the pBIN19 plasmid as well as 50 µg/ml Rifampicin for selection of the Ti helper plasmid. This ensures that cells growing on the media are AgII cells containing pBIN19. A negative control was also carried out, where water was used

instead of pBIN19. Thousands of colonies grew on the plate transformed with pBIN19, whilst no colonies were recovered from the negative control (results not shown). Fourteen colonies were picked and inoculated into liquid LB medium containing 50 µg/ml Kanamycin and Rifampicin, and grown at 28 °C.

Colony PCR using the enzyme GoTaq Flexi® was then carried out using the *eIF4E forward 2* and *reverse 3* sequencing primers (Table 3.3A). Fourteen colonies were chosen because of the low copy number of the plasmid to increase the ability to identify suitable transformants. Amplified fragments can be seen in Figure 5.8.

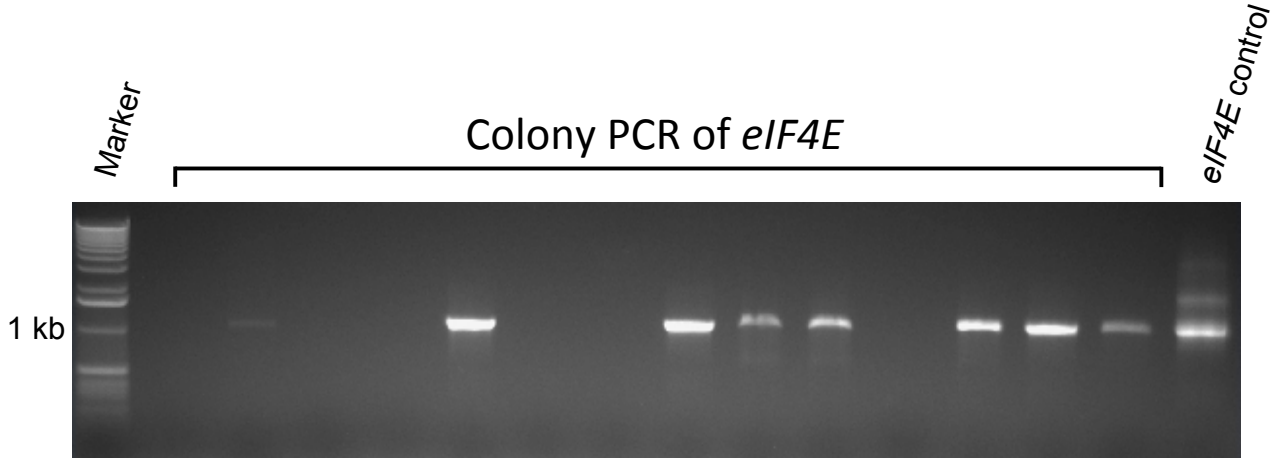


Figure 5.8 Amplification of *eIF4E* from pBin19 purified from AgII. The image shows 14 PCR amplification reactions from independent transformants and an *eIF4E* control reaction (amplified from Col-0 DNA template). The results show variation for the presence/absence of the eIF4E target sequence. Clones containing an *eIF4E* fragment of the predicted size (c. 1.1 kb) are visible in eight of the 14 reactions.

Figure 5.8 suggests that successful transformation of AgII with pBIN19+*eIF4E* has occurred and these cultures were used to transform *A. thaliana.*

5.5 Transformation of *A. thaliana*

With the assistance of Matthew Smoker and Jodie Pyke (*The Sainsbury Laboratory*, Norwich) the AgII pBIN19+*eIF4E* cultures were grown and transformed into an *A. thaliana.* Two transformations were carried out, one in

Col-0 (as a control) and one in the JIC62 line. The JIC62 line was created from an individual from the F_2 generation of a Col-0 x Sna-1 cross created by Percival-Alwyn (2010). This stock was created because of the difficulty when working with Sna-1, which requires extensive stratification in order to induce bolting and flowering of the plant. In order to carry out floral dip transformations, flowering plants were required, so the JIC62 line was much more suitable than the Sna-1 ecotype in this experiment. The JIC62 line was an early bolting line that has been confirmed by CAPS analysis to be homozygous for the Sna-1 version of the *eIF4E* allele, causing it to be resistant to BMYV (Percival-Alwyn, 2010).

The two Arabidopsis lines (Col-0 and JIC62) were transformed using the floral dip method (section 2.2.7) Seed was collected from these plants and gas sterilised (section 2.2.2) before being spread on GM media plates containing 50 mg.L kanamycin (section 2.2.2). Seeds were stratified for two days at 4 °C before being transferred to growth cabinets at 24°C, 16 hours light, until transformed plants could be observed.

Plants successfully transformed with pBIN19 should contain the kanamycin resistance gene, and would be able to grow on GM media plates containing kanamycin. If the plants were not transformed they would initially develop but then appear bleached and not survive more than two weeks. Example images of plates can be seen in Figure 5.9.

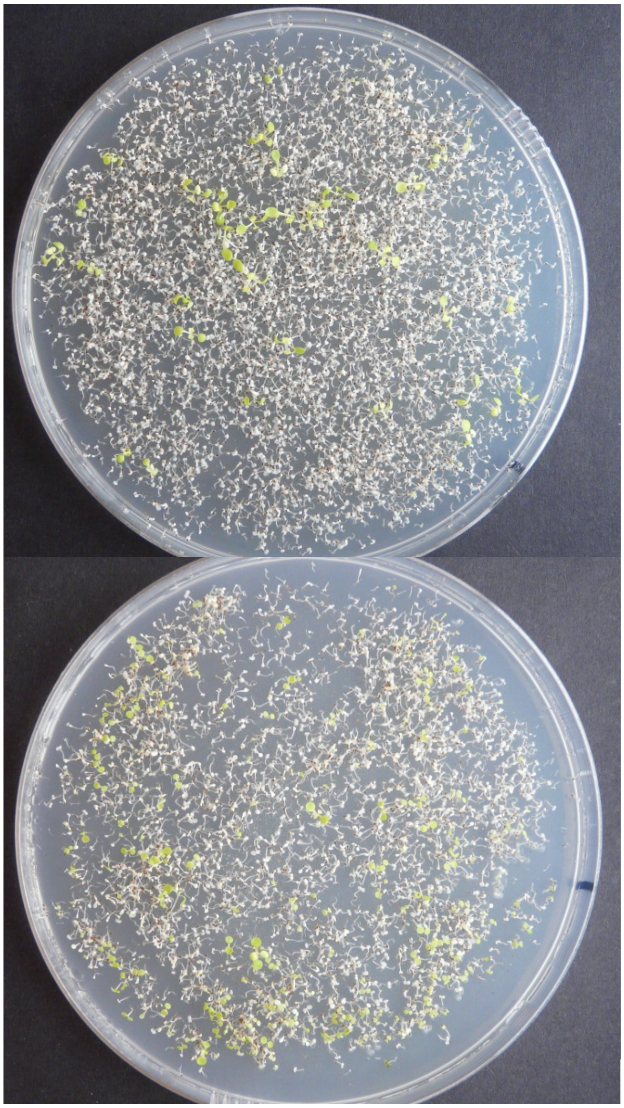


Figure 5.9 Transformed *A. thaliana* lines Col-0 and JIC62 Seeds collected from *A. thaliana* transformed with AgII pBin19+*eIF4E*. Top are Col-0 plants, bottom BMYV resistant JIC62. Green plants are Kanamycin resistant, and transformed with AgII, whilst white etiolated plants are susceptible to Kanamycin and did not survive.

In total 10 plates (around 4000 seeds per plate) from transformed Col-0 and 11 plates of seed (around 4000 seeds per plate) from JIC62 were analysed. Green plants from these plates were transferred into soil and seed was collected. During this time, DNA extractions were carried out on the re-potted plants that had survived. The surviving plants were given numbers 1-40, and CAPS analysis was performed to find out if any contained the Col-0 *eIF4E* allele (all should contain the Sna-1 *eIF4E* allele). CAPS analysis with the

restriction enzyme *BspHI* was used on PCR amplification products generated with *eIF4E* sequencing primer set 1 (Table 3.3A). Figure 5.10 shows the CAPS analysis results.

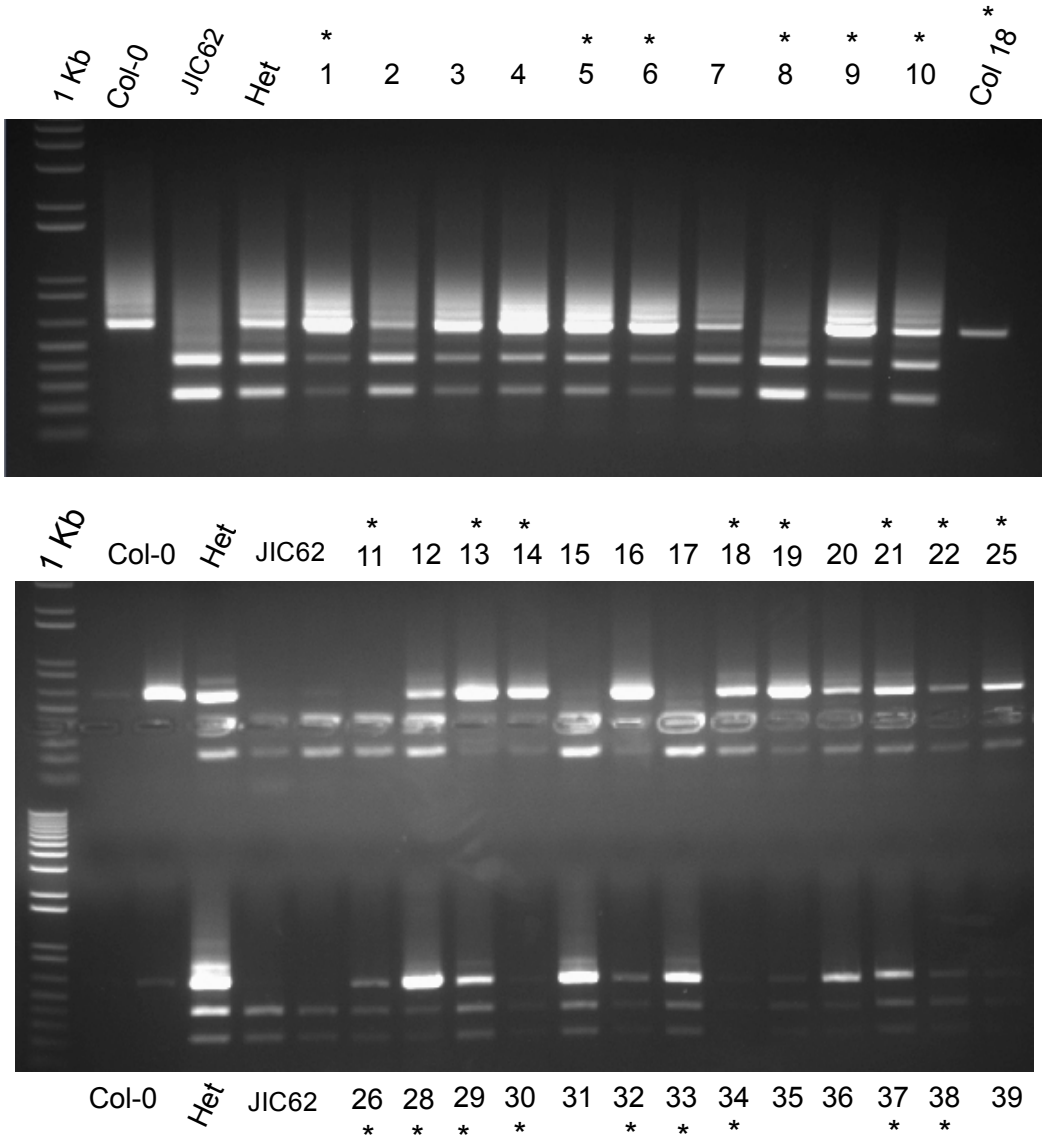


Figure 5.10 CAPS analysis of flower-dip inoculated JIC62. DNA extracted from inoculated plants was analysed for the presence of the *eIF4E* Col-0 allele. Col-0 and JIC62 DNA was used for controls (JIC62 is a line homozygous with Sna-1 *eIF4E*). Het = plants known to be Sna-1xCol-0, and therefore heterozygous for Col-0 and Sna-1 *eIF4E.* Plant numbers marked with '*' were used to generate seed. All plants are from the JIC62 flower-dip inoculation, except Col-0 18, which is a Col-0 plant that was flower dip inoculated as a control.

CAPS analysis results showed varying stoichiometry. In some cases the *eIF4E* stoichiometry is not 1:1 Col-0 version of the allele: Sna-1 version of the allele. This is because multiple insertion events of the Col-0 version of *eIF4E* may have occurred. In some instances, for example number eight, did not contain the Col-0 *eIF4E* allele at all. Figure 1.10 shows plants which were used in following experiments marked with a '*'. Most of these plants were selected because they showed an approximate 1:1 ratio (Col-0 *eIF4E* : Sna-1 *eIF4E*) rather than multiple insertions which might complicate further genetic characterisation. Other plants were chosen to act as controls (Col 18 and number 8). Seed was collected from the transformants marked '*', were sterilised (section 2.2.2) and plated onto GM media containing kanamycin, in order to select for the presence of the Col-0 *eIF4E* insert. If the seeds, and therefore plants, had successfully inherited the Col-0 *eIF4E* gene, they should segregate on the plate 3:1 (kanamycin resistant:kanamycin susceptible). Plants were stratified at 4 °C for two nights and then grown in growth cabinets at 24 °C, 16 hours light, until the segregation could be seen. The number of green and white plants were counted and a χ^2 analysis performed shown in Table 5.3 to show goodness of fit to the 3:1 ratio.

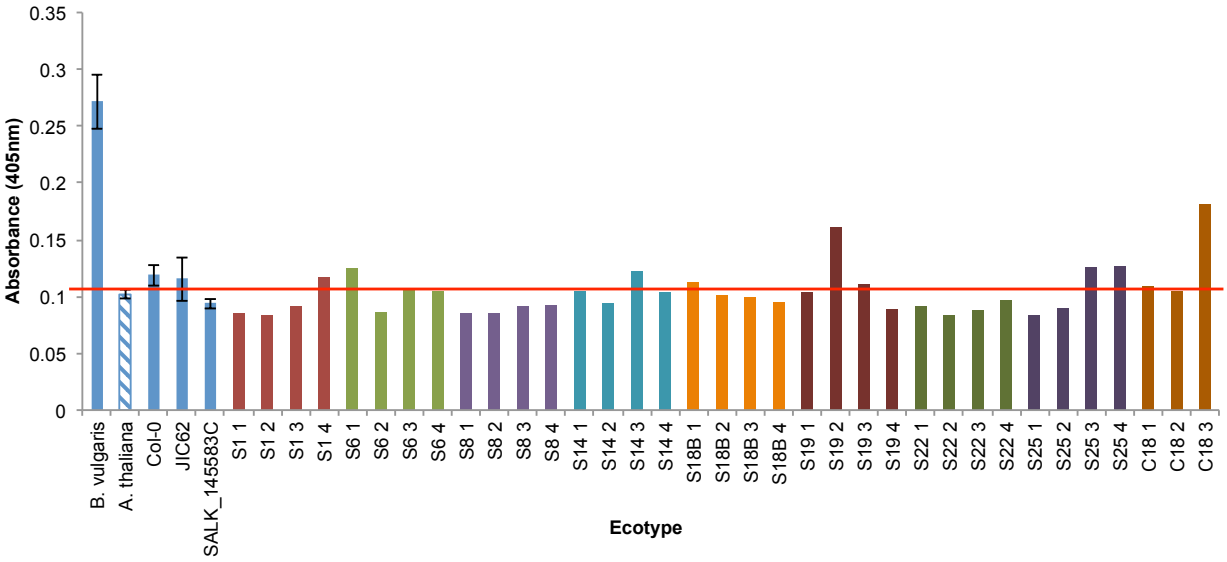
Table 5.3 χ^2 analysis of transformed plants. Goodness of fit was tested to a 3:1 green:white ratio. With 1 d.f and a 5% significance p value of 3.84.

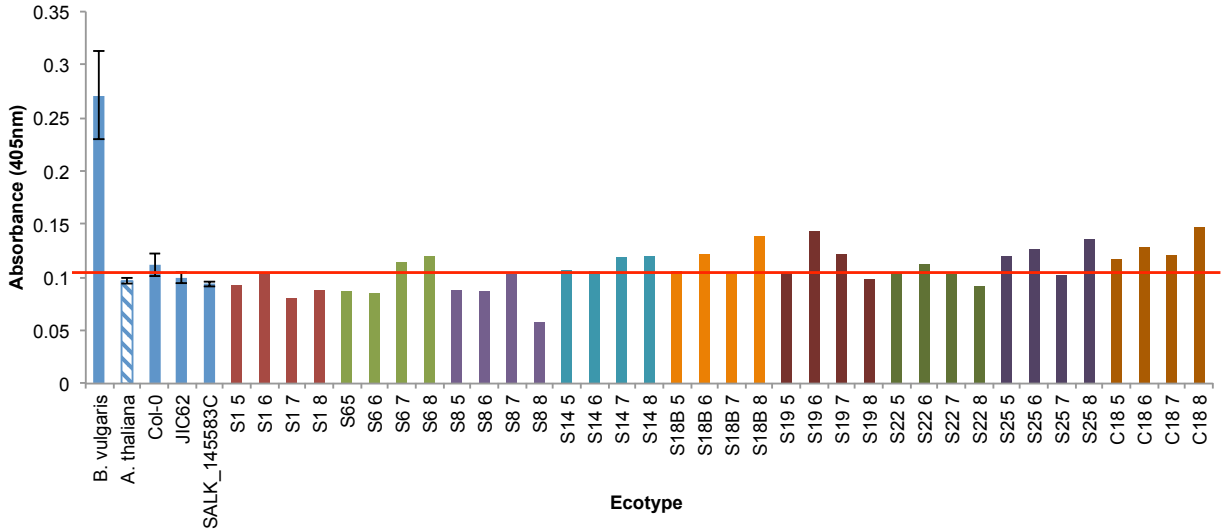
Plant parent number	Number of Green plants per plate	Number of White plants per plate	χ^2 value
S1	102	18	12.10
S5	38	34	18.96
S6	62	9	4.31
S8	57	23	1.70
S9	2	59	167.31
S11	54	37	11.89
S14	80	30	2.12
S18	84	20	1.43
S19	47	23	2.29
S21	30	45	58.00
S22	75	23	0.12
S25	65	21	0.02
S26	6	34	76.80
S28	72	52	19.97
S29	36	36	24.00
S30	10	60	103.20
S32	30	33	25.09
S33	15	55	80.36
S38	70	52	20.20
C18	71	24	0.00

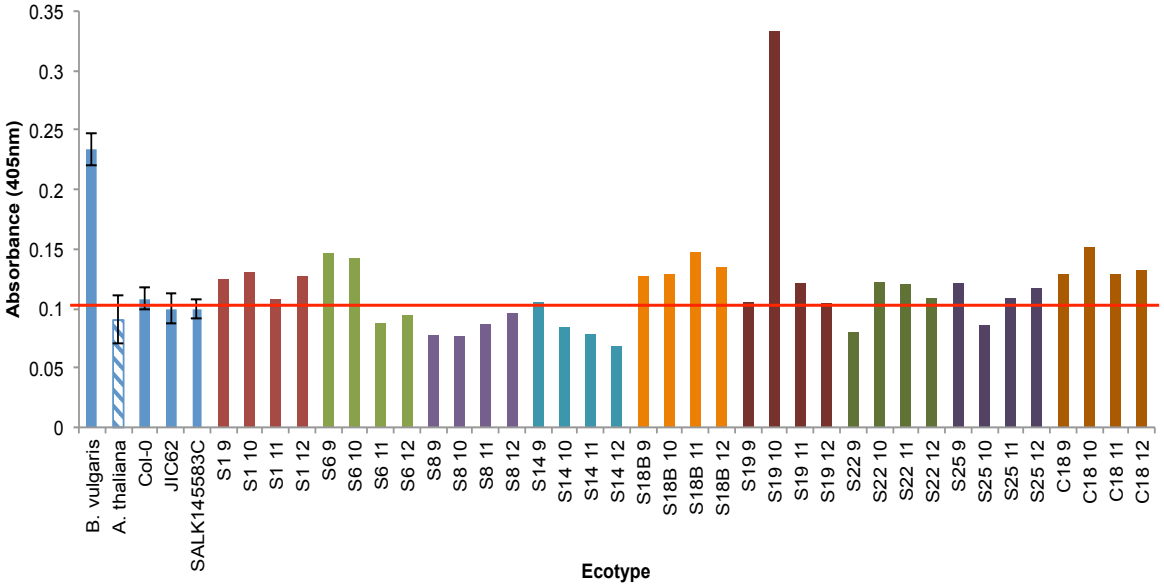
Table 5 indicates a significant 3:1 ratio was found in the transformed plants with progeny numbers 8, 14, 18, 19, 22 and 25. These plants were chosen for BMYV inoculation experiments using viruliferous aphids.

5.6 BMYV infection of transgenic *A. thaliana* plants

TAS-ELISA was used to monitor virus infection in transgenic plants. If *eIF4E* acts as a dominant susceptibility factor for BMYV infection, introducing the Col-0 *eIF4E* allele into resistant plants would be predicted to make these plants more susceptible. TAS-ELISA results can be seen in Figure 5.11 Infected *B. vulgaris* plants were used as a positive control, alongside uninfected *A. thaliana* plants. Col-0, JIC62 and *eIF4E* T-DNA insert (SALK_145583C) control plants were also used as controls. Seed from 8 independent transformants were chosen to study, as well as one Col-0 ecotype plant inoculated with Col-0 *eIF4E* as part of the control. The flower-dipped plants were designated as follows; S1, S6, S8, S14, S18B, S19, S22, S25, and Col-0 C18. Nineteen or twenty self-progeny of each transformant were infected with BMYV, and TAS-ELISA results are shown in Figure 5.11. Results are also summed up in Table 5.3.


Absorbance (405nm)
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
B. vulgaris
A. thaliana
Col-0
JIC62
SALK_145583C
S1 1
S1 2
S1 3
S1 4
S6 1
S6 2
S6 3
S6 4
S8 1
S8 2
S8 3
S8 4
S14 1
S14 2
S14 3
S14 4
S18B 1
S18B 2
S18B 3
S18B 4
S19 1
S19 2
S19 3
S19 4
S22 1
S22 2
S22 3
S22 4
S25 1
S25 2
S25 3
S25 4
C18 1
C18 2
C18 3
Ecotype


Absorbance (405nm)
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
B. vulgaris
A. thaliana
Col-0
JIC62
SALK_145583C
S1 5
S1 6
S1 7
S1 8
S65
S6 6
S6 7
S6 8
S8 5
S8 6
S8 7
S8 8
S14 5
S14 6
S14 7
S14 8
S18B 5
S18B 6
S18B 7
S18B 8
S19 5
S19 6
S19 7
S19 8
S22 5
S22 6
S22 7
S22 8
S25 5
S25 6
S25 7
S25 8
C18 5
C18 6
C18 7
C18 8
Ecotype


Absorbance (405nm)
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
B. vulgaris
A. thaliana
Col-0
JIC62
SALK145583C
S1 9
S1 10
S1 11
S1 12
S6 9
S6 10
S6 11
S6 12
S8 9
S8 10
S8 11
S8 12
S14 9
S14 10
S14 11
S14 12
S18B 9
S18B 10
S18B 11
S18B 12
S19 9
S19 10
S19 11
S19 12
S22 9
S22 10
S22 11
S22 12
S25 9
S25 10
S25 11
S25 12
C18 9
C18 10
C18 11
C18 12
Ecotype

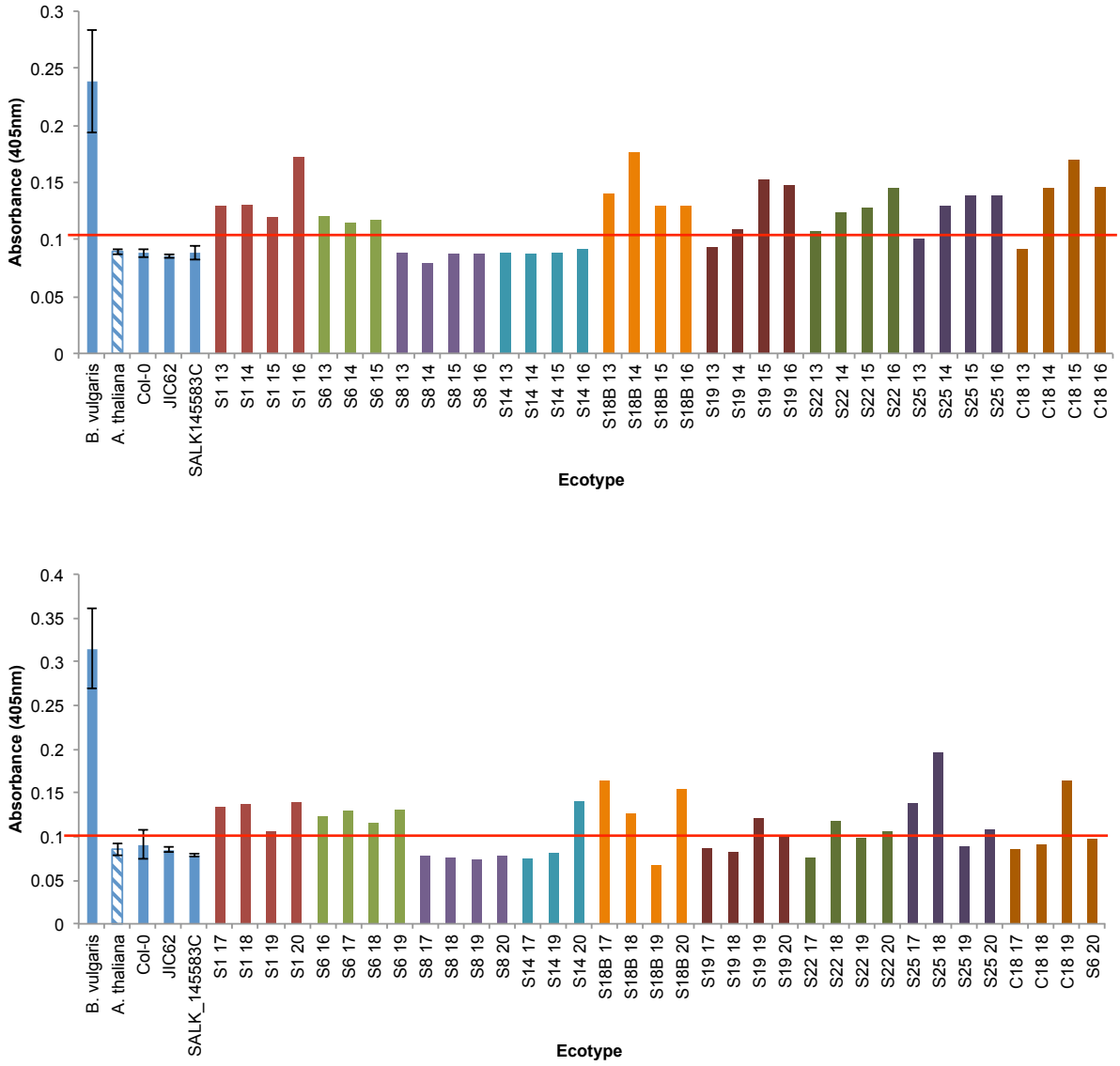


Figure 5.11 Five replicate plates of BMYV infection study TAS-ELISA results. Results show infection levels of plants inoculated with BMYV. Hashed bars represent plants that are not inoculated. The blue bars represent control plants, where three individual plants results were combined to create an average. *B. vulgaris* was used as a positive control as BMYV viruliferous insects were allowed to feed on it. Plants of different parental origin are shown as different coloured bars, which correspond throughout the five replicates for easier identification. The red line shows the computer-derived threshold of resistance to BMYV (above susceptible, below resistant).

Table 5.4 Summary of BMYV infection ELISA results of F_2 progeny from flower dip transformations. The number of uninfected and infected plants determined from TAS-ELISA data are shown in Figure 5.11. The number of infected and uninfected plants are given for each plant line, alongside their χ^2 value. With 2 d.f, a 5% p value is 3.84.

Plant Number	Number of Infected Plants	Number of Uninfected Plants	χ^2 Value
S1	16	4	0.26
S6	12	7	1.42
S8	0	20	60
S14	7	12	14.60
S18	15	5	0.00
S19	13	7	1.07
S22	11	9	4.20
S25	14	6	0.26
C18	14	5	0.018

The plants selected (apart from S8 and C18) were all confirmed to contain Col-0 *eIF4E* (Figure 5.10), and therefore contain at least one copy of the transgenic Col-0 allele. S8 plants were chosen as negative controls they contained no Col-0 *eIF4E*, but still grow on kanamycin media. It was expected that these plants would all be resistant to BMYV, as is shown in Figure 5.11 (in purple) and Table 5.4, as they contain no Col-0 allele. The Col-0 C18 transformed plant, which only contains the Col-0 *eIF4E* allele, should therefore be 100 % susceptible to BMYV infection. As seen in Figure 5.11 (in brown), and Table 5.4, this is not the case and is discussed below.

The progeny of transformed plants that contain Col-0 *eIF4E* at a single locus would be expected to segregate at a ratio of 3:1 susceptible:resistant i.e. as a simple Mendelian trait. The predicted genotypes would be 25% homozygous for the Col-0 *eIF4E* allele, 50% heterozygotes, and 25% containing no Col-0 *eIF4E*. Results show that progeny of transformant S18 fits the 3:1 ratio, shown in Table 5.4 clearly. Other transgenic plants with a ratio extremely close to 3:1

are S1, S6, S19 and S25. These results are good indicators that the plants contain a single Col-0 locus.

In order to confirm the presence of both the alleles of *eIF4E* a further CAPS analysis was carried out on plants selected according to their infection phenotype. The prediction was that for the progeny of each transformant there would be a correlation in the segregants between the presence of the Col-0 allele and susceptibility to BMYV. Uninfected and infected individual progeny from the transformants S1, S6, S8, S18 and S25 were tested using CAPS. The CAPS analysis was carried out using the *eIF4E* sequencing primer set 1 (Table 3.3A), and restriction enzyme digestion with *BspHI*. Digested fragments were run on a 1 % w/v agarose gel (Figure 5.12).

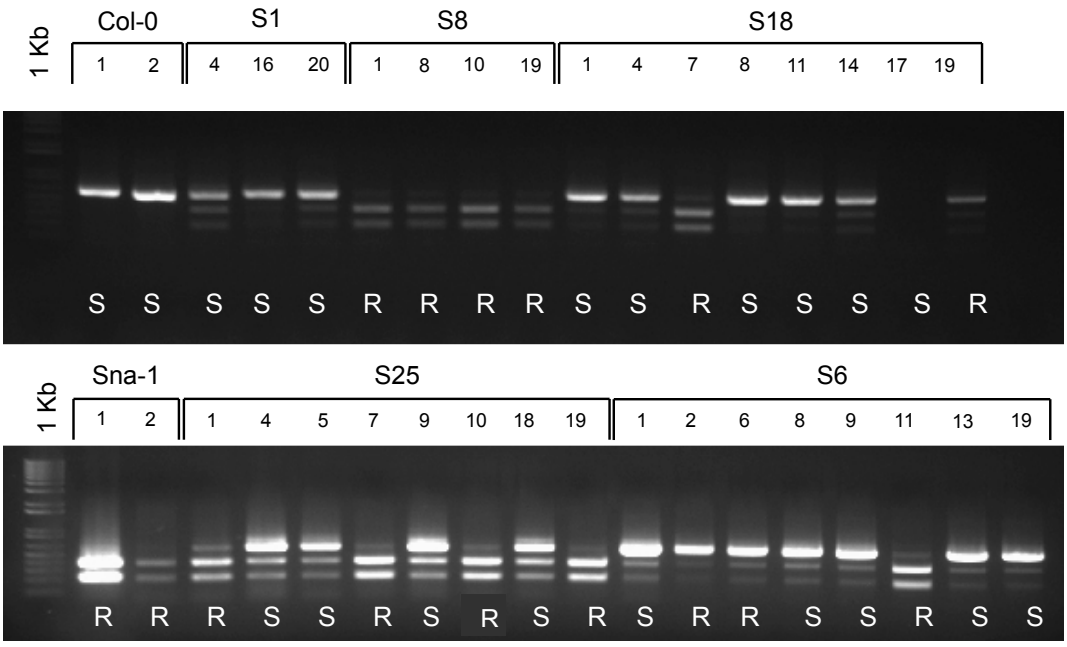


Figure 5.12 CAPS analysis of the DNA from the F_2 generation of JIC flower-dip transformed individuals. Groups of plants from the same parent are labelled S1, S8, S18, S25 and S6. Individual plants within this group are numbered, and directly reference them to the data shown in Figure 5.11. Sna-1 and Col-0 genomic DNA was also used in this analysis as controls. The white letters underneath the DNA bands represent the TAS-ELISA results, plants marked 'S' are susceptible to BMYV, and 'R' resistant to BMYV.

All of the plants should be homozygous for the Sna-1 version of *eIF4E*. This is clearer in some cases than others, but the Sna-1 version of *eIF4E* could be detected in all cases. The plants are therefore segregating for the presence of Col-0 *eIF4E* allele which could have been inserted anywhere in the *A. thaliana* JIC62 genome. S8, a control plant that was not transformed with Col-0 *eIF4E*, is seen in Figure 5.12 and only contains the Sna-1 *eIF4E* allele that was present in the JIC62 parent. Plants lacking Col-0 *eIF4E* are also uninfected with BMYV as shown by TAS-ELISA (Figure 5.11). There is a good correlation between susceptibility to BMYV and presence of the Col-0 *eIF4E* allele, suggesting that *eIF4E* is playing a role in the infection of Arabidopsis.

5.7 Discussion

Successful cloning of the Col-0 *eIF4E* allele into pBin19, and consequently the transformation into *A. thaliana* has allowed a functional complementation test to be carried out. In the previous Chapter, functional knock-outs of *eIF4E* in a susceptible plant resulted in resistance to BMYV. This work is supported by results from other groups (Reinbold et al., 2013). Functional complementation of this gene in *A. thaliana* has not previously been performed in relation to BMYV. The work discussed in this chapter has shown that a naturally resistant line (JIC62) containing the Sna-1 allele of *eIF4E* can be transformed with an allele that confers susceptibility to BMYV.

It was observed that some transgenics, including S8 in this study, are resistant to kanamycin but apparently lack the Col-0 eIF4E allele. The transgenic S8, and c.75% of its progeny were able to grow in media containing kanamycin. One explanation for this phenomenon is incomplete DNA transfer that incorporated the NPTII gene, but not Col-0 *eIF4E*. This would account for the resistance to kanamycin, and the lack of Col-0 *eIF4E* in the CAPS analysis.

During the infection studies, Col-0 (C18) was used as a positive control (Figure 5.11, Table 5.3). Infection however did not result in 100 % infected plants. This may be due to one of the limitations of the experiment where during the process of infection, the aphids have not successfully infected

plants with BMYV. This issue in other plants (such as knock outs performed in previous chapters) has hopefully been accounted for by the number of repeats carried out, but it is important to bare in mind that a plant which appears to be uninfected may not be resistant, but rather unsuccessfully inoculated. In this case, due to the number of previous experiments where Col-0 has been shown to be susceptible to BMYV, and the fact that the aphids used to inoculate the plants were older than ones which are usually used, it is presumed that the reason for low virus titre in these plants is because of unsuccessful inoculation, rather than resistance. This however does not rule out the fact that other plants in this study of functional complementation that appear to be resistant to BMYV, were also unsuccessfully inoculated. Plants; S18 number 4 and 19, and S6 numbers 2 and 6 all appear to be uninfected in TAS-ELISA studies (Figure 5.11), whilst CAPS analysis suggests the presence of Col-0 *eIF4E* (Figure 5.12). These plants could therefore be considered as susceptible to BMYV but unsuccessfully inoculated, although evidence for this would need to be gathered by a repeat infection study with younger (and more) aphids. Another result to notice is that the plant line JIC62 used in these infection studies have similar infection levels to Col-0. There could be two reasons for this. Either Col-0 is failing to be infected as it has previously been, due to the change in location of experiments being performed, or that the JIC62 line is not as resistant to BMYV infection as has previously been seen. Further investigation and genotyping of JIC62 should be performed, and more infection studies carried out.

This functional complementation analysis has provided further evidence that in *A. thaliana* the Sna-1 allele of *eIF4E* functions as a recessive resistance gene to BMYV. Resistant plants that were transformed and contained the dominant susceptible allele of the gene became susceptible to BMYV. Plants that did not successfully incorporate the susceptibility gene, shown by CAPS analysis (Figure 1.13, plant numbers S6: 11, S8: 2,8,10, 19, S18: 7, and S25: 1,7,10, 19), were also shown to be resistant to BMYV in this study (Figure 1.12). This recessive resistance gene has previously been shown to be of importance in plant potyviral resistance mechanisms (Gao et al., 2004; Piron et al., 2010; Ruffel et al., 2005), so resistance to other closely related poleroviruses, such

as TuYV, may exploit a similar resistance mechanism. This will be investigated using similar knock out plants to those described in chapter 4.

Chapter 6 Investigating the role of eIF components in TuYV Infection of *Arabidopsis thaliana*.

6.1 Introduction

Turnip yellows virus and *Beet mild yellowing virus* are both poleroviruses. They share a common genome structure (Figure 1.4), consisting of a 5.6 kb (TuYV) and 5.7 kb (BMYV) single stranded RNA genome (NCBI, 2013), with a 5' viral genome linked protein (VPg). Their sequences have 68.1 % identity (between the published sequences - NCBI, 2013), with most variation occurring at their 5' ends. This region encodes proteins involved in suppressing viral defence mechanisms, and generation of the VPg (NCBI, 2013, van der Wilk et al., 1997).

There have been many studies on plant-virus interactions involving eukaryotic translation elongation initiation factors (eIFs), and these have implicated a variety of eIFs as important recessive resistance genes to RNA virus infection (Albar et al., 2006; Gao et al., 2004; Piron et al., 2010; Ruffel, et al., 2005; Sato, et al., 2005). Recessive resistance genes are more commonly found to inhibit viral infection than fungal or bacterial because virus genomes encode relatively few proteins and require host cellular machinery in order to complete their lifecycle. The recessive resistance genes therefore reflect the loss of compatibility of the virus with host protein complexes, leaving the virus incapable of completing its lifecycle (Robaglia and Caranta, 2006).

The most studied viruses in relation to recessive resistance are those from the potyvirus genus. This is the largest genus of plant viruses, and their genome consists of a single stranded positive RNA, associated with a 5' VPg, and a 3' polyA tail. The potyvirus VPg has frequently been shown to bind to the host eukaryotic translation initiation factor 4E (eIF4E), or its isoform (eIF(iso)4E) (Léonard et al., 2000; Schaad et al., 2000; Wittmann et al., 1997). eIF4E, and its isoform have previously been discussed in relation to virus infection

(section 3.6). Recent studies have examined how potyviruses selectively use eIF components. It has been shown that whilst some potyviruses require a specific eIF protein in order to infect cells, others can use multiple proteins. On the whole, closely related viruses have been shown to require similar host molecular targets for infection, for example both isoforms of eIF4E (Reinbold et al., 2013).

It therefore seems reasonable to expect that, like BMYV, TuYV might target a similar eukaryotic translation initiation factor (eIF) component, or complex. It has also been shown that closely related viruses target the same components within the eIF complex in order to initiate translation of the virus genome. In this case, due to its association with BMYV, the eIF4E protein would be a good starting point to investigate targets for TuYV infection in *A. thaliana*. The aim of the experiments reported here is to use TuYV in infection studies with known *eIF* gene knock-outs in *A. thaliana.*

6.2 TuYV infection studies in Sna-1 and *eIF4E* and putative *eIF(iso)4E* mutants.

The Sna-1 allele of *eIF4E* has already been shown to be a recessive resistance gene to BMYV infection. An investigation to see if TuYV uses the same, or a similar infection mechanism, was therefore performed by infection studies of defined T-DNA insertion mutants. These insertions are located in *eIF4E* and also in the isoform of this gene, *eIF(iso)4E*.

The first study involved infection of the following plants with TuYV; Col-0, JIC62, Ler, SALK_145583C (an NASC line with a T-DNA insertion in *eIF4E*, described in chapter 3) and SALK_003528. The SALK_003528 line contains a T-DNA insert in the 3' untranslated region of the gene *eIF(iso)4E*, and has previously been shown to confer resistance to potyviruses (Reinbold et al., 2013). A schematic diagram of *eIF(iso)4E* is shown in **Error! Reference source not found.** 6.1, detailing T-DNA insertion sites.

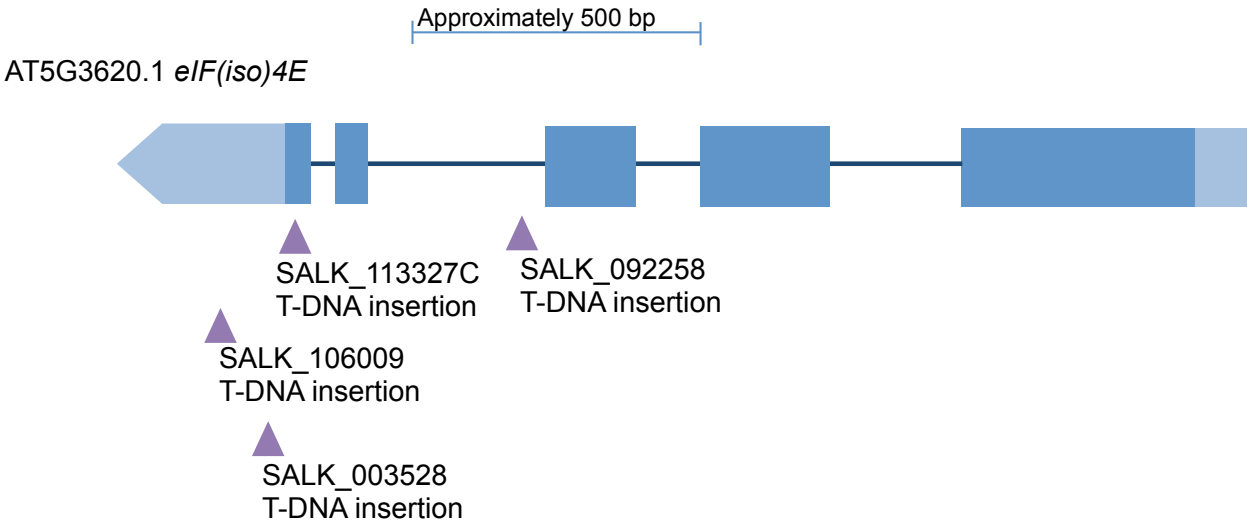


Figure 6.1. A schematic diagram of the *eIF(iso)4E* gene, and the T-DNA insert lines available. In the diagram purple arrows shows the T-DNA insertion sites in the SALK lines SALK_003528, SALK_106009, SALK_113327C and SALK_092258. The pale blue box represents the gene's 3' untranslated region, whilst the darker blue represent exons and the dark blue line represents introns.

Infection studies were carried out using viruliferous aphids to inoculate the plants as described in section 2.3. TAS-ELISA was carried out on infected plants and uninfected controls to measure the amount of virus present in the leaf sap at six weeks post-infection. TAS-ELISA results are shown in Figure 6.2

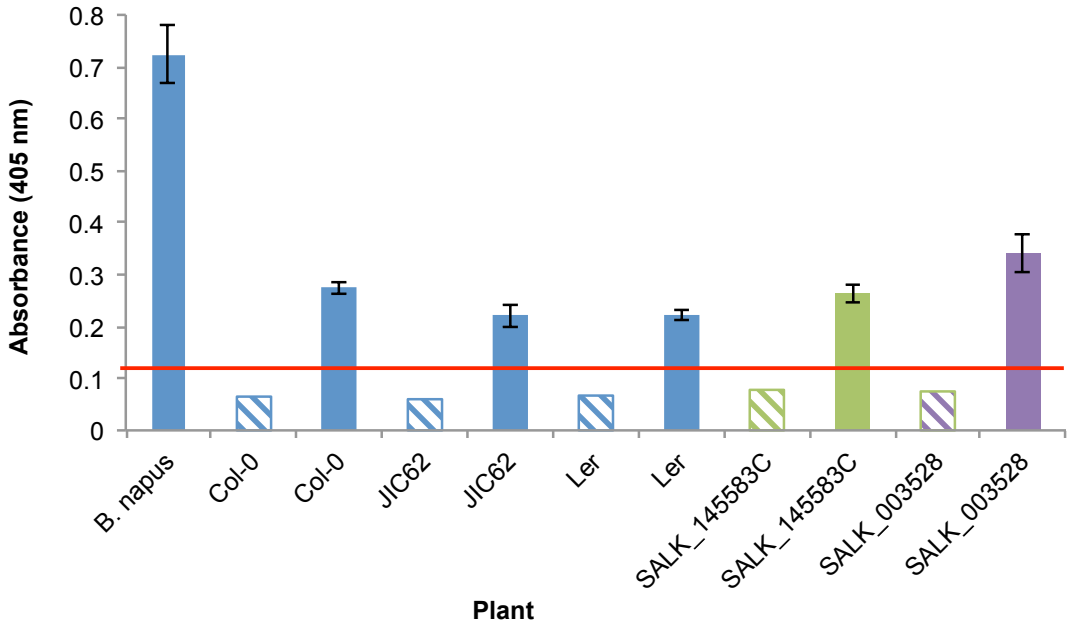


Figure 6.2 TuYV infection studies in control lines and ecotypes and *eIF4e* and *eIF(iso)4E* T-DNA insertion lines. The bars in blue represent controls. *B. napus* was used as a positive control as viruliferous aphids were allowed to feed on this plant. Hashed bars represent uninfected plants, and solid bars indicate plants that were exposed to the virus. Green colour represents the *eIF4E* T-DNA insert pant (SALK_145583C), whilst the purple bar represents the *eIF(iso)4E* T-DNA insert plant (SALK_003528). The red line indicates the threshold for susceptibility to TuYV. Plants with absorbance values greater than this line were considered to be infected with TuYV.

These results show that all plants appeared susceptible to TuYV infection. The four amino acid repeat sequence in the Sna-1 eIF4E protein (as discussed in Chapter 4) does not appear to inhibits infection of TuYV, as it does for BMYV. Loss of *eIF4E* function does not have an affect the ability of TuYV to infect *A. thaliana*.

The *eIF(iso)4E* T-DNA insertion stock (SALK_003528) was also susceptible to TuYV. This could mean that the eIF(iso)4E protein is not required for TuYV infection. However, the mutant line was discovered to be a segregating line for T-DNA insertion that had been ordered, so no conclusions can be drawn from this result. In order to find out if a true homozygous knock out in *eIF(iso)4E* conveys resistance to TuYV, more mutant lines were ordered. Due to time

constraints it was not possible to generate a stock of SALK_003528 pure breeding for the T-DNA insertion.

The lines ordered were the homozygous T-DNA insertion line SALK_113327C, and two segregating lines: SALK_106009 and SALK_092258. The details of the genomic locations of these T-DNA insertion sites are shown in Figure 6.1.

Infection studies with the new T-DNA insert lines were performed as previously described (section 2.3) and the level of virus particles found in plant leaf sap at 8 weeks post-infection was analysed by TAS-ELISA. The two heterozygous lines (SALK_106009 and SALK_092258) were allowed to self-pollinate, the seed collected, and the progeny were used in the infection study. However, as SALK_106009 and SALK_092258 were not pure breeding lines the progeny will still segregate for the T-DNA insertions. But if eIF(iso)4E is required for TuYV infection this might be indicated by a greater SE of virus titres in infected plants, however this was not apparent. The results of this infection experiment can be seen in Figure 6.3.

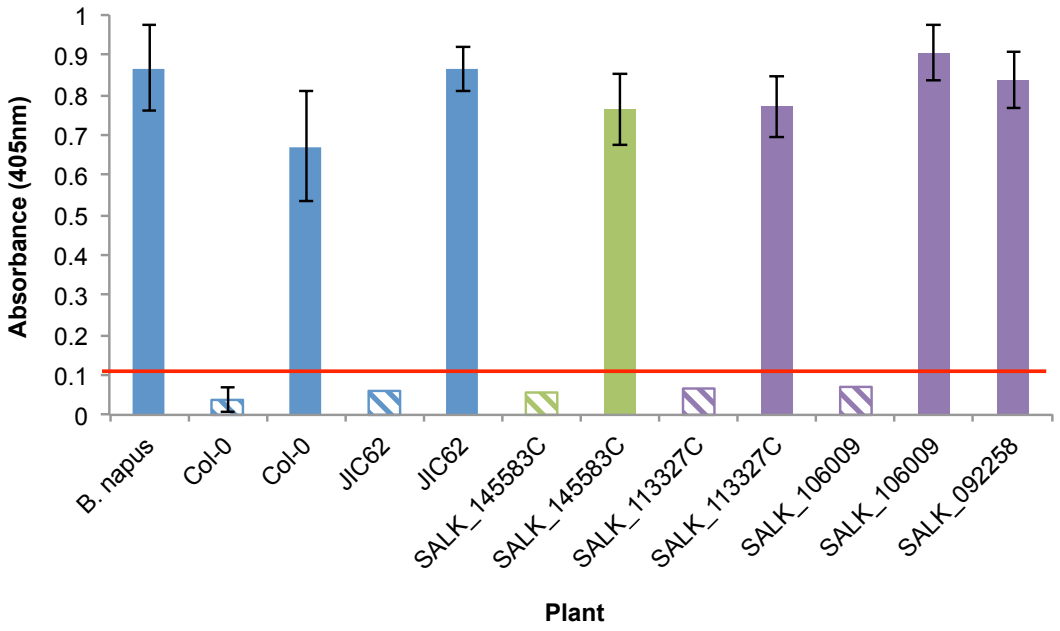


Figure 6.3 TuYV infection studies of *eIF(iso)4E* T-DNA insert lines. The bars in blue are controls. *B. napus* was used as a positive control viruliferous aphids were allowed to feed on this plant. Hashed bars represent uninfected plants, and solid bars indicate plants that were infected with the virus. Green colour represents the *eIF4E* T-DNA insert plant line (SALK_145583C), whilst purple bars represent the *eIF(iso)4E* T-DNA insert stocks (SALK_113327C, SALK_106009 and SALK_092258). The red line indicates the threshold for susceptibility to TuYV, plants with absorbance values greater than this line are considered infected with TuYV.

The results of the infection studies seen in Figure 6.3 indicate much the same result as Figure 6.2. TAS-ELISA indicated high levels of TuYV in mutant lines as well as the various controls. No conclusions can be drawn about eIF(iso)4E function in TuYV infection since none of the lines (including SALK_113327C (see below) are pure breeding for the T-DNA insert in eIFiso4E.

6.2.1 JIC62 x SALK_113327C crosses

Some viruses are able to use either eIF4E or eIF(iso)4E to infect cells (Duprat et al., 2002). Assuming the Sna-1 allele of eIF4E is still functional, and the mutation means that interaction of the TuYV VPg and eIF4E is inhibited in the same way as BMYV, a cross between JIC62 and SALK_113327C (*eIF(iso)4E* T-DNA insertion line) could give viable progeny, but with a mutated *eIF4E*

gene, and a non-functional *eIF(iso)4E*. This investigation might allow us to determine if TuYV can utilise both eIF4E isoforms in order to infect plants. In order to find out if TuYV can use both eIF isoforms of reciprocal crosses were performed between the JIC62 and SALK_113327C. Two crosses were successfully carried out and the seed collected. Ten F_1 plants from each cross were then analysed by CAPS to determine the *eIF4E* allele present. A selection of the results are shown in Figure 6.4.

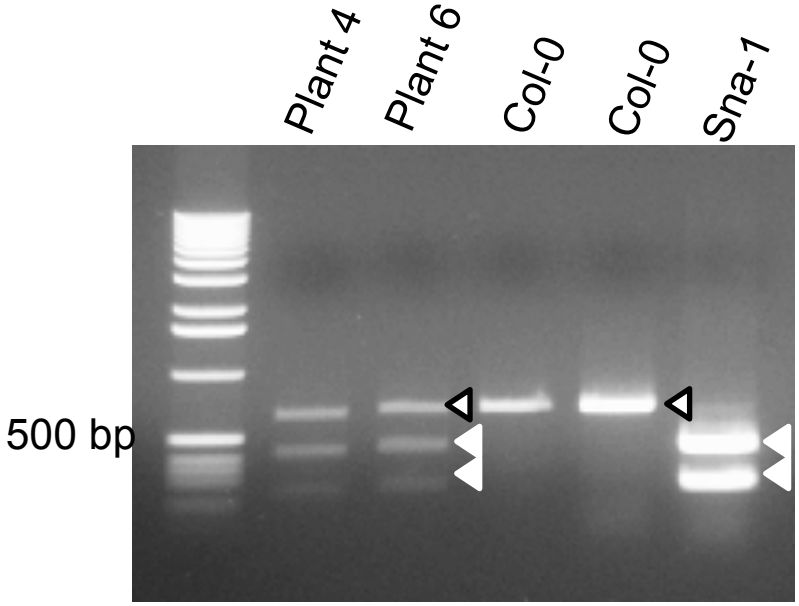


Figure 6.4 CAPS analysis of F1 progeny of JIC62 x SALK_113327C cross. DNA extracted from plants was analysed for the presence of the Col-0 and Sna-1 *eIF4E* allele. Plants 4 and 6 are two of the F_1 generation from this cross. Col-0 and Sna-1 plants are the controls. The white arrow with the black border indicates the Col-0 *eIF4E* allele. The solid white arrow indicates the fragment sizes of the two DNA fragments produced when the Sna-1 allele is digested with *BspHI* restriction enzyme.

Plants 4 and 6 are heterozygous for the Sna-1/Col-0 *eIF4E* alleles as predicted and must have arisen as a result of an authentic cross. About 50% of the progeny analysed were homozygous for the *eIF4E* allele of the maternal parent and therefore were self-pollination contaminants (results not shown). Molecular analysis of F_1s was performed in order to determine if the T-DNA insert of *eIF(iso)4E* was also present in any heterozygous *eIF4E* allele individuals. Two PCRs were performed in order to identify the T-DNA insert in

the *eIF(iso)4E* gene. The insertions present are from the pROK2 plasmid, so the same T-DNA left border primer was used as discussed in Section 4.5 (Figure 4.8). Two PCRs were carried out on F_1 DNAs. Reaction 'A' used primers to amplify the *eIF(iso)4E* gene, whilst reaction 'B' used the T-DNA left border primer, alongside the *eIF(iso)4E* forward primer. This is shown in Figure 6.5.

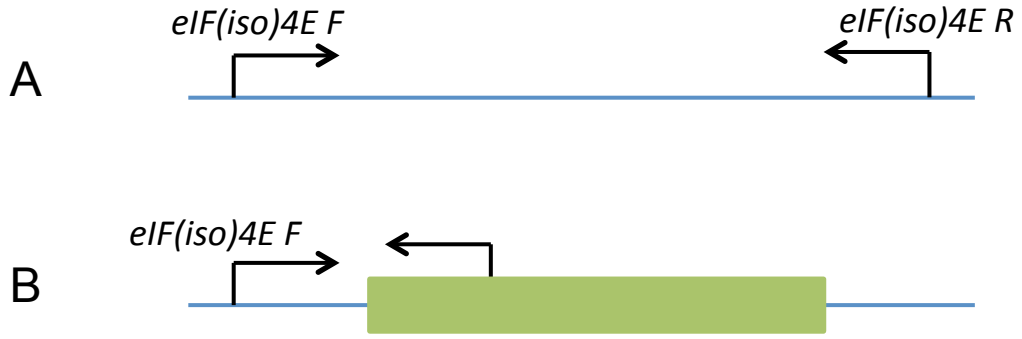


Figure 6.5 Primer sets used for amplification of *eIF(iso)4E* and its T-DNA insertion site. The green block in A and B represents the T-DNA insert, whilst the blue line represents the *eIF(iso)4E* sequence. The black arrows represent the 5'-3' orientation of the primers used.

Using these two primer sets T-DNA insertions can be identified in the *eIF(iso)4E* gene. If a DNA product is generated with primer set A (Figure 6.5), there is at least one wild type *eIF(iso)4E* allele present. If a product is seen using primer set B (Figure 6.5), there is at least one chromosome of the diploid set with a T-DNA insertion in *eIF(iso)4E.* Results of this analysis are seen in Figure 6.6, and the primer sequences are given in Table 6.1.

Table 6.1 Primer sequences for amplification of T-DNA insertions in *eIF(iso)4E*. Primer sequences were used in combinations seen in Figure 6.5

Primer name	Sequence (5'-3')
***eIF(iso)4E* F (I33)**	AGCTCTCCTTGTGGACTAGG
***eIF(iso)4E* R (I34)**	AAAGGTTCAAAATCACAGATACA
T-DNA LB R (H06)	TCCTTTCGCTTTCTTCCCTCCTTTCTC

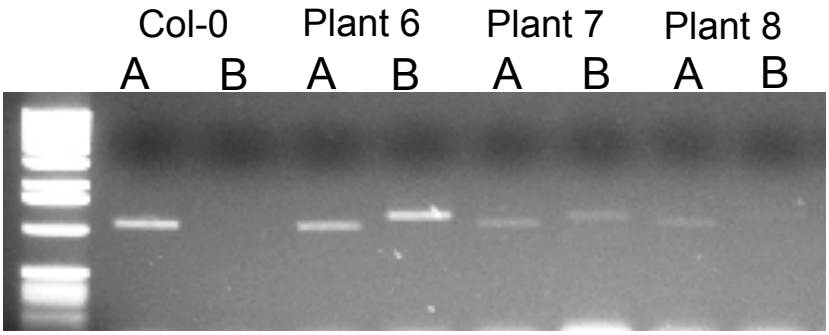


Figure 6.6 PCR results of JIC62 x SALK_113327C F_1 progeny T-DNA insertion verification. Four plants are shown here. Col-0 was used as a control, plants 6, 7 and 8 are all F_1 progeny of the JIC62 x SALK_113327C cross. 'A' represents primer set A used in PCR amplification, and 'B' represents primer set B. Details of primer sets are shown in Figure 6.5.

Results seen in Figure 6.6 are a selection of the results from the crosses described above. Col-0, which does not contain any *eIF(iso)4E* T-DNA insertion can only produce a DNA fragment with primer set A, as expected. Plants 6 and 7 are heterozygous for the T-DNA insertion in *eIF(iso)4E*, as DNA products are observed using both A and B primer sets. Plant 8 is representative of the F_1 plants that contained no T-DNA insertion. Only half of the heterozygotes tested contain the T-DNA suggesting that the SALK_113327C line may not be homozygous for the T-DNA insertion.

Four F_1 plants were identified as good candidates for continued research, i.e. numbers 1, 4, 6 and 11. These plants were heterozygous at both the *eIF4E* and *eIF(iso)4E* locus. The plants were allowed to self pollinate, and the F_2 seed collected. Seed was sown from each of the four parents, and CAPS/T-DNA analysis performed to identify individuals homozygous for the Sna-1 allele of *eIF4E*, and homozygous for the T-DNA insertion in *eIF(iso)4E*. The results of these two analyses can be seen in Figures 6.7 and 6.8.

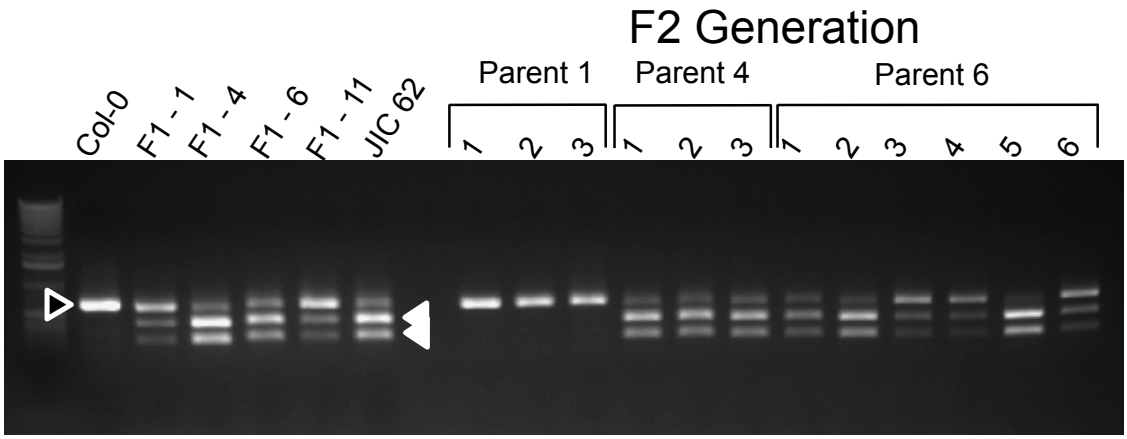


Figure 6.7 CAPS analysis of F_2 progeny of JIC62 x SALK_113327C cross. Col-0 and JIC62 plants were once again used as controls. The size of amplified Col-0 version of *eIF4E* is marked by a black arrow with a white border. The size of the Sna-1 (JIC62) *eIF4E* fragments after *BspHI* digestion is marked with solid white arrows. The plants numbered F1-1, F1-2, F1-4, F1-6 and F1-11 are the CAPS analysis of the F_1 generation from the JIC62 x SALK_113327C cross. The F_2 generation of the cross are labelled with their parent, and then given a number in numerical order.

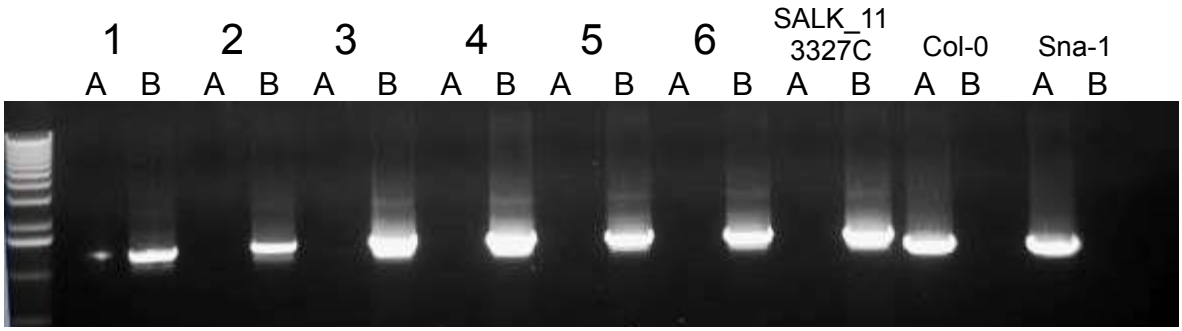


Figure 6.8 PCR results of JIC62 x SALK_113327C F_2 progeny T-DNA insertion verification. Plants numbered 1-6 all come from the parent line number 6. The SALK_113327C line is shown as a positive control, being homozygous for the T-DNA insertion in *eIF(iso)4E*. Two other control plants, Col-0 and Sna-1 plants are also shown. 'A' represents primer set A that and 'B' represents primer set B. Details of primer sets are seen in Figure 6.5.

Plants which were apparently homozygous for the T-DNA insertion in *eIF(iso)4E* (Figure 6.8) and homozygous for the Sna-1 *eIF4E* allele (Figure 6.7) were selected for infection studies. Three plants from the parent number 6 (1, 2, and 5) appeared homozygous for the Sna-1 allele of *eIF4E* (Figure 6.7). Surprisingly, all of these plants also appeared to contain the T-DNA

(using primer set B) and the lack of a PCR product with primer set A suggests they are homozygous for the T-DNA insertion (Figure 6.8). This would need to be verified by molecular analysis of the progeny. Plants 1, 2 and 5 were allowed to self pollinate, and seed was collected to generate lines UEA1, UEA2 and UEA3 respectively. These plants were then used for infection studies. Infection studies using viruliferous aphids were conducted as described section 2.3, and analysed by TAS-ELISA (Figure 6.9).

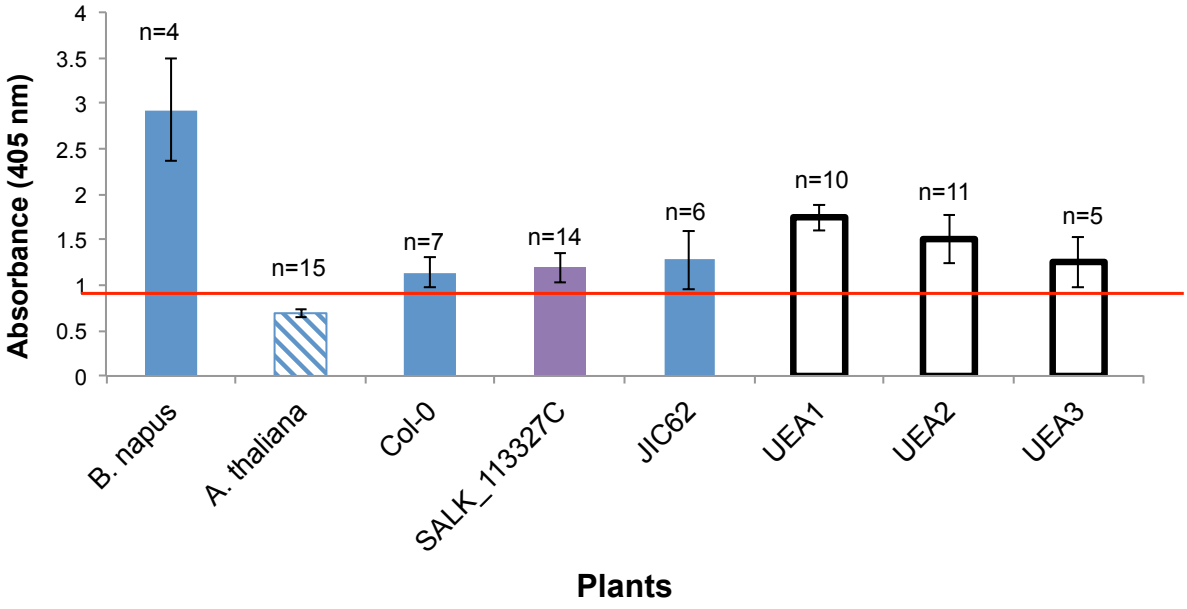


Figure 6.9 TuYV infection studies of UEA lines 1, 2 and 3. The bars in blue are controls. *B. napus* was used as a positive control. Hashed bars represent uninfected plants, and solid bars indicate plants that were infected with the virus. The purple bars represent the *eIF(iso)4E* T-DNA insert plant line, SALK_113327C. The white bars with the black outline represent the UEA lines generated from the original JIC62 x SALK_113327C cross confirmed to be homozygous for the Sna-1 allele of *eIF4E,* and T-DNA insertions in *eIF(iso)4E*. The red line indicates the threshold for susceptibility to TuYV, plants with absorbance values greater than this line are considered infected. The 'n' indicates the number of plants infected, and *A. thaliana* refers to uninfected controls of each plant line and ecotype infected.

Figure 6.9 shows the original parents from the JIC62 x SALK_113327C cross, both are susceptible to TuYV infections as are all three UEA lines but their predicted genotypes (homozygous for the T-DNA insertion and the Sna-1 allele of *eIF4E*) have yet to be confirmed.

6.3 Tools for verifying loss of *eIF* function in T-DNA insertion mutants

A number of molecular resources are described above which can be used to determine the zygosity of the T-DNA insertion mutants described in this study. Ultimately the knock-out phenotype of homozygous T-DNA insertions would need to be confirmed by RT-PCR analysis to confirm that the insertions result in loss of a functional mRNA for each gene. These resources were generated to confirm the T-DNA knock out phenotype of line SALK_115583C and the *eIF(iso)4E* T-DNA insertion line SALK_113327C.

RNA was extracted from Col-0 plants and plants heterozygous for the T-DNA insertions described above (section 2.2.5). RT-PCR was carried out using a mixture of four reverse primers (section 2.4.3) shown in Table 6.2. PCR was then carried out on the cDNA products to amplify specific sections of the cDNA shown in Table 6.2. Results of RT-PCR amplifications are shown in Figure 6.10. The controls used were Col-0 wild type, with primers to the endogenous gene β-tubulin primers were used as controls.

Table 6.2 Primer pair sequences and their targeted sequences for amplification of plant cDNA.

Primer Set	Primer name	Sequence (5'-3')	Target site	cDNA amplification size
A	eIF4E T-DNA F (J54)	GAGAAGGAGACGGAAACGTT	*eIF4E* T-DNA	338 bp
	eIF4E T-DNA R (J55)	GTAGAGCCAGCTCTTATCAG		
B	*eIF4E* F (J56)	GCTTGCATTGATTGGAGAGC	Sna-1 *eIF4E* insert	348 bp
	eIF4E R (J57)	CACTAGCAAAGACAGACTGTC		
C	*eIF(iso)4E* T-DNA F (J70)	CAACCACACAAGTCGAAAG	*eIF(iso)4E* T-DNA insert	510 bp
	eIF(iso)4E T-DNA F (J71)	CCGACCAAACAGTATCACA		
D	β - tubulin F	GAAAGGAATGAGGTTCACTG	β - tubulin	340 bp
	β - tubulin R	TGGGAACTCGCTCATATCT		

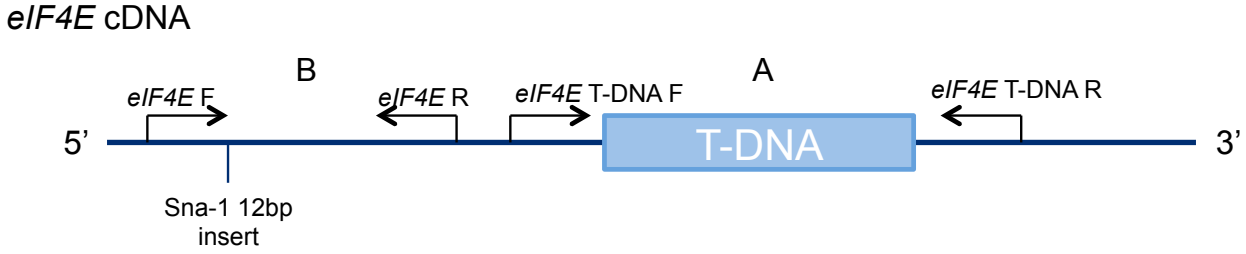


Figure 6.10. The targeted primer sites of the *eIF4E* primer sets A and B given in Table 6.2. The black arrows represent the 5'-3' orientation of the primers used. The blue T-DNA box represents the T-DNA site which would be present in a successful T-DNA insertion plant of the SALK_145583C line.

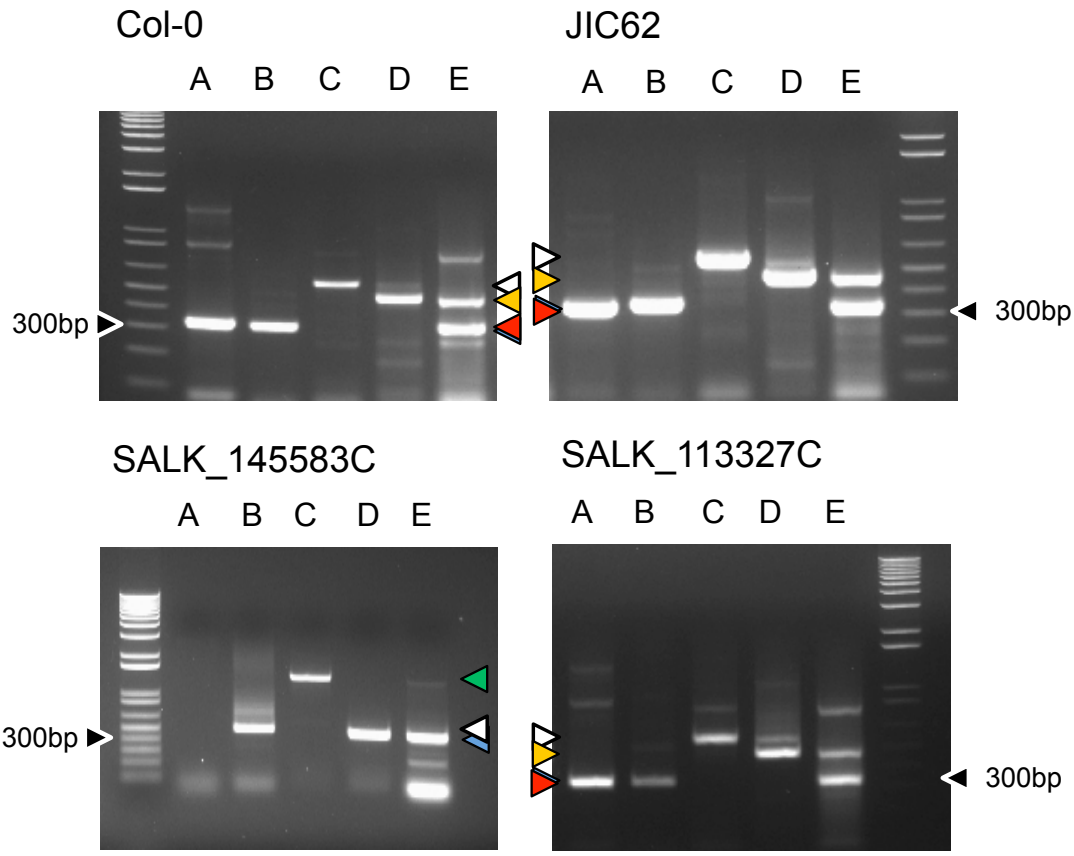


Figure 6.11 PCR of cDNA from plant lines used in TuYV infection studies. Primer sets correspond to those detailed in Table 6.2. Template DNA came from the F_3 selfed generation of the SALK_145583C plants, and the F_2 selfed generation of the SALK_113327C line. Individual plants were used for PCR. The red arrow indicates the product made by primer set A. The blue arrow indicates the product made by primer set B. The white arrow indicates the product made by primer set C. The yellow arrow indicates the product made by primer set D. The green arrow represents the product of primer set C with a genomic DNA template. Primer set E includes all primers. Black arrows with white borders indicate the 300 bp fragment.

Figure 6.11 shows that as expected, all cDNAs can be amplified from Col-0 and JIC62 mRNA, as these lines contain no T-DNA insertions. The SALK_145583C line (*eIF4E* T-DNA insert line) is able to amplify all fragments, except those using primer sequence A to the *eIF4E* sequence which flank the predicted T-DNA insertion. This result confirms that this T-DNA insertion is an authentic knock-out that produces no detectable eIF4E mRNA. This DNA sample must also contain contaminating genomic DNA as the fragment produced in conjunction with primer set C is the predicted size for the DNA

sequence of *eIF(iso)4E*. This sample should be DNase digested and re-amplified. The SALK_113327C line (*eIF(iso)4E* T-DNA) also shows amplification of all fragments, including *eIF(iso4E)*. Since no characterised plants are available that are confirmed to be homozygous for the T-DNA insertion in *eIF(iso4E)* it Is not possible to confirm if the insertion generates a true a gene knock-out.

6.4 TuYV infection studies in *eIF4G*, *eIF(iso)4G1*, and *eIF(iso)4G2* T-DNA insertion lines.

A recent investigation into TuYV infection in Arabidopsis screened a variety of *eIF* T-DNA insertion mutants(Reinbold et al., 2013). Their studies implied that instead of requiring eIF4E or eIF(iso4E) for translation, TuYV in fact targets eIF(iso)4G1. Following these findings the same T-DNA insertion lines were used in an infection study to see if these results could be replicated. The T-DNA insertion lines used are described in Table 6.3.

Table 6.3 Details of T-DNA insert lines in *eIF4G* complexes for TuYV inoculations.

SALK line	T-DNA insert Gene	Chromosome location	Insert in gene exon	Homozygous/ Heterozygous
SALK_112882C	*eIF4G*	AT3G60240	7	Homozygous
SALK_009905C	*eIF(iso)4G1*	AT1G17330	6	Homozygous
SALK_076633C	*eIF(iso)4G2*	AT2G24050	4	Homozygous

All lines are homozygous and the same reported lines as used by Reinbold et al. (2013). Unfortunately, no *eIF(iso)4G1G2* double mutants were available, and time limitations prevented the creation of a double mutant line. Infections with TuYV were carried out as previously described (section 2.3), and results can be seen in Figure 6.12.

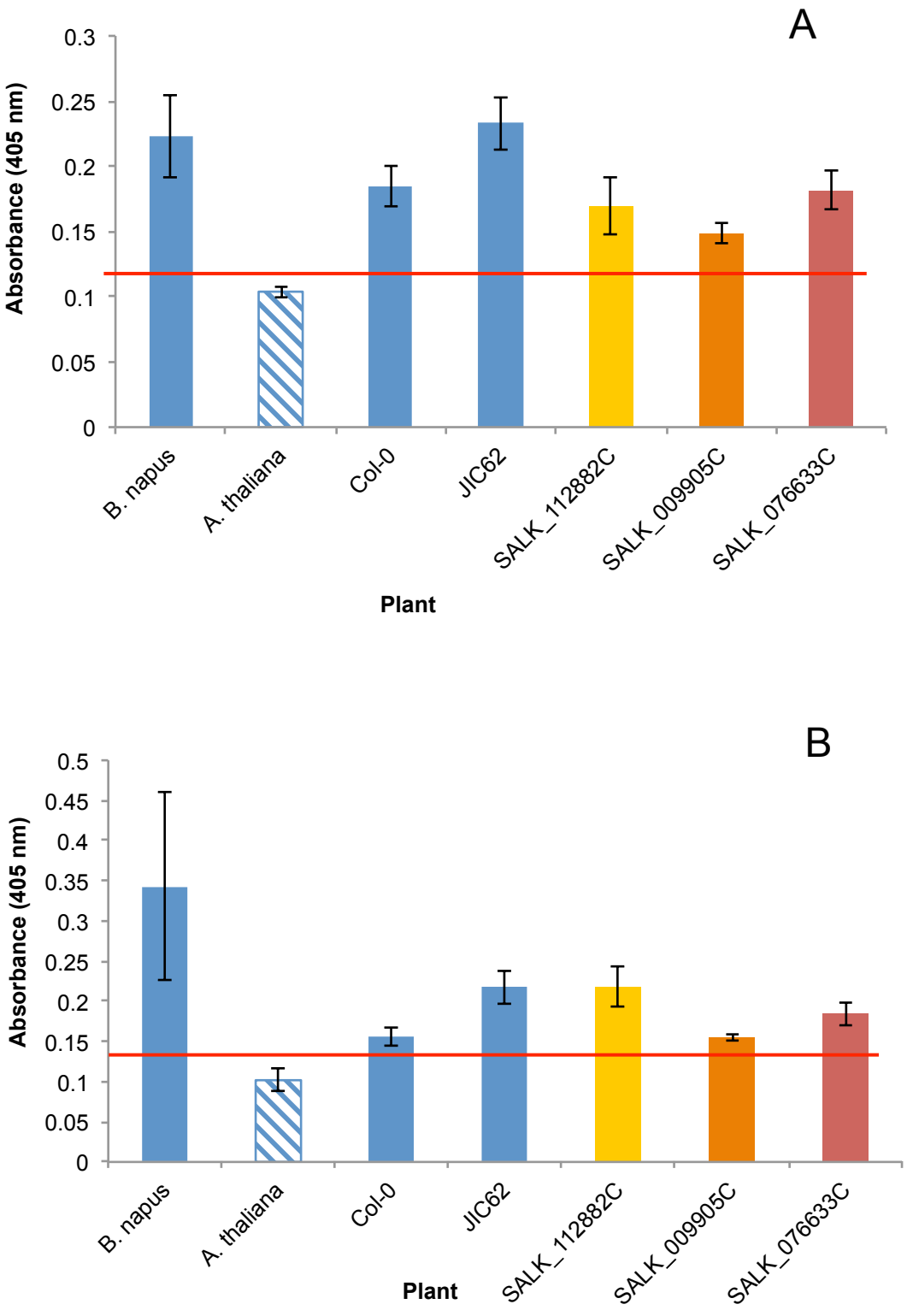


Figure 6.12 TAS-ELISA infection studies of TuYV and various T-DNA insert lines of *A. thaliana*. The two repeats are shown (A and B). The bars in blue are controls. *B. napus* was used as a positive control as viruliferous TuYV aphids were allowed to feed on this plant. Hashed bars represent uninfected plants, and solid bars indicate plants that were exposed to the virus. The yellow bars represent *eIF4G* T-

DNA insertion line (SALK_112882C). The orange bars represent the *eIF(iso)4G1* T-DNA inserts (SALK_009905C), and the red bars represent *eIF(iso)4G2* T-DNA insertion lines (SALK_076633C). The 'n' indicates the number of repeats for each plant, and standard error bars are shown. The red line indicates the threshold for susceptibility to TuYV, plants with absorbance values greater than this line are considered infected with TuYV and *A. thaliana* refers to uninfected controls of each plant line and ecotype infected.

The results of the infection study (Figure 6.12) show that according to the TAS-ELISA method used, all mutants are susceptible to TuYV infection. The *eIF(iso)4G1* mutant line does appear to have the lowest value for the accumulation of virus particles, but will need to be tested to find if this result is significantly different to the Col-0 infection levels.

In order to determine the significance of the different absorbance levels between the *eIF(iso)4G1* T-DNA insertion line (SALK_009905C, shown in orange in Figure 6.12) and Col-0 a T-test was performed. Results are shown in Table 6.4.

Table 6.4 T-test results comparing virus accumulation in Col-0 plants Vs. SALK_009905C. Results are all shown to three significant figures. Average values are taken from those seen in Figure 6.12.

Repeat	Plant	Average Abs (405nm)	Standard Deviation	t value	Degrees of freedom	5% significance p value
A	Col-0	0.185	0.035	**0.373**	12	2.179
	SALK_009905C	0.149	0.023			
B	Col-0	0.156	0.027	**0.002**		
	SALK_009905C	0.155	0.013			

Results from Table 6.4 suggest that there is no significant difference between the virus accumulation levels of Col_0 and SALK_009905C for either repeat.

These plants have not yet been tested by RT-PCR to indicate the presence of a homozygous T-DNA knock-out of any of these genes. These results are assuming that the lines ordered are true knock-outs.

6.5 Discussion

The aim of these experiments was to determine if TuYV, a very significant pathogen of oilseed rape, targets similar eIF components during infection as BMYV. This knowledge could have important practical applications in engineering resistant oilseed rape varieties. The eIF4E protein in Arabidopsis was previously shown to be an important virus susceptibility factor for BMYV, so infection studies were carried out with putative *eIF* T-DNA knock out lines. As was previously shown in potyvirus infection, closely related viruses often share similar infection strategies in plants (Reinbold et al., 2013; Robaglia and Caranta, 2006). Given the sequence similarities observed between BMYV and TuYV it was therefore assumed that TuYV will target either eIF4E, eIF(iso)4E, or both, during infection.

The initial study of TuYV infection of an *eIF4E* knock-out line and a putative *eIF(iso)4E* T-DNA insertion line (Figure 6.2 and 6.3) showed that TuYV was able to successfully infect both. This result also confirmed that BMYV resistance in the Sna-1 ecotype is due to virus resistance, rather than a more general aphid resistance phenomenon. The observation that TuYV is able to infect both the Sna-1 ecotype and SALK_145583C (*eIF4E* T-DNA insert line) is consistent with a number of different models. Clearly, TuYV does not rely solely on eIF4E for establishing infection in Arabidopsis and it may use a different molecular mechanism for infection of Arabidopsis than BMYV. Closely related viruses have previously been shown to use similar, but not identical, infection methods e.g. certain potyviruses have been shown to target either eIF4E or eIF(iso)4E during infection. This is possibly the case with TuYV and the aim was to test this genetically.

In the initial infection experiment (Figure 6.2) the stock SALK_003528 (*eIF(iso)4E* T-DNA insertion) line was also found to allow TuYV accumulation

within cells. However, this line was subsequently shown to be heterozygous for the T-DNA insertion, and can still express *eIF(iso)4E*. The experiments were therefore repeated on other stocks that might breed true for the T-DNA insertion in *eIF(iso)4E*, such as SALK_11327C (Figure 6.3). These experiments replicated the results previously seen (Figure 6.2) where the T-DNA insert in *eIF(iso)4E* had no effect on the ability for TuYV to accumulate within the plants. This result was fairly unexpected as it has not previously been reported that such closely related viruses would use different such different infection methods. It was subsequently shown that SALK_113327C may not be homozygous (Section 6.3) and therefore no conclusions can be drawn from this analysis. If these plants are segregating for the presence/absence of the T-DNA, and the gene is important for TuYV infection of Arabidopsis, this be manifested by a wider range of standard deviation within the samples infected. This was not apparent (Figure 6.3), but could be explained by the small sample size used. The experiment should be repeated with a confirmed homozygous T-DNA insertion line in the *eIF(iso)4E*, with a large sample size.

In order to investigate if TuYV can use both isoforms of eIF4E to infect a plant, attempts were made to create new lines from the cross JIC62 x SALK_113327C. The F_2 generation of this cross was examined to find individual plants that were homozygous for both the Sna-1 *eIF4E* allele, and the T-DNA insertion in *eIF(iso)4E* (Figure 6.7 and 6.8). Three F_2 individuals of the desired genotype were identified by molecular analysis (Figures 6.4 and 6.6) and used to created UEA1, UEA2 and UEA3 lines. The self progeny of the these lines were infected with TuYV. Results showed that all the generated UEA lines in this infection study were susceptible to TuYV (Figure 6.9). This could mean that TuYV is using a different molecular mechanism to infect Arabidopsis, for example a different family of proteins from the eIF complex, or that the Sna-1 version of eIF4E is still compatible with TuYV. The assumption in this study has been that the BMYV and TuYV VPgs are similar enough that the disruption of the plant eIF4E cap binding pocket that results in resistance to BMYV, has the same effect on TuYV. Further investigations of the possible interaction of TuYV and eIF4E are required to confirm this. Also,

due to time constraints, it has not been possible to confirm the genotypes of the UEA1, UEA2 and UEA 3 progeny and, despite being homozygous for Sna-1 eIF4E, they might still be segregating for the T-DNA insertion in *eIF(iso)4E*. Once the genotypes of these plants are confirmed with the molecular markers generated in this research (Figure 6.10) an important test will be to confirm that the eIF(iso)4E mRNA is no longer produced.

Shortly after the completion these infection studies, Reinbold et. al. (2013) published results implicating a different eIF component in TuYV infection. This study reported that knock-outs of *eIF(iso)4G1* resulted in a four-fold decrease in the accumulation of virus particles compared to Col-0. *eIF(iso)4G1* is one of the proteins required to create the isoform of the eIF4F complex, eIF(iso)4F. The *eIF(iso)4G1* gene is encoded for on Arabidopsis chromosome V, as is *eIF(iso)4E.* These two components form the *eIF(iso)4F* complex. The proteins involved in Arabidopsis eIF complex formation and their chromosomal origins are shown in Figure 6.13.

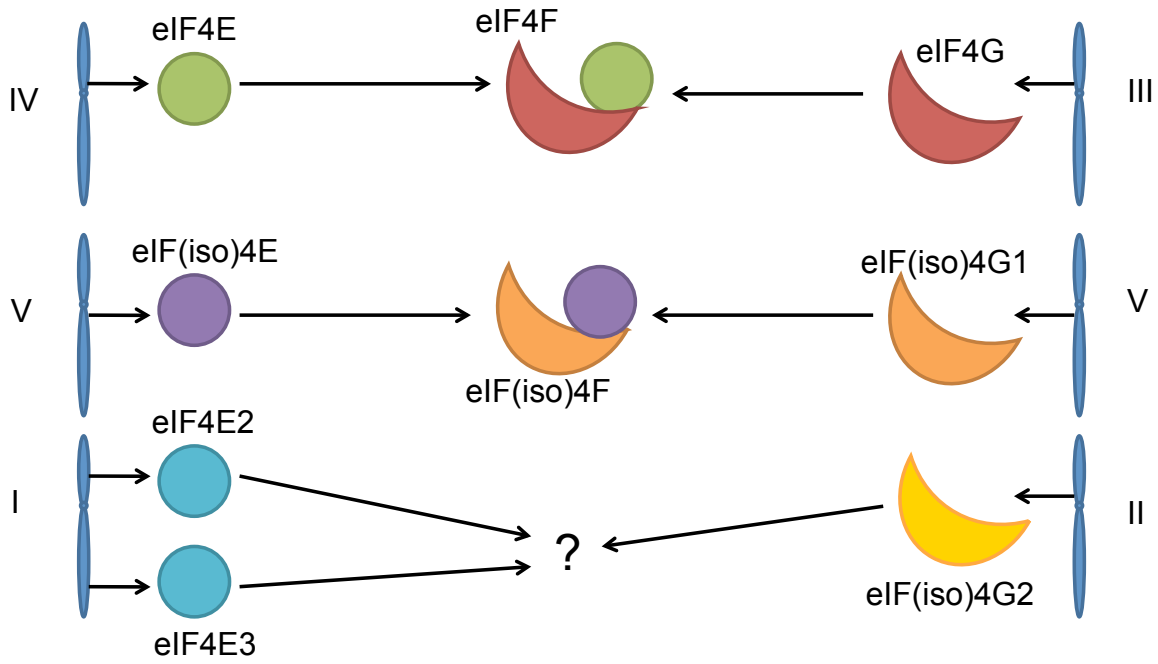


Figure 6.13. The structure of the eIF4F complexes, and the chromosomal locations of the corresponding genes. Arrows show the structures that form proven interactions. It is not currently known what eIF4E2 and 3, and eIF(iso)4G2 complexes may form, if any. Figure modified from (Robaglia and Caranta, 2006).

In their study, Reinbold et al. (2013) infected T-DNA insertion lines of all components shown in Figure 6.13. Similar to this study, knocking out *eIF4E* and *eIF(iso)4E* individually had no effect on host resistance to TuYV. Their experiments suggested that eIF(iso)4G1 was required by TuYV in order to infect Arabidopsis, and a T-DNA insertion in this gene directly resulted in a statistically significant four-fold reduction in virus accumulation. The study also created a double knockout mutant line of *eIF(iso)4G1G2.* This line exhibited severe growth defects, as well as high resistance levels to TuYV. These growth defects were not reported in either of the *eIF(iso)4G* T-DNA insertion lines. Reinbold et al., have therefore suggested that eIF(iso)4G2 is only present, or utilised, in the absence of eIF(iso)4G1, but that at least one version of eIF(iso)4G is required in order for normal plant development, although as yet, a direct association between eIF(iso)4E and eIF(iso)4G2 has not been demonstrated. Reinbold et. al. also carried out yeast two-hybrid experiments to confirm the interaction between eIF(iso)4G1 and TuYV VPg. This interaction is not completely novel, as a similar interaction has been documented for susceptibility of rice to the potyvirus *Rice yellow mottle virus* (RYMV) (Hébrard et al., 2009).

This study attempted to replicate Reinbold's findings, using the same T-DNA insert lines. The isolate used by Reinbold was the French isolate TuYV-FL1, whilst the isolate used in these tests was the UK isolate UK-BB. Unfortunately no double knock out mutation lines of *eIF(iso)4G1G2* are available, and time constraints meant that crosses could not be performed. Results seen in Figure 6.12 suggest that none of the T-DNA insertion lines resulted in resistance to TuYV in Arabidopsis. Further analysis of the data to confirm whether this experiment led to a significant decrease in TuYV accumulation in Arabidopsis, using the statistical t-Test (Table 6.4) was performed. These results showed that there was no significant difference found between Col-0 and *eIF(iso)4G2* (SALK_009905C) lines virus accumulation levels in either repeat. These results suggest that the knock out of *eIF(iso)4G2* has no effect on the ability of the TuYV UK-BB isolate to infect Arabidopsis. Another possibility, is that these lines are not pure breeding as claimed and this would need to be verified by molecular analysis of their progeny.

The TuYV isolate used in studies by Reinbold et al. (2013) was the isolate formally known as BWYV-FL1 (Veidt et al., 1992). This isolate is the current reference sequence for TuYV. Attempted cloning experiments performed in this study of the UK-BB TuYV isolate suggests that the sequences of these two viruses differ quite considerably. This difference might have consequences on the mechanism of infection of these two strains, but this seems very unlikely. In the study by Reinbold, virus accumulation was found to be impacted most significantly by removing *eIF(iso)4G1,* and less so by *eIF(iso)4G2*, whilst a double mutant *eIF(iso)4G1G2* was shown to have the greatest effect on decreasing virus accumulation. One explanation of the results seen here with the UK-BB TuYV isolate may be that this isolate is able to use both eIF(iso)4G1 and eIF(iso)4G2 proteins equally in order to initiate replication of the virus mRNA. In order to confirm this, further investigation and yeast two-hybrid experiments need to be performed to confirm an interaction between the eIF(iso)4G2 protein and TuYV.

Chapter 7 General Discussion

7.1 Investigating resistance BMYV resistance in Arabidopsis

7.1.1 Introduction

One of the original aims of this study was to characterise a gene controlling resistance or susceptibility to *Beet mild yellowing virus* (BMYV) in *Arabidopsis thaliana*. This was achieved by screening natural variation within *Arabidopsis thaliana*. This is a useful method to find resistance to viruses, without the costly and time-consuming creation of mutant lines. An added advantage to this approach is the understanding of how orthologs of these genes in crop plants could be identified and targeted (or altered) to avoid the use of controversial genetic modification technologies. A previous study of natural variation within Arabidopsis was carried out, and identified an ecotype which appeared to be resistant to BMYV infection (Percival-Alwyn, 2010). This resistance could have been caused by an *R* gene response to infection, or as a recessive resistance gene. By analysing a segregating population of the progeny of Sna-1 x Col-0 (JIC62) plants, it was found that a 1:2:1 ratio of high resistance:intermediate susceptibility:high susceptibility was observed after infection with BMYV. This indicates that a single semi-dominant monogenic trait controls susceptibility to BMYV in these plants. The results meant that the most likely resistance mechanism was a recessive (or passive) resistance gene in the resistant Sna-1 ecotype. A molecular analysis showed this was most likely the consequence of a duplicated 12 bp sequence within the gene and a four amino acid duplication in the mature protein (Section 3.2).

7.1.2 *eIF4E* identification

This study began by investigating the differences between F_2 progeny from the Col-0 x Sna-1 cross either resistant or susceptible to BMYV by carrying out an AFLP analysis of DNA bulks. The DNA bulks were taken from the 20 most infected and 20 least BMYV infected plants. By using these bulks, it was

predicted that unlinked markers in the Arabidopsis ecotypes would be accounted for in each bulk, and that the major segregating factor between the two bulks would be the gene responsible for susceptibility to BMYV. The AFLP analysis identified six fragments of DNA that were present exclusively in susceptible plants (Figure 3.1). Recessive resistance is often caused by the absence of a gene, or a mutation in a gene, meaning the virus is unable to interact with the host in order to complete its viral lifecycle. We were therefore looking for a fragment of DNA, which is present in susceptible plants, but missing in the resistant plant. The six fragments were sequenced, and it was shown that four mapped to Arabidopsis chromosome 4, in a region known to containing the *eukaryotic translation initiation factor 4E* gene (Figure 3.2, Table 3.2). This gene had previously been implicated as an important factor for other virus infections such as potyviruses. This identified a candidate gene for further investigation using the considerable genetic resources available for A. thaliana.

In order to find out if there was any clear difference between the Sna-1 and Col-0 *eIF4E*, the gene from Sna-1 was sequenced, and directly compared with the published sequence of Col-0. Sequence analysis indicated a 12 bp insert, which was a direct repeat of the Sna-1 sequence (Figures 3.3 and 3.4). This repeat sequence probably occurred during DNA replication, causing the polymerase to repeat the replication of the same fragment of DNA (Kornberg, 2005). It is presumed that as this repeat did not cause any loss of fitness to the plant, it has been maintained. As Sna-1 was found growing in a region that has large amounts of sugar beet cultivation (Suffolk), the repeat may have increased plant fitness as it confers resistance to BMYV, which is prevalent in the area. By modelling the Arabidopsis eIF4E (based on pea eIF4E, Figure 3.5), the 12 bp insert repeat, which gave rise to a 4 amino acid insert repeat, was found to create a loop in the cap-binding region of the protein. Mutations in the cap-binding region have previously been shown to convey resistance to other plant virus infections (Browning, 2004).

With the identification of *eIF4E* through AFLP analysis, and subsequently the sequencing differences between the Sna-1 and Col-0 *eIF4E,* infection studies

were carried out on T-DNA knock out lines of Col-0 in the *eIF4E* gene. These studies showed that by knocking out function of the *eIF4E* gene in the susceptible Col-0 plant, BMYV virions could not accumulate inside the plant (Figure 4.6). These results were supported by functional complementation analysis where plants resistant to BMYV became susceptible with the introduction of the susceptible Col-0 *eIF4E* allele (Figure 5.11).

7.1.3 Translation initiation factors

The translation initiation protein eIF4E has previously been shown to act as an important recessive resistance trait to many viruses (Le Gall et al., 2011). For this reason breeders often exploit naturally occurring mutations within this gene to create effective and sustainable resistance to plant diseases (Michelmore, 2003). Around half of the 200 known resistance genes are recessive, and these genes have been reported in barley, rice, maize and yam crop species, as well as many others. More importantly, recessive resistance genes are increasingly important for creating crops that are resistant to RNA viruses, with 70% of recessive resistance genes known to influence virus infection or movement (Charron et al., 2008; Diaz-Pendon et al., 2004; Le Gall et al., 2011). Studies investigating translation initiation proteins and recessive resistance have only so far implicated eIF4E and eIF4G (and their isoforms) as virus susceptibility factors (Le Gall et al., 2011).

All eukaryotic cells encode translation initiation factor proteins, but only plants contain a second copy, or isoform, of the eIF4F proteins. Many studies have been carried out to try and understand the purpose of having two copies (Browning, 2004). eIF4E and eIF(iso)4E have been shown to have 50% similarity in their protein sequences, and both have a molecular mass of around 24kDa. The eIF4G and eIF(iso)4G proteins are considerably different with molecular masses in wheat of 180 kDa and 86 kDa respectively. This suggests that these proteins have diverged to have different functions within the plant (Le Gall et al., 2011). The functional significance of two eIF4F complexes is still not fully understood, but recent evidence suggests that they are able to discriminate between, and selectively recruit, different mRNA

structures, although the eIF4F complexes seem to be functionally interchangeable (Gallie and Browning, 2001; Le Gall et al., 2011). Most viruses studied so far require a single host protein in order to infect the plant, but some viruses have evolved the ability to utilise more than one protein, although this has only been shown in the more similar eIF proteins eIF4E and eIF(iso)4E (Duprat et al., 2002).

In eukaryotic cells the role of eIF4E is to bind the m^7G 5' cap of mRNA and recruit other eIF proteins such as eIF4G (see Figure 3.8) in order to initiate translation, and to recruit ribosomal subunits. The eIF4E and eIF4G protein complex, also known as eIF4F, recruit proteins eIF4A, and eIF3 and the polyA binding protein (PABP). The PABP plays an important role in mRNA translation as it allows circularisation of the structure enabling efficient translation of the mRNA (Wells et al.,1998). Most plant viruses studied to date belong to the potyvirus family. These viruses have a 5' VPg (which is thought to substitute for the 5' m^7g cap of mRNA) and a polyadenylated tail (Le Gall et al., 2011). These viruses have been shown to directly interact with eIF4E and eIF(iso)4E through their VPgs (Léonard et al., 2000). Mutation in these genes results in the virus being unable to accumulate within plants, suggesting the interaction plays a role in the virus lifecycle within the plant (Bruun-Rasmussen et al., 2007; Kang et al., 2005; Kanyuka et al., 2005; Ruffel et al., 2005, 2002, 2006). Most of the naturally occurring recessive resistance genes are caused by a change of one to five non-conserved amino acids of eIF4E (or eIF(iso)4E) in the cap-binding pocket (Charron et al., 2008; Le Gall et al., 2011). The ability of potyviruses to overcome these natural mutations has also been demonstrated by sequence variation on the surface of the VPg (Roudet-Tavert et al., 2007). This accumulated evidence gives a strong indication of the importance of the interaction between the VPg and eIF4E or eIF(iso)4E in potyvirus infection.

Similarly to potyviruses, poleroviruses have a 5' VPg but poleroviruses do not have a 3' polyadenylated tail. Previously the polyA tail has been shown to act as an important translational enhancer, which coordinates with the 5' VPg and stabilises the RNA (Gallie, 1991). It is thought that another factor in the 3'

UTR must therefore substitute for other ssRNA viruses that lack a polyA tail. In some viruses such as *Alfalfa mosaic virus* (AMV), the viral coat protein has been identified as performing this role, although the specifics of these interactions have not yet been identified (Le Gall et al., 2011). It has also been observed that the insertion of an artificial polyA tail in AMV removes the need for the coat protein to initiate translation (Neeleman et al., 2001). Although these 3' UTR structures have not yet been identified in poleroviruses, stem loop structures in the 3' UTR of the polerovirus *Pepper vein yellow virus* have been observed (Murakami et al., 2011), and the 3' UTR of the closely related luteovirus *Barley yellow dwarf virus* contains a translation enhancer sequence that fulfils a similar role (Domier et al., 2002).

7.1.4 *eIF4E* and Viral VpG associations

It is largely undisputed that eIF4E or eIF(iso)4E interacts with the VPg of viral genomes, and that this association is vital in order for infection of the plant, but understanding exactly why this interaction takes place is disputed. The VPg may be present to directly mimic the m^7g of the host mRNA to initiate translation, but there is also evidence that interaction of VPg and eIF4E/(iso)4E facilitates cap-independent translation of the virus (Lellis et al., 2002). It has also been argued that the interaction may prevent cap-dependent translation of host mRNA, freeing ribosomes for the translation of viral RNA (Dreher and Miller, 2006), or that the interaction is not important for translation at all, but for cell-to-cell movement of viral RNA (Gao et al., 2004).

It has been previously shown by yeast-two hybrid experiments that the BMYV VPg interacts directly with eIF4E (Reinbold et al., 2013). This supports the observation that the removal of the *eIF4E* gene or the truncation of this protein abolishes the interaction, causing Arabidopsis to become resistant to BMYV infection. Whether the interaction is important for replication of the virus is still unknown. Potyvirus replication has been shown to be successful in the absence of the VPg by replication being initiated by cap-independent translation mechanisms, although VPg is likely to assist the recruitment of

translation initiation factors and stimulate translation (Thivierge et al., 2005). A protein sequence alignment comparison of the P1-P2 read-through proteins of BMYV and TuYV was performed, the proteins encoded by ORF-1 and ORF-2 that are the precursor to the VPg. This sequence comparison showed a sequence identity of 49.1 % and a similarity of 64.4 %. This suggests that the proteins are only around 50 % similar to each other, however this analysis does not take into account the post translational modifications of these proteins so more investigation into the similarity of these proteins is required before any conclusions drawn (EMBL-EBI, EMBOSS Water alignment).

It is also possible that poleroviruses are able to use cap-independent translation due to the presence of an internal ribosomal entry site (IRES) in the genome of *Potato leafroll virus* (Jaag et al., 2003). IRESs often require interaction with the eIF4G protein in order to recruit ribosomal subunits, but the IRES of *Potato leafroll virus* looks like it interacts directly with the host ribosome, although this has not been fully understood (Jaag et al., 2003). There is a problem with this strategy in poleroviruses. Cap-independent translation, is usually initiated in the 5' UTR, or in the 3' UTR (Kneller et al., 2006). Neither of these mechanisms has so far been identified in poleroviruses, and instead the IRES site of *Potato leafroll virus* is located at the end of ORF-1 (Jaag et al., 2003). This therefore does not explain how the ORF-0 or the majority of the ORF-1 proteins would be expressed. Further investigation of the polerovirus IRES, and the role of eIF4E and VPg are required to understand the mechanism being used to replicate poleroviruses. If it is shown that the polerovirus does in fact contain an IRES that allows the functional expression of the virus genome, as appears common in single stranded RNA viruses (Kneller et al., 2006), and the IRES was interacting with eIF4G, we would expect to find that the deletion of this protein would lead to the plant becoming resistant to the virus. This was not the case, and deletion of eIF4G did not effect virus accumulation in the plant (Kneller et al., 2006). If the IRES is able to directly interact with the ribosome itself, the knock-out of eIF4G should have no effect on virus infectivity.

It is possible eIF4E does not play a role in translation of the virus, and there is evidence to suggest that it plays an important role in virus cell-to-cell movement. It has been frequently shown that mammalian eIF4E proteins are responsible for trafficking specific mRNA sequences through the nuclear pores into the cell cytoplasm (Rousseau et al., 1996; Topisirovic et al., 2003), and recent evidence suggests that the potyvirus *Pea seed borne mosaic virus* (PSBMV) is using eIF4E in a similar way. The translocation of plant viruses through cells is not particularly well understood but is thought to comprise a mixture of host proteins and virally encoded proteins. Viral proteins that are known to play a role in movement include the coat proteins and VPg. Gao et al. (2004) proposed that the host eIF4E protein interacts with the viral VPg in order to assist in viral movement, although the exact mechanism of this, is not understood. This idea was proposed after it was discovered that PSBMV was unable to spread around plants containing mutations in eIF4E (Gao et al., 2004). It has been proposed that eIF4E aids in transport of viruses by its strong association with eIF4G, which in turn can interact with microtubules. Microtubules have previously been implicated with virus spread, but as of yet this interaction for virus movement has not been shown (Gao et al., 2004; Lellis et al., 2002).

Future work in this area would need to investigate the polerovirus IRES sequence already identified in *Potato leafroll virus* in order to establish if eIF4G is required to initiate translation. A study of virus accumulation in the cell, using protoplasts, would further our understanding of eIF4E's role in resistance to the BMYV. If the virus was not able to accumulate in a cell that lacks eIF4E, then it could be concluded that eIF4E is required for BMYV replication and accumulation within a cell. A further study of virus movement around the plant would discern whether eIF4E was an important factor in cell-to-cell movement of BMYV. A deeper understanding of this mechanism is required, as not much evidence of the mechanism of virus movement involving eIF4E is available, and a better understanding of the role of eIF4E in relation to poleroviruses is urgently required.

Now that eIF4E has been identified as an important infection factor for BMYV this knowledge can be put into use in crop plants. Methods such as targeting induced local lesions in genomes (TILLING) enable plants that are mutated in specific genes to be identified. This process would be made more complicated by the fact that many sugar beet varieties are autotetraploids ($2n = 4x = 36$), and would require a mutation in each duplicated copy of eIF4E. This would probably require extensive back-crossing in order to ensure only mutant eIF4E was present in the plant and reduce the chance of any other mutations occurring within the plant. The plant would then have to be tested for susceptibility to BMYV, and to ensure that any mutations do not detrimentally affect root yield. Even after this process has been performed, it is not guaranteed that this gene will give the same resistance in sugar beet plants as observed in Arabidopsis. It has previously been shown that *Tobacco etch virus* (TEV) and *Lettuce mosaic virus*, although requiring eIF(iso)4E to infect Arabidopsis, actually required eIF4E in order to infect pepper, tomato and lettuce (Duprat et al., 2002; Lellis et al., 2002). This process is a slow and long one, with no guaranteed agronomic outcome. This information however, will be useful to plant breeders in order to try and create BMYV genetically resistant sugar beet crops.

7.2 TuYV infection studies

The results of BMYV infection analysis in this study clearly show that disruption to *eIF4E*, either by a natural mutation, a truncated protein, or complete knock-out, causes plants to lose susceptibility to the virus. Previous evidence in other plant viruses suggest that closely related plant viruses often require the same, or similar host proteins to infect plants (Kneller et al., 2006; Robaglia and Caranta, 2006). As BMYV and TuYV are closely related poleroviruses, it stands to reason that they may require similar host translation factors in order to infect a plant. Infection studies of TuYV in plants containing T-DNA knock-outs in *eIF4E* and *eIF(iso)4E* were therefore performed. These experiments showed that even with truncated, mutated, or the absence of either of these proteins, TuYV was still able to successfully infect cells (Figure 6.3), although the subsequent discovery that the line SALK113327C may still

be segregating for the presence of a T-DNA insertion in *eIF(iso)4E* means that no firm conclusions can be drawn from the work with this line, or the other *eIF(iso)4E* T-DNA insertion lines as they were also segregating.

The evidence from this study strongly suggests that the removal or mutation of eIF4E (and possibly eIF(iso)4E) alone is not capable of producing resistant plants. It is possible that TuYV, if it is infecting plants by a similar method to BMYV, can use either isoform of eIF4E, as seen in other viruses (Duprat et al., 2002). It is predicted that an *eIF4E* and *eIF(iso)4E* T-DNA double knock out line would not be viable, so instead lines containing the Sna-1 version of *eIF4E* alleles and T-DNA insertion in *eIF(iso)4E* were created. These lines named UEA1, UEA2 and UEA3 (Section 6.2.1) were created from the F2 segregants of a JIC62 x SALK_113327C cross. The genotypes of these lines have yet to be verified so work still needs to be carried out to confirm the relevance of the findings from this cross. The F_1 generation of the UEA lines were used in infection studies. All the UEA lines (as well as the parent plants) were susceptible to TuYV. This could mean one of three possibilities; firstly the T-DNA insertion in *eIF(iso)4E* may not be a 'true' insertion (this needs to be verified), secondly the TuYV may use different eIF proteins in order to infect Arabidopsis, or thirdly the insert repeat found in the Sna-1 version of *eIF4E* does not disrupt TuYV infection in the same way it disrupts BMYV. Further investigation into a possible interaction between TuYV and eIF4E (both the Col-0 and the Sna-1 versions of the protein) and the eIF(iso)4E would be useful. This could be carried out using yeast-two hybrid analysis to find out if there is an interaction between the eIF4E isoforms and the VPg of TuYV. If no interaction is seen then further evidence is gained to imply TuYV uses a different infection mechanism to BMYV. Another method to investigate this interaction could use co-imunoprecipitation of the virus VPg and eIF4E proteins *in planta*. If antibodies were made to the VPg of the virus, proteins interactions with the VPg could be investigated (Weigel and Glazebrook, 2002).

Soon after the completion of these experiments, a paper was published indicating that instead of the predicted eIF4E protein being a requirement for

TuYV infection, a different family of protein was responsible for TuYV infection, the eIF(iso)4G family (Reinbold et al., 2013). The eIF(iso)4G1 interaction would be unusual because previously published evidence of such closely related viruses requiring a different family of protein in order to infect plants is unprecedented. With this knowledge, T-DNA knock out lines in the same *eIF4G* genes as published were ordered in order to find out if the UK-BB TuYV isolate acted in the same way as the TuYV-FL1 isolate used in their study. Published results indicated that a T-DNA insertion in *eIF(iso)4G1* reduced the accumulation of TuYV four-fold in comparison to Col-0 (Reinbold et al., 2013). These findings could not be replicated within this study, as no significant difference in the levels of virus accumulation could be detected between SALK_009905C and Col-0 (Figure 6.5, Table 6.3). Reinbold et al. (2013) suggested that TuYV may be able to utilise either eIF(iso)4G1 or eIF(iso)4G, but preferentially selected eIF(iso)4G1. One explanation of results seen here could be that the UK-BB TuYV isolate is able to use both the eIF(iso)4G proteins equally, so removal of either protein has no effect on virus accumulation. This may be possible because there is evidence to suggest that other viruses are able to use both isoforms of eIF proteins to complete the virus lifecycle (Duprat et al., 2002). Reinbold et. al. (2013) used a double knock-out plant *eIF(iso)4G1G2*, which was found to contain extremely low levels of virus. This plant also had obvious growth deficiencies that would not be viable in a crop plant. A good way to test if TuYV was using these isoforms would be to find an Arabidopsis ecotype containing natural variation in at least one of these sequences. Crosses and infection studies could then be performed in a similar method to those described here using eIF4E and eIF(iso)4E (Section6.2.1).

This study was not able to identify any single protein that was able to cause loss of susceptibility in a plant to TuYV. The fact that Reinbold et al., were able to identify eIF(iso)4G1 as a potential target of TuYV infection brings up interesting questions about the sequences of the two virus isolates used in these studies. A comparative sequence analysis between the two isolates (TuYV-FL1 from France and the UK-BB TuYV isolate) would be interesting and reveal distinct differences inn these viruses. If the sequences were

significantly different, which is unlikely, identifying key elements of the virus genome that differ between the isolates would help identify areas of the virus genome important for infection of plants, and the method of infection.

Another interesting point to consider if TuYV uses eIF4G instead of eIF4E in order to infect plant cells, the eIF of plants may be important determinants of host range of these viruses. TuYV and BMYV, although having overlapping host ranges, also have very distinct plants that they are able to infect. For example BMYV is a prolific infector of sugar beet, whilst TuYV is unable to infect sugar beet. This theory was previously suggested in the case of *Melon necrotic spot virus* and its ability to infect *Nicotiana benthamiana* using eIF4E (Nieto et al., 2011; Reinbold et al., 2013).

7.3 Synthesising an infectious clone of TuYV

One of the original aims of this study was to develop an infectious TuYV clone, capable of being delivered by *A. tumefaciens* into plants in order to study infection. This has previously been carried out using several poleroviruses, but most recently by Percival-Alwyn (2010) with BMYV. The use of infectious clones would mean that aphid inoculations would no longer be required, and the study of virus infection would become easier to perform. This part of the study has not been discussed in this thesis because of the difficulties faced in the cloning. The initial aim was to clone TuYV in two sections, into the plasmid pGreen (compatible with *E. coli* and *A. tumefaciens*). This process would also use overlap PCR in order to introduce the *Cauliflower mosaic virus* (CaMV) 35S promoter, and the *Agrobacterium nopaline synthetase (nos)* transcription terminator sequence alongside a hammerhead ribosome sequence (*ribo*) designed to cleave any non-viral RNA, and any 3' polyadenylated sequences (Leiser et al., 1992; Percival-Alwyn, 2010). A diagram of the cloning pathway is shown in Figure 7.1.

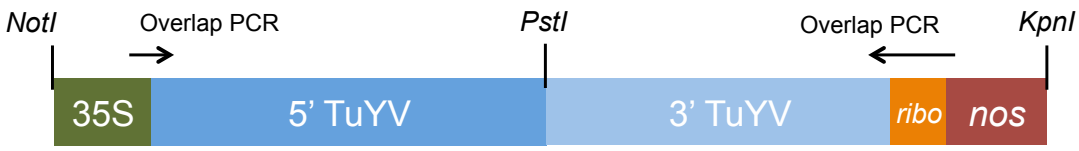


Figure 7.1 CAMV 35S promoter is cloned upstream of the TuYV sequence (blue) using overlap PCR to connect the 35S and 5' TuYV fragments. The *ribo-nos* terminator sequences are cloned downstream of the TuYV sequence, and connected to the 3'TuYV sequence. The two DNA fragments would then be digested with *PstI* and ligated together. The fragment is inserted into a plasmid vector using directional cloning with the use of the *NotI* and *KpnI* digestion sites found in the 35S promoter and *nos-ribo* terminal sequences respectively.

This is similar to the method described by Percival-Alwyn (2010) where two virus fragments were cloned into the pGreen plasmid vector downstream of the strong CaMV 35S promoter sequence, and upstream of the *nos-ribo* terminator sequence.

The reference sequence used for the identification of restriction enzyme sites, and to design primers was the TuYV-FL1 strain. It was soon apparent that this sequence was significantly different to the sequence of the UK Brooms Barn TuYV isolate, as predicted restriction target sites were not present in the TuYV sequence, and extra sites were found in several places within the sequence. For these reasons the cloning strategy was modified to amplify three fragments of TuYV instead of two. These fragments were named 5', mid, and 3' in accordance to their positioning within the TuYV genome, and a schematic diagram of the proposed cloning technique is seen in Figure 7.2. Each of the three TuYV fragments were first cloned in order to insert them into a plasmid so that a bacterial stock containing sections of viral DNA could be maintained.

Figure 7.2 Second cloning plan of TuYV. The CAMV 35S promoter is cloned upstream of the TuYV sequence (blue) using overlap PCR to connect the 35S and a small section of the 5' TuYV fragment. This overlapped fragment would then be ligated to the larger 5' fragment. The *ribo-nos* terminator sequences are cloned downstream of the TuYV sequence, and Connected to the 3'TuYV sequence. The full fragment, when completed would be inserted into a plasmid vector using directional cloning with the use of the *NotI* and *KpnI* digestion sites found in the 35S promoter and *nos-ribo* terminal sequences respectively.

It was possible to amplify TuYV cDNA from RNA of infected leaf extract, but only at very low concentrations, the three amplified fragments of TuYV are shown in Figure 7.3, alongside primers used in Table 7.1.

Table 7.1 Primer sets used in the creation of TuYV, 35S and *nos-ribo* DNA fragments for cloning.

Primer Name	Sequence (5'-3')	Polarity
TuYV 5' F (J07)	GAATTCGGATCCTTGGAGAGGACAAAAGA AACCAGGAGGGAATCC	F
TuYV 5' Pst/EcoRI (J16)	TGAAGAGCCTGCGACTGCAGGCTTC	R
TuYV Mid Forward (J28)	TCCGGAAGCCCCTACCTT	F
TuYV Mid Reverse (J29)	TGGGTTGTGGAGAGGGAGAA	R
TuYV 3' R (J31)	GAATTCGGTACCACACCGAACTCGGCTAG GGATTT	R
TuYV 3' Forward New (G65)	CGAGGACCAATTCAGGATCC	F
Upstream 35S	GTTGAAGATGCCTCTGCC	F
35S – BMYV	TGGTTTCTTTTGTCCTCTCCAAATGAAATGA ACTTCC	R
RiboTuYV3'	ACGGACTCATCAGTAGACATGTGA ATCAT GTCTAGACACCGAAGTGCCGTAGG	R
M13 Forward (A11)	GTAAAACGACGGCCAGT	F

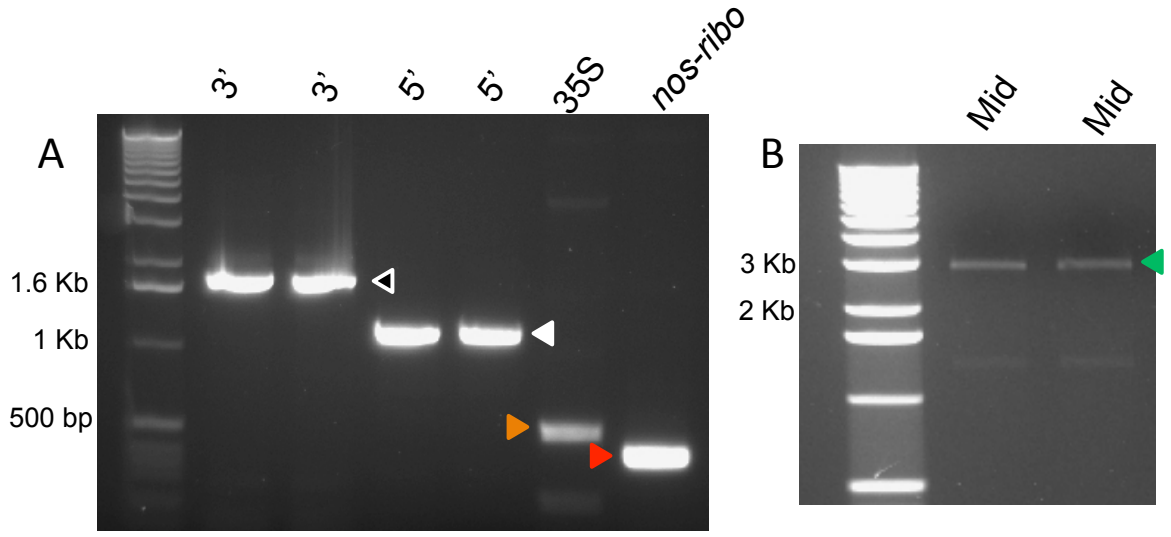


Figure 7.3. PCR amplified cDNA fragments of TuYV, and PCR amplified fragments of CamV 35S and *nos-ribo*. The cDNA amplification of 3' TuYV fragments (shown in image A, indicated with a black arrow, white boarder) were 1.6 Kb, 5' TuYV fragments (shown in image A, indicated with a white arrow) were 1.1 Kb, and the Mid TuYV fragments (shown in image B, indicated with a green arrow) were 3 Kb. All of these TuYV fragments were created from reverse transcription of RNA extracted from leaf sap. The 35S DNA fragment (shown in image A indicated with an orange arrow), at 450 bp and the *nos-ribo* fragment (shown in image A, indicated with a red arrow) at 300 bp were amplified from a previous successful infections clone, pSLJ4K1 cassette.

Two thirds of the virus was able to be cloned into plasmid vectors (5' and mid sections) and transformed into *E. coli*, but overlap of the 5' sequence and the 35S CaMV promoter has proved difficult. The 3' sequence has not successfully been cloned into *E. coli*. The 3' fragment has been PCR amplified, and appears to be able to ligate into plasmid vectors (seen by DNA gel electrophoresis), but *E. coli* does not appear to be able to survive containing this fragment of DNA. Other investigations by scientists at the John Innes Centre (personal communication) attempted to clone the TuYV coat protein gene (located in the 3' section of the RNA) and were abandoned due to the difficulties encountered in the cloning steps (data unpublished). It is currently unknown why such difficulties are being faced with this sequence, as other poleroviruses have been able to be cloned in this manner. Attempted sequencing of the TuYV Brooms Barn isolate is required in order to fully

understand any sequence divergence from the current TuYV reference sequence, although this may prove difficult unless direct sequencing of the cDNA was attempted.

7.4 Summary

This study has identified a susceptibility factor, eIF4E, required for BMYV infection of *A. thaliana* through the study of natural variation. The mutation of the *eIF4E* gene has been shown to create BMYV resistant plants, which can be functionally complemented with the addition of the susceptible version of *eIF4E.* It has also been shown that TuYV, a closely related polerovirus, does not use the same eIF4E mechanism to infect plants, and evidence from this study suggests that it is likely to be able to use more than one eIF protein isoform, whether that be eIF4E or eIF4G has yet not been fully determined for the Brooms Barn TuYV isolate.

The cloning of TuYV to create an *Agrobacterium* mediated infectious clone has so far proved unsuccessful due to unforeseen sequence variation between the UK-BB TuYV isolate, and the reference sequence isolate. Problems have also been encountered in cloning the 3' end of the virus, which has also been encountered by other groups.

There is still a lot of work to be done to fully understand polerovirus modes of infection. Further infection studies of selected mutants, and the discovery of other naturally occurring ecotypes will aid understanding of processes of infection, and association between viral and host proteins. The successful creation of a virus clone of TuYV will support this investigation by making infection processes easier, cheaper, faster and more reliable to perform. The information gathered from this study implicating eIF4E as a susceptibility factor for BMYV infection could allow future breeding programs of sugar beet to analyse sequence variation in this gene. This could aid the creation of durable genetic resistance to BMYV and reduce the yield damaging effect of this virus on one of the UKs most important arable crop species.

Appendix

A. Sna-1 *elf4E* nucleotide sequence (Exons shown in yellow)

ATGGCGGTAGAAGACACTCCCAAATCTGTTGTAACGGAAGAAGCTAAGCCTAATTCA
ATAGAGAATCCGATTGATCGATACCATGAGGAAGGTGATGATGCCGAAGAAGGAGAG
ATCGCCGGAGGAGAAGGAGACGGAAACGTTGACGAATCGAGCAAATCCGGTGTTCCT
GAATCGCATCCTCTGGAACATTCATGGACTTTCTGGTTCGATAATCCTGCTGTGAAA
TCGAAACAAACCTCTTGGGGAAGTTCCTTGCGACCCGTGTTTACGTTTTCAACTGTT
GAGGAATTTTGGAGGTTAGGTTTTTGATTTTATTTTATTTCCGACTCAATATCTGGT
TTGTTCAATTATTCTGCATCTGGGTTTTGTTATAGGTTTCGATTTGTTGAGGAAAGT
TATGTTCTTTATTGGGGGATTAGAAGATCCCATTGAAGTCATTATATGTTTTTGATG
AATTGCTATGTTTGGTGTTTGAATTCGTAGCTAAAGCTTATGTTAGGGTTTAGCTTT
GATATTCTGTTCACTTGTTTGTGAAGTAAAGTAGAAGAGCAAAGTTTGTGAGAGAAA
GAGACCAACTTTGAAATCTTTCTTAGTGGTTTTCTAGGTATCAGAATTTGAGCAAAC
ACTTTCTTGATGACTGAGGGTTGATGTTGTATAGTTCTTGCTCTTCCCAATGAGATT
CATAGGTTTGTGTATTGTTCTTTCGACTTCTTATTTTAAAAGACATTTTGGTTTGCA
GTTTGTACAACAACATGAAGCATCCGAGCAAGTTAGCTCACGGAGCTGACTTCTACT
GTTTCAAACACATCATTGAACCTAAGTGGGAGGATCCTATTTGTGCTAATGGAGGAA
AATGGACTATGACTTTCCCTAAGGAGAAGTCTGATAAGAGCTGGCTCTACACGGTAC
GGTTTCTATTCTTCTTTTATTTTGACTCGTAACTCCTGCGTCATCATCCAATTGAAT
CTCACCGGTTTTTCTTTTTACATGCTTGGTTTAGTTGCTTGCATTGATTGGAGAGCA
GTTTGATCATGGAGATGAAATATGTGGAGCAGTTGTCAACATTAGAGGAAAGCAAGA
AAGGATATCTATTTGGACTAAAAATGCTTCAAACGAAGCTGCTCAGGTAAATAGAAA
AGACCTTTCTCATCAAAACTCATCAAACAGTCTTCTCTGTGTAAAAATAAGACTTTA
ATTCTCGTCTGCATCAAATGTTGCAGGTGAGCATTGGAAAACAATGGAAGGAGTTTC
TCGATTACAACAACAGCATAGGTTTCATCATCCATGTAAGAAGAAAGCTTCTCATGA
TTCTAATTCAAAAGTCTTCATTTTCTTCAGATCTCTCATTGTTTGGTTTGATTTCTT
TTCTTTCAGGAGGATGCGAAGAAGCTCGACAGGAAGCTCGACAGGAATGCAAAGAAC
GCTTACACCGCTTGA

B. Sna-1 and Col-0 *eIF4E* nucleotide sequence alignment (Exons shown in yellow)

```
Sna-1EIF4E          1 ATGGCGGTAGAAGACACTCCCAAATCTGTTGTAACGGAAGAAGCTAAGCC     50
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E          1 ATGGCGGTAGAAGACACTCCCAAATCTGTTGTAACGGAAGAAGCTAAGCC     50

Sna-1EIF4E         51 TAATTCAATAGAGAATCCGATTGATCGATACCATGAGGAAGGTGATGATG     100
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E         51 TAATTCAATAGAGAATCCGATTGATCGATACCATGAGGAAGGTGATGATG     100

Sna-1EIF4E        101 CCGAAGAAGGAGAGATCGCCGGAGGAGAAGGAGACGGAAACGTTGACGAA     150
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E        101 CCGAAGAAGGAGAGATCGCCGGAGGAGAAGGAGACGGAAACGTTGACGAA     150

Sna-1EIF4E        151 TCGAGCAAATCCGGTGTTCCTGAATCGCATCCTCTGGAACATTCATGGAC     200
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E        151 TCGAGCAAATCCGGTGTTCCTGAATCGCATCCTCTGGAACATTCATGGAC     200

Sna-1EIF4E        201 TTTCTGGTTCGATAATCCTGCTGTGAAATCGAAACAAACCTCTTGGGGAA     250
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E        201 TTTCTGGTTCGATAATCCTGCTGTGAAATCGAAACAAACCTCTTGGGGAA     250

Sna-1EIF4E        251 GTTCCTTGCGACCCGTGTTTACGTTTTCAACTGTTGAGGAATTTTGGAGG     300
                      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E        251 GTTCCTTGCGACCCGTGTTTACGTTTTCAACTGTTGAGGAATTTTGGAGG     300

Sna-1EIF4E        301 TTAGGTTTTTGATTTTATTTTATTTCCGACTCAATATCTGGTTTGTTCAA     350
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E        301 TTAGGTTTTTGATTTTATTTTATTTCCGACTCAATATCTGGTTTGTTCAA     350

Sna-1EIF4E        351 TTATTCTGCATCTGGGTTTTGTTATAGGTTTCGATTTGTTGAGGAAAGTT     400
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E        351 TTATTCTGCATCTGGGTTTTGTTATAGGTTTCGATTTGTTGAGGAAAGTT     400

Sna-1EIF4E        401 ATGTTCTTTATTGGGGGATTAGAAGATCCCATTGAAGTCATTATATGTTT     450
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E        401 ATGTTCTTTATTGGGGGATTAGAAGATCCCATTGAAGTCATTATATGTTT     450

Sna-1EIF4E        451 TTGATGAATTGCTATGTTTGGTGTTTGAATTCGTAGCTAAAGCTTATGTT     500
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E        451 TTGATGAATTGCTATGTTTGGTGTTTGAATTCGTAGCTAAAGCTTATGTT     500

Sna-1EIF4E        501 AGGGTTTAGCTTTGATATTCTGTTCACTTGTTTGTGAAGTAAAGTAGAAG     550
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E        501 AGGGTTTAGCTTTGATATTCTGTTCACTTGTTTGTGAAGTAAAGTAGAAG     550

Sna-1EIF4E        551 AGCAAAGTTTGTGAGAGAAAGAGACCAACTTTGAAATCTTTCTTAGTGGT     600
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E        551 AGCAAAGTTTGTGAGAGAAAGAGACCAACTTTGAAATCTTTCTTAGTGGT     600

Sna-1EIF4E        601 TTTCTAGGTATCAGAATTTGAGCAAACACTTTCTTGATGACTGAGGGTTG     650
                      |||||||||||||||||||||||||||||||.||||||||||||||||||
Col-0EIF4E        601 TTTCTAGGTATCAGAATTTGAGCAAACACTTTCGTGATGACTGAGGGTTG     650

Sna-1EIF4E        651 ATGTTGTATAGTTCTTGCTCTTCCCAATGAGATTCATAGGTTTGTGTATT     700
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E        651 ATGTTGTATAGTTCTTGCTCTTCCCAATGAGATTCATAGGTTTGTGTATT     700

Sna-1EIF4E        701 GTTCTTTCGACTTCTTATTTTAAAAGACATTTTGGTTTGCAGTTTGTACA     750
                      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E        701 GTTCTTTCGACTTCTTATTTTAAAAGACATTTTGGTTTGCAGTTTGTACA     750

Sna-1EIF4E        751 ACAACATGAAGCATCCGAGCAAGTTAGCTCACGGAGCTGACTTCTACTGT     800
                      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E        751 ACAACATGAAGCATCCGAGCAAGTTAGCTCACGGAGCTGACTTCTACTGT     800
```

```
Sna-1EIF4E       801 TTCAAACACATCATTGAACCTAAGTGGGAGGATCCTATTTGTGCTAATGG    850
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E       801 TTCAAACACATCATTGAACCTAAGTGGGAGGATCCTATTTGTGCTAATGG    850

Sna-1EIF4E       851 AGGAAAATGGACTATGACTTTCCCTAAGGAGAAGTCTGATAAGAGCTGGC    900
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E       851 AGGAAAATGGACTATGACTTTCCCTAAGGAGAAGTCTGATAAGAGCTGGC    900

Sna-1EIF4E       901 TCTACACGGTACGGTTTCTATTCTTCTTTTATTTTGACTCGTAACTCCTG    950
                     |||||||.||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E       901 TCTACACTGTACGGTTTCTATTCTTCTTTTATTTTGACTCGTAACTCCTG    950

Sna-1EIF4E       951 CGTCATCATCCAATTGAATCTCACCGGTTTTTCTTTTTACATGCTTGGTT   1000
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E       951 CGTCATCATCCAATTGAATCTCACCGGTTTTTCTTTTTACATGCTTGGTT   1000

Sna-1EIF4E      1001 TAGTTGCTTGCATTGATTGGAGAGCAGTTTGATCATGGAGATGAAATATG   1050
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E      1001 TAGTTGCTTGCATTGATTGGAGAGCAGTTTGATCATGGAGATGAAATATG   1050

Sna-1EIF4E      1051 TGGAGCAGTTGTCAACATTAGAGGAAAGCAAGAAAGGATATCTATTTGGA   1100
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E      1051 TGGAGCAGTTGTCAACATTAGAGGAAAGCAAGAAAGGATATCTATTTGGA   1100

Sna-1EIF4E      1101 CTAAAAATGCTTCAAACGAAGCTGCTCAGGTAAATAGAAAAGACCTTTCT   1150
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E      1101 CTAAAAATGCTTCAAACGAAGCTGCTCAGGTAAATAGAAAAGACCTTTCT   1150

Sna-1EIF4E      1151 CATCAAAACTCATCAAACAGTCTTCTCTGTGTAAAAATAAGACTTTAATT   1200
                     |||||||||||||||||||||| |||||||||||||||||||||||||||||
Col-0EIF4E      1151 CATCAAAACTCATCAAACAGTC-TCTCTGTGTAAAAATAAGACTTTAATT   1199

Sna-1EIF4E      1201 CTCGTCTGCATCAAATGTTGCAGGTGAGCATTGGAAAACAATGGAAGGAG   1250
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E      1200 CTCGTCTGCATCAAATGTTGCAGGTGAGCATTGGAAAACAATGGAAGGAG   1249

Sna-1EIF4E      1251 TTTCTCGATTACAACAACAGCATAGGTTTCATCATCCATGTAAGAAGAAA   1300
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E      1250 TTTCTCGATTACAACAACAGCATAGGTTTCATCATCCATGTAAGAAGAAA   1299

Sna-1EIF4E      1301 GCTTCTCATGATTCTAATTCAAAAGTCTTCATTTTCTTCAGATCTCTCAT   1350
                     |||||||||.||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E      1300 GCTTCTCATAATTCTAATTCAAAAGTCTTCATTTTCTTCAGATCTCTCAT   1349

Sna-1EIF4E      1351 TGTTTGGTTTGATTTC-TTTTCTTTCAGGAGGATGCGAAGAAGCTCGACA   1399
                     ||||||||||||.|||| ||||.||||||||||||||||
Col-0EIF4E      1350 TGTTTGGTTTGGTTTCTTTTTTTTTCAGGAGGATGCGAA-----------   1388

Sna-1EIF4E      1400 GGAAGCTCGACAGGAATGCAAAGAACGCTTACACCGCTTGA           1440
                       |||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4E      1389 -GAAGCTCGACAGGAATGCAAAGAACGCTTACACCGCTTGA           1428
```

C. Sna-1 eIF4E predicted amino acid sequence

MAVEDTPKSVVTEEAKPNSIENPIDRYHEEGDDAEEGEIAGGEGDGNVDESSKSGVP
ESHPLEHSWTFWFDNPAVKSKQTSWGSSLRPVFTFSTVEEFWSLYNNMKHPSKLAHG
ADFYCFKHIIEPKWEDPICANGGKWTMTFPKEKSDKSWLYTLLALIGEQFDHGDEIC
GAVVNIRGKQERISIWTKNASNEAAQVSIGKQWKEFLDYNNSIGFIIHEDAKKLDRK
LDRNAKNAYTA

D. Sna-1 *eIF4G* nucleotide sequence (Exons shown in yellow)

CATTCAATAAACAACAAAGTTTTTTTTGTATCCATATTTCATGTACTAAGTCTCCACAAATTATAGCCA
AATTCACACGTGTTAATATTTTATCAATGAAGTTAATTTAAATTTACATTTATTTGGTTTTTAGTGCGG
AAAATTCAGCTTTTTGTTATTGTTTTGTCAACGCATTATGTTTTATCTCTTCAAATGATAGTAAAAAAT
AATATTTTAATTAAAATTAAAATTAAAAATTTGGCTTAAAAAAGTTTGTTGAGTTTTTTGGGGTGACCA
CCTGAAACCTCTCTTTGCACATTGACCGGAAAAAAAGTTTTTTGATTGAGTAACGAGATAAACTTAATG
AGTTGTCGAATTTACCCCTAAGACATAGCGTAATTTACTAATCAAGTACTTTAACCAGACTCTTTCACT
AATACAAAATCTCTATATTTTTTTTTTCTTTCTTTAATTTCCAATTTATTTTTTGATTACACATAGTTT
CTCTTCATTCATCTCTTTCTCTCCCCGTCTTTTCTTGTTCCTATAATCATGTCCGTCCAATAAAACCCT
AATTTTATCTTTCCTTGCGTCGTTTCACTCCATTACAGAGTTGTAGAATTTAGCTTGCCAAAAAGTGTG
AAGATTAGGGTTTGTTCTAGATGTCCTACAATCAATCCAGACCCGACAGAAGCGAGACTCAATATCGTA
GAACTGGTCGATCCACCGGTAACCAACAACAACAACAACAACACCGATCTTCTTCCGCCGCCGGTTACG
GTAAGGGCGCCGGCGCTCCTGGTTCTGCGCCTGCCCCTTCCACTTATCCTGATAATTCTTCCTTGTCTT
CCAATCGCAGGTTTTCATTTCAAATTCGACTTTTTTTTTTTTGGATGATTCTCTGAGCCATTATTGGAC
ATGATTAATTATGCAAAGTTTGATCCTTTATTGTTATTTATGACAGTTTTAAGAAGCCCGGCAATGCTC
AAGGAGGAGGGCAGCCTCGGGTGAATCTGCCACCTGTGAATCATCCTAATAATCACAACAATGGTCCCA
ATGCTCACTCTCGCTCTCAAGGTATGATCGTTATTCACGGAGTGCCGCAAAGTCTTTAAATTTGTATAA
TTTCAGATTAATTTGGTCAGTGTGTGAGTGATCGTTTGGTTTAGTTCAATCTATCTTACACTGTTAGTT
ACAGGAGAACCGGGTGTTGGTGGACCAACCAATCCAACTGAATCGTTCAACAGAAACACCGGACCTATT
CCAAAGGCTCCAACTTCTCAGTCTACCGTCATGAGTTCCAAGATCAATGAGACGCCCAACACAGCTAAA
GGTAGGTTTTTGTTGCATCGTCTATTTTTTTGTATGCATACTGAGCTTTGAAGTTGAACTAAATTCCTC
TGTAATTCCATAAAGGTATAAGTTTTATAATGTTTATTTTGTTTTGTTTAGTGGCAGCCTCTGGAGACG
CTTCTCAGGCATTTCCTCTCCAGTTTGGGTCACTTGGTCCTGATTTGATGGTAATGCTGTTGTTCCTTC
CTTTGTTTGTTTGATTTTTCCATCAACACCTGAATAATCTTCCCTGTTGGTTTTCTAGGTTCCTGCTCG
AACTACCTCAGCACCTCCGAATATGGATGACCAGAAACGTGCCCAGGTGGAACACCTTTTAGTATTATG
CATCTGCCGTTACTGTGATTGTTGCCTACATAAATAATTCTTTCTTATAGGATGATGAATAGTCTGTGT
CTCGAACTTTGCTTCTGACTTTTAAATGCTCTATCATTTGTCTTGTATTGCGATTTGCTGGCCTGATTG
CTTGATCTTTAACTCAAATTGGTTTAAAGAATATATACTGAAACTTAAGATGTCTTTTATTCACTTTGT
TCTGAACTTGACCTTTCGTTTATGACACGAGCAAACATTGTTTTGCAGATGCAGCAATCTTCTTTAAGA
ACGGCGTCAAATGTGCCAGCTTCTGTACCCAAAAAAGATTCATCAAATAAGGGTGCAGATAATCAATTG
ATGAGGAAAGAGGGGCACAATCCATCGAGTGAAAAAGCTGATATCCAAGTCCCACATATAGCCCCTCCA
AGTCAAACGCAGAAGTCTCCAATTACAAATATTCGCATGCCTTCTGTGCAGACACCATATCAGCATACT
CAGGTCCCTCACCCTGTACATTTTGGTGGGCCGAATATGCATATGCAGCCTCCCGTGACTGCAACCTCG
TTTCAGATGCCAATGCCAATGGCATTATCTATGGGAAATACTCCTCAAATCCCGCCGCAGGTGTTTTAT
CAGGGACATCCACCACATCCGATGCATCATCAGGGTATGATGCATCAGGCTCAGGGACATGGTTTTGCA
ACTCCAATGGGTGCTCAGATTCATCCTCAGTTAGGCCATGTGGGTGTGGGTTTGAGCCCTCAGTATCCC
CAGCAGCAAGGTGGAAAATATGGTGGGGCACGCAAGACCACCCCTGTAAAGATTACACATCCTGACACA
CACGAAGAGCTGAGGCTTGATCGACGTGGTGACCCGTATTCAGAAGGCGATTCAACGGCTTTAAAACCA
CATTCTAACCCACCTCCCAGATCACAGCCAGTCTCATCATTTGCTCCAAGACCAGTCAATTTGGTGCAA
CCCTCATATAACTCCAATACCATGATATATCCCCCGGTTTCGGTACCGTTAAATAATGGTCCAATGTCA
TCCGCTCAGGCACCGAGATATCATTACCCAGTTATTGATGGGTCTCAGAGAGTACAACTTATCAACCAA
CCTGCTCATACTGCTCCACAGCTTATCAGACCCGCTGCTCCTGCACATCTTTCCTCTGATTCGACTTCC
TCTGTGAAAGCACGCAATGCCCAAAATGTAATGTCATCTGCTCTACCTGTAAATGCGAAGGTATCAGTG
AAGCCAGCTGGGGTTTCTGAAAAGCTTGGATCACCAAAAGACAGGTCACATGGAGAAGTTAACATTTCT
CTGTCACAAAAGAACGTGGAGGCATGTTCGTTGAGCTCTTCCCAGCAGCCGAAACCTAGCTTTGTCTCT
GGAGTACCAAATTCGTCTGCTCCGCCAGCAAAGTCGCCTGTGGAGACTGTTCCGCTAGCAAAGTCGTCT
GTGGAGACTGTTCCGCCAGTAAAGTCGTCTGTGGAGACTGCTCCAGTTACAACGACTGAAATCAGAAGA
GCGGAAATGGTGAGTGAGTCGATCTCAGTTGAAGATCAGACATGTAAGGTGGAACCCCCTCATAATCTG
ACTGAGGTATGATACTGTGTTTTGTTTTTGGGTATTATTCATTAATTCTTTTTCTTATTTGATTAAATT
ATTTTCTTTTTTTGGCTGGTTCACTCAGAATCGTGGACAGACTATGCCAGACTCTCTGGTCTCTGATCC
TGAAACAGCAACCGTTGCTGCCAAGGAAAATTTATCACTCCCAGCTACCAACGGGTTTAGGAAGCAACT
CCTGAAGGTGTCTACTACATCTGATGCTCCAACTTCTGACTCAGTAGATACAAGTATTGACAAATCTAC
GGAAGGTTCAAGCCATGCCTCATCGGAGATTTCTGGTTCTTCACCGCAAGAGAAAGACCTAAAATGTGA
TAACCGGACTGCTTCTGACAAGCTCGATGAAAGGTCTGTAATTTCTGATGCAAAACACGAAACACTGTC
AGGTGTGCTTGAGAAGGCACAGAATGAGGTAGATGGTGCCACAGATGTCTGTCCTGTCTCTGAAAAACT
AGCTGTTACAGATGATACGAGCTCTGACCTTCCACATTCTACTCATGTTCTGTCTTCTACTGTTCCTCT
TGGACATTCGGAAACACATAAATCTGCTGTTGAAACAAACACGAGAAGAAATACTTCTACAAAAGGAAA
GAAGAAGATAAAAGAAATCCTTCAAAAAGCAGATGCTGCAGGGACAACTTCTGATCTCTATATGGCTTA
CAAAGGGCCTGAGGAAAAGAAAGAGAGCTCAAATGTTGTTCATGATGTTTCGAACCAGAACCTGTTACC

TGCCATACCTCAGGCTGTTGAAGCCATTGTGGATACTGAACCAGTGAAAAATGAACCAGAAGACTGGGA
AGATGCAGCCGATGTTTCTACACCAAAGCTGGAAACTGCAGATAATTCTGTGAATGCAAAGAGAGGTTC
CTCAGATGAGGTCAGCGACAACTGCATCAATACAGAAAAGAAGTACTCCCGGGATTTCCTCCTAAAGTT
TGCAGACCTGTGTACTGCTCTTCCTGAGGGATTTGACGTTTCGCCTGATATTGCTAATGCCTTGATTGT
TGCATATATGGGTGCATCACATCATGAACATGATTCATATCCTACTCCTGGAAAGGTTATGGATCGCCA
AGCAAGTGGTGCTCGTTTAGATCGCCGTCCCAGCAACGTGGCTGGTGATGATAGATGGACGAAGAATCA
GGGTTCTCTTCCAGCAGGATATGGGGGTAACGTAGGTTTCCGACCTGGTCAAGGAGGAAACTCGGGAGT
TTTAAGAAACCCTCGTATGCAGGGACCAATTATATCTAGACCGATGCAACCTGTGGGTCCTATGGGAGG
AATGGGTAGAAATACCCCCGACTTAGAAAGGTGGCAACGTGGTTCAAATTTCCAACAAAAAGGACTTTT
TCCTTCTCCGCACACTCCTATGCAAGTGATGCACAAAGCCGAGAGAAAATACCAAGTGGGGACAATTGC
AGATGAAGAACAAGCAAAACAAAGGCAGTTAAAGAGCATCCTGAACAAGTTGACCCCACAAAACTTTGA
GAAACTGTTTGAGCAAGTTAAAAGTGTCAACATTGACAACGCTGTTACACTTTCTGGTGTCATTTCACA
GATATTTGACAAAGCCTTGATGGAGCCAACATTCTGTGAGATGTATGCAGATTTCTGTTTTCATCTCTC
TGGGGCGTTACCTGATTTTAATGAGAATGGTGAAAAGATTACCTTCAAAAGATTGCTTCTCAATAAATG
TCAGGAAGAATTCGAGAGGGGGGAGAAAGAAGAGGAGGAAGCCAGTAGAGTTGCCGAAGAAGGTCAAGT
AGAACAAACCGAGGAGGAAAGGGAAGAGAAAAGACTTCAGGTGCGAAGGAGAATGCTTGGTAACATCAG
ACTTATTGGTGAGTTATACAAGAAAAGGATGTTGACTGAGAAAATCATGCACGCATGCATCCAGAAGTT
GCTCGGGTATAATCAAGATCCACATGAAGAGAATATTGAAGCTCTGTGTAAACTAATGAGTACGATAGG
AGTTATGATCGATCACAACAAAGCTAAGTTCCAGATGGATGGATATTTTGAGAAAATGAAAATGCTATC
ATGCAAACAAGAATTGTCTTCTAGGGTGAGGTTCATGTTGATCAATGCCATCGATCTGAGAAAGAACAA
ATGGCAGGAGAGAATGAAGGTCGAAGGGCCGAAAAAAATTGAGGAAGTGCACAGAGATGCTGCACAAGA
ACGCCAAACTCAAGCGAATAGGCTTTCACGTGGACCCTCAATGAATTCGTCAGGAAGAAGAGGGCATAT
GGAGTTTAGTAGTCCTAGGGGAGGAGGAGGAATGCTATCACCTCCAGCTGCCCAAATGGGTAGTTACCA
TGGACCACCTCAAGGTCGTGGCTTTAGTAATCAGGACATTCGATTTGATGACAGGCCATCTTATGAGCC
TAGGATGGTTCCAATGCCGCAAAGGTCAGTATGTGAGGAGCCTATTACCTTGGGTCCGCAAGGTGGTCT
TGGTCAGGGAATGTCTATTAGAAGGCCTGCAGTAGCATCAAACACTTATCAGTCTGATGCTACTCAGGC
CGGTGGTGGAGATTCTAGGCGACCGGCCGGTGGTTTGAATGGTTTTGGCTCACATAGACCTGCAAGTCC
TGTTACTCACGGACGGTCAAGCTTTCAAGAGCGGGGAACAGCTTATGTTCATAGGGAATTTGCAAGTCT
GTCGCGTGCTTCTGATCTGTCACCAGAAGTTTCGTCCGCTAGGCAAGTACTACAAGGGCCATCAGCTAC
AGTAAACAGTCCTCGAGAAAATGCTTTGTCTGAAGAACAGTTAGAGAATCTGTCATTGTCCGCAATTAA
GGAATATTACAGGTACTATATCTCTCCTTTCTTGCTGGTCATTTGTTTCATTCTTCGCAAAGTCATCAG
ATACTGCCACATTTAGGAAATTATTCTGTGGGAACCTGTGTAGGTATAATTGGATCAAATTACAGTTCA
TTCGTTTACAGTTAGAGCGTTTGTTTACAACGACAGATCATCTGATAAAGTGTATATTTCTTGAAGTAT
GTAAATTAGAATTTGCTCCAAACAAAACGGCTTAAAAATACATATGAGAGACTTTGTTCTTTGACTTTA
ATAGAATCTCTAATACAGGGGTAAACTCTGTTTTTTTTACTGACTGTGCGTTTTATGTATGATGTTAGT
TTCTATTCGTTTACCCTCTTTATTATTGATAAGCACTCTGGTGCGTTCCAATGTGTTCTTGTGCAGTGC
CCGAGATGAGAATGAGATTGGTATGTGCATGAAAGATATGAATTCACCAGCTTACCACCCAACAATGAT
TTCTCTCTGGGTAACTGATTCGTTTGAGAGAAAAGACAAAGAAAGGGATCTCTTAGCAAAGCTCCTTGT
GAACCTCGTGAAATCTGCTGACAACGCCTTAAACGAAGTCCAGCTAGTGAAAGGGTAAGTAAACAAAGC
CCATAAATCTTGAAGCCCTAGTCTAATCAAACGTCTGTTGCTTAAATCTTTTACTGCTTTTTTTTTCAC
AGGTTTGAATCGGTTTTGAAAACCCTGGAGGATGCAGTAAATGATGCTCCAAAAGCAGCAGAGTTTCTT
GGTAGAATATTTGGGAAAAGTGTGACAGAGAAAGTAGTGACATTGACAGAGATTGGTCGGTTAATCCAG
GAAGGAGGAGAAGAACCAGGAAGTCTGATAGAGTTTGGATTAGGCGGCGATGTTCTTGGGAGTGTTTTG
GAGATGATAAAAACAGAAGCTGGAGAAGAAACGTTGGTTGAGATTCGCCGGAGCTCAGGTCTGAGGATT
GAAAATTTCAAACCTCATGCACCTAACCGGTCTAAGATATTAGAGAAATTTACTTAGGAAAAAAAAATG
GAACCATCTTTTGGGTTCCTTTCTTCTTCTCTTTTTTTGTTTCTCTCTTAAAAGTCTTTTCTCTTTTCA
AGTGCTTCAAACAAAACTAATTTGTTATAAAGGGAGTTTCTCTATTTTATTATATAGCAAAAAACTTCC
AAAATTTCTCATTTCTGTGTTTAACCTTTTCGTACATCAGTTTTAAAGCACAGAGAGCTCAATGTTCTT
CCAATATCGTTATTAATAAATTTTGATTAAATTCAATCAAATCGGAGTTATATTACCACATGAGGTTAA
AGGGCCATATTAAAAAGTCTGCACTTCATATGAGCAACAAGGCTTTTATGTCTTTATGGTTGATTTGAT
GGCCCATATATGATAGTTCAAAGGCCCATATTAAAAAATGCCCTAACA

E. Sna-1 and Col-0 *eIF4G* nucleotide alignment (Exons shown in yellow)

```
Sna-1EIF4G       534 GTCCGTCCAATAAAACCCTAATTTTATCTTTCCTTGCGTCGTTTCACTCC    583
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G         1 GTCCGTCCAATAAAACCCTAATTTTATCTTTCCTTGCGTCGTTTCACTCC     50

Sna-1EIF4G       584 ATTACAGAGTTGTAGAATTTAGCTTGCCAAAAAGTGTGAAGATTAGGGTT    633
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G        51 ATTACAGAGTTGTAGAATTTAGCTTGCCAAAAAGTGTGAAGATTAGGGTT    100

Sna-1EIF4G       634 TGTTCTAGATGTCCTACAATCAATCCAGACCCGACAGAAGCGAGACTCAA    683
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       101 TGTTCTAGATGTCCTACAATCAATCCAGACCCGACAGAAGCGAGACTCAA    150

Sna-1EIF4G       684 TATCGTAGAACTGGTCGATCCACCGGTAACCAACAACAACAACAACAACA    733
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       151 TATCGTAGAACTGGTCGATCCACCGGTAACCAACAACAACAACAACAACA    200

Sna-1EIF4G       734 CCGATCTTCTTCCGCCGCCGGTTACGGTAAGGGCGCCGGCGCTCCTGGTT    783
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       201 CCGATCTTCTTCCGCCGCCGGTTACGGTAAGGGCGCCGGCGCTCCTGGTT    250

Sna-1EIF4G       784 CTGCGCCTGCCCCTTCCACTTATCCTGATAATTCTTCCTTGTCTTCCAAT    833
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       251 CTGCGCCTGCCCCTTCCACTTATCCTGATAATTCTTCCTTGTCTTCCAAT    300

Sna-1EIF4G       834 CGCAGGTTTTCATTTCAAATTCGACTTTTTTTTTTTTTGGATGATTCTCTG    883
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       301 CGCAGGTTTTCATTTCAAATTCGACTTTTTTTTTTTTTGGATGATTCTCTG    350

Sna-1EIF4G       884 AGCCATTATTGGACATGATTAATTATGCAAAGTTTGATCCTTTATTGTTA    933
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       351 AGCCATTATTGGACATGATTAATTATGCAAAGTTTGATCCTTTATTGTTA    400

Sna-1EIF4G       934 TTTATGACAGTTTTAAGAAGCCCGGCAATGCTCAAGGAGGAGGGCAGCCT    983
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       401 TTTATGACAGTTTTAAGAAGCCCGGCAATGCTCAAGGAGGAGGGCAGCCT    450

Sna-1EIF4G       984 CGGGTGAATCTGCCACCTGTGAATCATCCTAATAATCACAACAATGGTCC   1033
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       451 CGGGTGAATCTGCCACCTGTGAATCATCCTAATAATCACAACAATGGTCC    500

Sna-1EIF4G      1034 CAATGCTCACTCTCGCTCTCAAGGTATGATCGTTATTCACGGAGTGCCGC   1083
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       501 CAATGCTCACTCTCGCTCTCAAGGTATGATCGTTATTCACGGAGTGCCGC    550

Sna-1EIF4G      1084 AAAGTCTTTAAATTTGTATAATTTCAGATTAATTTGGTCAGTGTGTGAGT   1133
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       551 AAAGTCTTTAAATTTGTATAATTTCAGATTAATTTGGTCAGTGTGTGAGT    600

Sna-1EIF4G      1134 GATCGTTTGGTTTAGTTCAATCTATCTTACACTGTTAGTTACAGGAGAAC   1183
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       601 GATCGTTTGGTTTAGTTCAATCTATCTTACACTGTTAGTTACAGGAGAAC    650

Sna-1EIF4G      1184 CGGGTGTTGGTGGACCAACCAATCCAACTGAATCGTTCAACAGAAACACC   1233
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       651 CGGGTGTTGGTGGACCAACCAATCCAACTGAATCGTTCAACAGAAACACC    700

Sna-1EIF4G      1234 GGACCTATTCCAAAGGCTCCAACTTCTCAGTCTACCGTCATGAGTTCCAA   1283
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       701 GGACCTATTCCAAAGGCTCCAACTTCTCAGTCTACCGTCATGAGTTCCAA    750

Sna-1EIF4G      1284 GATCAATGAGACGCCCAACACAGCTAAAGGTAGGTTTTTGTTGCATCGTC   1333
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       751 GATCAATGAGACGCCCAACACAGCTAAAGGTAGGTTTTTGTTGCATCGTC    800
```

```
Sna-1EIF4G      1334 TATTTTTTTGTATGCATACTGAGCTTTGAAGTTGAACTAAATTCCTCTGT   1383
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       801 TATTTTTTTGTATGCATACTGAGCTTTGAAGTTGAACTAAATTCCTCTGT    850

Sna-1EIF4G      1384 AATTCCATAAAGGTATAAGTTTTATAATGTTTATTTTGTTTTGTTTAGTG   1433
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       851 AATTCCATAAAGGTATAAGTTTTATAATGTTTATTTTGTTTTGTTTAGTG    900

Sna-1EIF4G      1434 GCAGCCTCTGGAGACGCTTCTCAGGCATTTCCTCTCCAGTTTGGGTCACT   1483
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       901 GCAGCCTCTGGAGACGCTTCTCAGGCATTTCCTCTCCAGTTTGGGTCACT    950

Sna-1EIF4G      1484 TGGTCCTGATTTGATGGTAATGCTGTTGTTCCTTCCTTTGTTTGTTTGAT   1533
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G       951 TGGTCCTGATTTGATGGTAATGCTGTTGTTCCTTCCTTTGTTTGTTTGAT   1000

Sna-1EIF4G      1534 TTTTCCATCAACACCTGAATAATCTTCCCTGTTGGTTTTCTAGGTTCCTG   1583
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1001 TTTTCCATCAACACCTGAATAATCTTCCCTGTTGGTTTTCTAGGTTCCTG   1050

Sna-1EIF4G      1584 CTCGAACTACCTCAGCACCTCCGAATATGGATGACCAGAAACGTGCCCAG   1633
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1051 CTCGAACTACCTCAGCACCTCCGAATATGGATGACCAGAAACGTGCCCAG   1100

Sna-1EIF4G      1634 GTGGAACACCTTTTAGTATTATGCATCTGCCGTTACTGTGATTGTTGCCT   1683
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1101 GTGGAACACCTTTTAGTATTATGCATCTGCCGTTACTGTGATTGTTGCCT   1150

Sna-1EIF4G      1684 ACATAAATAATTCTTTCTTATAGGATGATGAATAGTCTGTGTCTCGAACT   1733
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1151 ACATAAATAATTCTTTCTTATAGGATGATGAATAGTCTGTGTCTCGAACT   1200

Sna-1EIF4G      1734 TTGCTTCTGACTTTTAAATGCTCTATCATTTGTCTTGTATTGCGATTTGC   1783
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1201 TTGCTTCTGACTTTTAAATGCTCTATCATTTGTCTTGTATTGCGATTTGC   1250

Sna-1EIF4G      1784 TGGCCTGATTGCTTGATCTTTAACTCAAATTGGTTTAAAGAATATATACT   1833
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1251 TGGCCTGATTGCTTGATCTTTAACTCAAATTGGTTTAAAGAATATATACT   1300

Sna-1EIF4G      1834 GAAACTTAAGATGTCTTTTATTCACTTTGTTCTGAACTTGACCTTTCGTT   1883
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1301 GAAACTTAAGATGTCTTTTATTCACTTTGTTCTGAACTTGACCTTTCGTT   1350

Sna-1EIF4G      1884 TATGACACGAGCAAACATTGTTTTGCAGATGCAGCAATCTTCTTTAAGAA   1933
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1351 TATGACACGAGCAAACATTGTTTTGCAGATGCAGCAATCTTCTTTAAGAA   1400

Sna-1EIF4G      1934 CGGCGTCAAATGTGCCAGCTTCTGTACCCAAAAAAGATTCATCAAATAAG   1983
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1401 CGGCGTCAAATGTGCCAGCTTCTGTACCCAAAAAAGATTCATCAAATAAG   1450

Sna-1EIF4G      1984 GGTGCAGATAATCAATTGATGAGGAAAGAGGGGCACAATCCATCGAGTGA   2033
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1451 GGTGCAGATAATCAATTGATGAGGAAAGAGGGGCACAATCCATCGAGTGA   1500

Sna-1EIF4G      2034 AAAAGCTGATATCCAAGTCCCACATATAGCCCCTCCAAGTCAAACGCAGA   2083
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1501 AAAAGCTGATATCCAAGTCCCACATATAGCCCCTCCAAGTCAAACGCAGA   1550

Sna-1EIF4G      2084 AGTCTCCAATTACAAATATTCGCATGCCTTCTGTGCAGACACCATATCAG   2133
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1551 AGTCTCCAATTACAAATATTCGCATGCCTTCTGTGCAGACACCATATCAG   1600
```

```
Sna-1EIF4G      2134 CATACTCAGGTCCCTCACCCTGTACATTTTGGTGGGCCGAATATGCATAT   2183
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1601 CATACTCAGGTCCCTCACCCTGTACATTTTGGTGGGCCGAATATGCATAT   1650

Sna-1EIF4G      2184 GCAGCCTCCCGTGACTGCAACCTCGTTTCAGATGCCAATGCCAATGGCAT   2233
                     ||||.||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1651 GCAGACTCCCGTGACTGCAACCTCGTTTCAGATGCCAATGCCAATGGCAT   1700

Sna-1EIF4G      2234 TATCTATGGGAAATACTCCTCAAATCCCGCCGCAGGTGTTTTATCAGGGA   2283
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1701 TATCTATGGGAAATACTCCTCAAATCCCGCCGCAGGTGTTTTATCAGGGA   1750

Sna-1EIF4G      2284 CATCCACCACATCCGATGCATCATCAGGGTATGATGCATCAGGCTCAGGG   2333
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1751 CATCCACCACATCCGATGCATCATCAGGGTATGATGCATCAGGCTCAGGG   1800

Sna-1EIF4G      2334 ACATGGTTTTGCAACTCCAATGGGTGCTCAGATTCATCCTCAGTTAGGCC   2383
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1801 ACATGGTTTTGCAACTCCAATGGGTGCTCAGATTCATCCTCAGTTAGGCC   1850

Sna-1EIF4G      2384 ATGTGGGTGTGGGTTTGAGCCCTCAGTATCCCCAGCAGCAAGGTGGAAAA   2433
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1851 ATGTGGGTGTGGGTTTGAGCCCTCAGTATCCCCAGCAGCAAGGTGGAAAA   1900

Sna-1EIF4G      2434 TATGGTGGGGCACGCAAGACCACCCCTGTAAAGATTACACATCCTGACAC   2483
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1901 TATGGTGGGGCACGCAAGACCACCCCTGTAAAGATTACACATCCTGACAC   1950

Sna-1EIF4G      2484 ACACGAAGAGCTGAGGCTTGATCGACGTGGTGACCCGTATTCAGAAGGCG   2533
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      1951 ACACGAAGAGCTGAGGCTTGATCGACGTGGTGACCCGTATTCAGAAGGCG   2000

Sna-1EIF4G      2534 ATTCAACGGCTTTAAAACCACATTCTAACCCACCTCCCAGATCACAGCCA   2583
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2001 ATTCAACGGCTTTAAAACCACATTCTAACCCACCTCCCAGATCACAGCCA   2050

Sna-1EIF4G      2584 GTCTCATCATTTGCTCCAAGACCAGTCAATTTGGTGCAACCCTCATATAA   2633
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2051 GTCTCATCATTTGCTCCAAGACCAGTCAATTTGGTGCAACCCTCATATAA   2100

Sna-1EIF4G      2634 CTCCAATACCATGATATATCCCCCGGTTTCGGTACCGTTAAATAATGGTC   2683
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2101 CTCCAATACCATGATATATCCCCCGGTTTCGGTACCGTTAAATAATGGTC   2150

Sna-1EIF4G      2684 CAATGTCATCCGCTCAGGCACCGAGATATCATTACCCAGTTATTGATGGG   2733
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2151 CAATGTCATCCGCTCAGGCACCGAGATATCATTACCCAGTTATTGATGGG   2200

Sna-1EIF4G      2734 TCTCAGAGAGTACAACTTATCAACCAACCTGCTCATACTGCTCCACAGCT   2783
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2201 TCTCAGAGAGTACAACTTATCAACCAACCTGCTCATACTGCTCCACAGCT   2250

Sna-1EIF4G      2784 TATCAGACCCGCTGCTCCTGCACATCTTTCCTCTGATTCGACTTCCTCTG   2833
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2251 TATCAGACCCGCTGCTCCTGCACATCTTTCCTCTGATTCGACTTCCTCTG   2300

Sna-1EIF4G      2834 TGAAAGCACGCAATGCCCAAAATGTAATGTCATCTGCTCTACCTGTAAAT   2883
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2301 TGAAAGCACGCAATGCCCAAAATGTAATGTCATCTGCTCTACCTGTAAAT   2350

Sna-1EIF4G      2884 GCGAAGGTATCAGTGAAGCCAGCTGGGGTTTCTGAAAAGCTTGGATCACC   2933
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2351 GCGAAGGTATCAGTGAAGCCAGCTGGGGTTTCTGAAAAGCTTGGATCACC   2400
```

```
Sna-1EIF4G      2934 AAAAGACAGGTCACATGGAGAAGTTAACATTTCTCTGTCACAAAAGAACG   2983
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2401 AAAAGACAGGTCACATGGAGAAGTTAACATTTCTCTGTCACAAAAGAACG   2450

Sna-1EIF4G      2984 TGGAGGCATGTTCGTTGAGCTCTTCCCAGCAGCCGAAACCTAGCTTTGTC   3033
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2451 TGGAGGCATGTTCGTTGAGCTCTTCCCAGCAGCCGAAACCTAGCTTTGTC   2500

Sna-1EIF4G      3034 TCTGGAGTACCAAATTCGTCTGCTCCGCCAGCAAAGTCGCCTGTGGAGAC   3083
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2501 TCTGGAGTACCAAATTCGTCTGCTCCGCCAGCAAAGTCGCCTGTGGAGAC   2550

Sna-1EIF4G      3084 TGTTCCGCTAGCAAAGTCGTCTGTGGAGACTGTTCCGCCAGTAAAGTCGT   3133
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2551 TGTTCCGCTAGCAAAGTCGTCTGTGGAGACTGTTCCGCCAGTAAAGTCGT   2600

Sna-1EIF4G      3134 CTGTGGAGACTGCTCCAGTTACAACGACTGAAATCAGAAGAGCGGAAATG   3183
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2601 CTGTGGAGACTGCTCCAGTTACAACGACTGAAATCAGAAGAGCGGAAATG   2650

Sna-1EIF4G      3184 GTGAGTGAGTCGATCTCAGTTGAAGATCAGACATGTAAGGTGGAACCCCC   3233
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2651 GTGAGTGAGTCGATCTCAGTTGAAGATCAGACATGTAAGGTGGAACCCCC   2700

Sna-1EIF4G      3234 TCATAATCTGACTGAGGTATGATACTGTGTTTTGTTTTTGGGTATTATTC   3283
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2701 TCATAATCTGACTGAGGTATGATACTGTGTTTTGTTTTTGGGTATTATTC   2750

Sna-1EIF4G      3284 ATTAATTCTTTTTCTTATTTGATTAAATTATTTTCTTTTTTTGGCTGGTT   3333
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2751 ATTAATTCTTTTTCTTATTTGATTAAATTATTTTCTTTTTTTGGCTGGTT   2800

Sna-1EIF4G      3334 CACTCAGAATCGTGGACAGACTATGCCAGACTCTCTGGTCTCTGATCCTG   3383
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2801 CACTCAGAATCGTGGACAGACTATGCCAGACTCTCTGGTCTCTGATCCTG   2850

Sna-1EIF4G      3384 AAACAGCAACCGTTGCTGCCAAGGAAAATTTATCACTCCCAGCTACCAAC   3433
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2851 AAACAGCAACCGTTGCTGCCAAGGAAAATTTATCACTCCCAGCTACCAAC   2900

Sna-1EIF4G      3434 GGGTTTAGGAAGCAACTCCTGAAGGTGTCTACTACATCTGATGCTCCAAC   3483
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2901 GGGTTTAGGAAGCAACTCCTGAAGGTGTCTACTACATCTGATGCTCCAAC   2950

Sna-1EIF4G      3484 TTCTGACTCAGTAGATACAAGTATTGACAAATCTACGGAAGGTTCAAGCC   3533
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      2951 TTCTGACTCAGTAGATACAAGTATTGACAAATCTACGGAAGGTTCAAGCC   3000

Sna-1EIF4G      3534 ATGCCTCATCGGAGATTTCTGGTTCTTCACCGCAAGAGAAAGACCTAAAA   3583
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3001 ATGCCTCATCGGAGATTTCTGGTTCTTCACCGCAAGAGAAAGACCTAAAA   3050

Sna-1EIF4G      3584 TGTGATAACCGGACTGCTTCTGACAAGCTCGATGAAAGGTCTGTAATTTC   3633
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3051 TGTGATAACCGGACTGCTTCTGACAAGCTCGATGAAAGGTCTGTAATTTC   3100

Sna-1EIF4G      3634 TGATGCAAAACACGAAACACTGTCAGGTGTGCTTGAGAAGGCACAGAATG   3683
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3101 TGATGCAAAACACGAAACACTGTCAGGTGTGCTTGAGAAGGCACAGAATG   3150

Sna-1EIF4G      3684 AGGTAGATGGTGCCACAGATGTCTGTCCTGTCTCTGAAAAACTAGCTGTT   3733
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3151 AGGTAGATGGTGCCACAGATGTCTGTCCTGTCTCTGAAAAACTAGCTGTT   3200
```

```
Sna-1EIF4G      3734 ACAGATGATACGAGCTCTGACCTTCCACATTCTACTCATGTTCTGTCTTC   3783
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3201 ACAGATGATACGAGCTCTGACCTTCCACATTCTACTCATGTTCTGTCTTC   3250

Sna-1EIF4G      3784 TACTGTTCCTCTTGGACATTCGGAAACACATAAATCTGCTGTTGAAACAA   3833
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3251 TACTGTTCCTCTTGGACATTCGGAAACACATAAATCTGCTGTTGAAACAA   3300

Sna-1EIF4G      3834 ACACGAGAAGAAATACTTCTACAAAAGGAAAGAAGAAGATAAAAGAAATC   3883
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3301 ACACGAGAAGAAATACTTCTACAAAAGGAAAGAAGAAGATAAAAGAAATC   3350

Sna-1EIF4G      3884 CTTCAAAAAGCAGATGCTGCAGGGACAACTTCTGATCTCTATATGGCTTA   3933
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3351 CTTCAAAAAGCAGATGCTGCAGGGACAACTTCTGATCTCTATATGGCTTA   3400

Sna-1EIF4G      3934 CAAAGGGCCTGAGGAAAAGAAAGAGAGCTCAAATGTTGTTCATGATGTTT   3983
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3401 CAAAGGGCCTGAGGAAAAGAAAGAGAGCTCAAATGTTGTTCATGATGTTT   3450

Sna-1EIF4G      3984 CGAACCAGAACCTGTTACCTGCCATACCTCAGGCTGTTGAAGCCATTGTG   4033
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3451 CGAACCAGAACCTGTTACCTGCCATACCTCAGGCTGTTGAAGCCATTGTG   3500

Sna-1EIF4G      4034 GATACTGAACCAGTGAAAAATGAACCAGAAGACTGGGAAGATGCAGCCGA   4083
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3501 GATACTGAACCAGTGAAAAATGAACCAGAAGACTGGGAAGATGCAGCCGA   3550

Sna-1EIF4G      4084 TGTTTCTACACCAAAGCTGGAAACTGCAGATAATTCTGTGAATGCAAAGA   4133
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3551 TGTTTCTACACCAAAGCTGGAAACTGCAGATAATTCTGTGAATGCAAAGA   3600

Sna-1EIF4G      4134 GAGGTTCCTCAGATGAGGTCAGCGACAACTGCATCAATACAGAAAAGAAG   4183
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3601 GAGGTTCCTCAGATGAGGTCAGCGACAACTGCATCAATACAGAAAAGAAG   3650

Sna-1EIF4G      4184 TACTCCCGGGATTTCCTCCTAAAGTTTGCAGACCTGTGTACTGCTCTTCC   4233
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3651 TACTCCCGGGATTTCCTCCTAAAGTTTGCAGACCTGTGTACTGCTCTTCC   3700

Sna-1EIF4G      4234 TGAGGGATTTGACGTTTCGCCTGATATTGCTAATGCCTTGATTGTTGCAT   4283
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3701 TGAGGGATTTGACGTTTCGCCTGATATTGCTAATGCCTTGATTGTTGCAT   3750

Sna-1EIF4G      4284 ATATGGGTGCATCACATCATGAACATGATTCATATCCTACTCCTGGAAAG   4333
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3751 ATATGGGTGCATCACATCATGAACATGATTCATATCCTACTCCTGGAAAG   3800

Sna-1EIF4G      4334 GTTATGGATCGCCAAGCAAGTGGTGCTCGTTTAGATCGCCGTCCCAGCAA   4383
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3801 GTTATGGATCGCCAAGCAAGTGGTGCTCGTTTAGATCGCCGTCCCAGCAA   3850

Sna-1EIF4G      4384 CGTGGCTGGTGATGATAGATGGACGAAGAATCAGGGTTCTCTTCCAGCAG   4433
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3851 CGTGGCTGGTGATGATAGATGGACGAAGAATCAGGGTTCTCTTCCAGCAG   3900

Sna-1EIF4G      4434 GATATGGGGGGTAACGTAGGTTTCCGACCTGGTCAAGGAGGAAACTCGGGA   4483
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3901 GATATGGGGGGTAACGTAGGTTTCCGACCTGGTCAAGGAGGAAACTCGGGA   3950

Sna-1EIF4G      4484 GTTTTAAGAAACCCTCGTATGCAGGGACCAATTATATCTAGACCGATGCA   4533
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      3951 GTTTTAAGAAACCCTCGTATGCAGGGACCAATTATATCTAGACCGATGCA   4000
```

```
Col-0EIF4G      4001 ACCTGTGGGTCCTATGGGAGGAATGGGTAGAAATACCCCCGACTTAGAAA   4050

Sna-1EIF4G      4584 GGTGGCAACGTGGTTCAAATTTCCAACAAAAAGGACTTTTTCCTTCTCCG   4633
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4051 GGTGGCAACGTGGTTCAAATTTCCAACAAAAAGGACTTTTTCCTTCTCCG   4100

Sna-1EIF4G      4634 CACACTCCTATGCAAGTGATGCACAAAGCCGAGAGAAAAATACCAAGTGGG   4683
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4101 CACACTCCTATGCAAGTGATGCACAAAGCCGAGAGAAAAATACCAAGTGGG   4150

Sna-1EIF4G      4684 GACAATTGCAGATGAAGAACAAGCAAAACAAAGGCAGTTAAAGAGCATCC   4733
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4151 GACAATTGCAGATGAAGAACAAGCAAAACAAAGGCAGTTAAAGAGCATCC   4200

Sna-1EIF4G      4734 TGAACAAGTTGACCCCACAAAACTTTGAGAAACTGTTTGAGCAAGTTAAA   4783
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4201 TGAACAAGTTGACCCCACAAAACTTTGAGAAACTGTTTGAGCAAGTTAAA   4250

Sna-1EIF4G      4784 AGTGTCAACATTGACAACGCTGTTACACTTTCTGGTGTCATTTCACAGAT   4833
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4251 AGTGTCAACATTGACAACGCTGTTACACTTTCTGGTGTCATTTCACAGAT   4300

Sna-1EIF4G      4834 ATTTGACAAAGCCTTGATGGAGCCAACATTCTGTGAGATGTATGCAGATT   4883
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4301 ATTTGACAAAGCCTTGATGGAGCCAACATTCTGTGAGATGTATGCAGATT   4350

Sna-1EIF4G      4884 TCTGTTTTCATCTCTCTGGGGCGTTACCTGATTTTAATGAGAATGGTGAA   4933
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4351 TCTGTTTTCATCTCTCTGGGGCGTTACCTGATTTTAATGAGAATGGTGAA   4400

Sna-1EIF4G      4934 AAGATTACCTTCAAAAGATTGCTTCTCAATAAATGTCAGGAAGAATTCGA   4983
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4401 AAGATTACCTTCAAAAGATTGCTTCTCAATAAATGTCAGGAAGAATTCGA   4450

Sna-1EIF4G      4984 GAGGGGGGAGAAAGAAGAGGAGGAAGCCAGTAGAGTTGCCGAAGAAGGTC   5033
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4451 GAGGGGGGAGAAAGAAGAGGAGGAAGCCAGTAGAGTTGCCGAAGAAGGTC   4500

Sna-1EIF4G      5034 AAGTAGAACAAACCGAGGAGGAAAGGGAAGAGAAAAGACTTCAGGTGCGA   5083
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4501 AAGTAGAACAAACCGAGGAGGAAAGGGAAGAGAAAAGACTTCAGGTGCGA   4550

Sna-1EIF4G      5084 AGGAGAATGCTTGGTAACATCAGACTTATTGGTGAGTTATACAAGAAAAG   5133
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4551 AGGAGAATGCTTGGTAACATCAGACTTATTGGTGAGTTATACAAGAAAAG   4600

Sna-1EIF4G      5134 GATGTTGACTGAGAAAATCATGCACGCATGCATCCAGAAGTTGCTCGGGT   5183
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4601 GATGTTGACTGAGAAAATCATGCACGCATGCATCCAGAAGTTGCTCGGGT   4650

Sna-1EIF4G      5184 ATAATCAAGATCCACATGAAGAGAATATTGAAGCTCTGTGTAAACTAATG   5233
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4651 ATAATCAAGATCCACATGAAGAGAATATTGAAGCTCTGTGTAAACTAATG   4700

Sna-1EIF4G      5234 AGTACGATAGGAGTTATGATCGATCACAACAAAGCTAAGTTCCAGATGGA   5283
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4701 AGTACGATAGGAGTTATGATCGATCACAACAAAGCTAAGTTCCAGATGGA   4750

Sna-1EIF4G      5284 TGGATATTTTGAGAAAATGAAAATGCTATCATGCAAACAAGAATTGTCTT   5333
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4751 TGGATATTTTGAGAAAATGAAAATGCTATCATGCAAACAAGAATTGTCTT   4800

Sna-1EIF4G      5334 CTAGGGTGAGGTTCATGTTGATCAATGCCATCGATCTGAGAAAGAACAAA   5383
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4801 CTAGGGTGAGGTTCATGTTGATCAATGCCATCGATCTGAGAAAGAACAAA   4850
```

```
Sna-1EIF4G      5384 TGGCAGGAGAGAATGAAGGTCGAAGGGCCGAAAAAAATTGAGGAAGTGCA   5433
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4851 TGGCAGGAGAGAATGAAGGTCGAAGGGCCGAAAAAAATTGAGGAAGTGCA   4900

Sna-1EIF4G      5434 CAGAGATGCTGCACAAGAACGCCAAACTCAAGCGAATAGGCTTTCACGTG   5483
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      4901 CAGAGATGCTGCACAAGAACGCCAAACTCAAGCGAATAGGCTTTCACGTG   4950

Sna-1EIF4G      5484 GACCCTCAATGAATTCGTCAGGAAGAAGAGGGCATATGGAGTTTAGTAGT   5533
                     |||||||||||||||||||||||||.||||||||||||||||||||||||||
Col-0EIF4G      4951 GACCCTCAATGAATTCGTCAGGAAGAAGGGGGCATATGGAGTTTAGTAGT   5000

Sna-1EIF4G      5534 CCTAGGGGAGGAGGAGGAATGCTATCACCTCCAGCTGCCCAAATGGGTAG   5583
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5001 CCTAGGGGAGGAGGAGGAATGCTATCACCTCCAGCTGCCCAAATGGGTAG   5050

Sna-1EIF4G      5584 TTACCATGGACCACCTCAAGGTCGTGGCTTTAGTAATCAGGACATTCGAT   5633
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5051 TTACCATGGACCACCTCAAGGTCGTGGCTTTAGTAATCAGGACATTCGAT   5100

Sna-1EIF4G      5634 TTGATGACAGGCCATCTTATGAGCCTAGGATGGTTCCAATGCCGCAAAGG   5683
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5101 TTGATGACAGGCCATCTTATGAGCCTAGGATGGTTCCAATGCCGCAAAGG   5150

Sna-1EIF4G      5684 TCAGTATGTGAGGAGCCTATTACCTTGGGTCCGCAAGGTGGTCTTGGTCA   5733
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5151 TCAGTATGTGAGGAGCCTATTACCTTGGGTCCGCAAGGTGGTCTTGGTCA   5200

Sna-1EIF4G      5734 GGGAATGTCTATTAGAAGGCCTGCAGTAGCATCAAACACTTATCAGTCTG   5783
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5201 GGGAATGTCTATTAGAAGGCCTGCAGTAGCATCAAACACTTATCAGTCTG   5250

Sna-1EIF4G      5784 ATGCTACTCAGGCCGGTGGTGGAGATTCTAGGCGACCGGCCGGTGGTTTG   5833
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5251 ATGCTACTCAGGCCGGTGGTGGAGATTCTAGGCGACCGGCCGGTGGTTTG   5300

Sna-1EIF4G      5834 AATGGTTTTGGCTCACATAGACCTGCAAGTCCTGTTACTCACGGACGGTC   5883
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5301 AATGGTTTTGGCTCACATAGACCTGCAAGTCCTGTTACTCACGGACGGTC   5350

Sna-1EIF4G      5884 AAGCTTTCAAGAGCGGGGAACAGCTTATGTTCATAGGGAATTTGCAAGTC   5933
                     |||...|||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5351 AAGTCCTCAAGAGCGGGGAACAGCTTATGTTCATAGGGAATTTGCAAGTC   5400

Sna-1EIF4G      5934 TGTCGCGTGCTTCTGATCTGTCACCAGAAGTTTCGTCCGCTAGGCAAGTA   5983
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5401 TGTCGCGTGCTTCTGATCTGTCACCAGAAGTTTCGTCCGCTAGGCAAGTA   5450

Sna-1EIF4G      5984 CTACAAGGGCCATCAGCTACAGTAAACAGTCCTCGAGAAAATGCTTTGTC   6033
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5451 CTACAAGGGCCATCAGCTACAGTAAACAGTCCTCGAGAAAATGCTTTGTC   5500

Sna-1EIF4G      6034 TGAAGAACAGTTAGAGAATCTGTCATTGTCCGCAATTAAGGAATATTACA   6083
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5501 TGAAGAACAGTTAGAGAATCTGTCATTGTCCGCAATTAAGGAATATTACA   5550

Sna-1EIF4G      6084 GGTACTATATCTCTCCTTTCTTGCTGGTCATTTGTTTCATTCTTCGCAAA   6133
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5551 GGTACTATATCTCTCCTTTCTTGCTGGTCATTTGTTTCATTCTTCGCAAA   5600

Sna-1EIF4G      6134 GTCATCAGATACTGCCACATTTAGGAAATTATTCTGTGGGAACCTGTGTA   6183
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5601 GTCATCAGATACTGCCACATTTAGGAAATTATTCTGTGGGAACCTGTGTA   5650
```

```
Sna-1EIF4G      6184 GGTATAATTGGATCAAATTACAGTTCATTCGTTTACAGTTAGAGCGTTTG   6233
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5651 GGTATAATTGGATCAAATTACAGTTCATTCGTTTACAGTTAGAGCGTTTG   5700

Sna-1EIF4G      6234 TTTACAACGACAGATCATCTGATAAAGTGTATATTTCTTGAAGTATGTAA   6283
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5701 TTTACAACGACAGATCATCTGATAAAGTGTATATTTCTTGAAGTATGTAA   5750

Sna-1EIF4G      6284 ATTAGAATTTGCTCCAAACAAAACGGCTTAAAAATACATATGAGAGACTT   6333
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5751 ATTAGAATTTGCTCCAAACAAAACGGCTTAAAAATACATATGAGAGACTT   5800

Sna-1EIF4G      6334 TGTTCTTTGACTTTAATAGAATCTCTAATACAGGGGTAAACTCTGTTTTT   6383
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5801 TGTTCTTTGACTTTAATAGAATCTCTAATACAGGGGTAAACTCTGTTTTT   5850

Sna-1EIF4G      6384 TTTACTGACTGTGCGTTTTATGTATGATGTTAGTTTCTATTCGTTTACCC   6433
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5851 TTTACTGACTGTGCGTTTTATGTATGATGTTAGTTTCTATTCGTTTACCC   5900

Sna-1EIF4G      6434 TCTTTATTATTGATAAGCACTCTGGTGCGTTCCAATGTGTTCTTGTGCAG   6483
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5901 TCTTTATTATTGATAAGCACTCTGGTGCGTTCCAATGTGTTCTTGTGCAG   5950

Sna-1EIF4G      6484 TGCCCGAGATGAGAATGAGATTGGTATGTGCATGAAAGATATGAATTCAC   6533
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      5951 TGCCCGAGATGAGAATGAGATTGGTATGTGCATGAAAGATATGAATTCAC   6000

Sna-1EIF4G      6534 CAGCTTACCACCCAACAATGATTTCTCTCTGGGTAACTGATTCGTTTGAG   6583
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6001 CAGCTTACCACCCAACAATGATTTCTCTCTGGGTAACTGATTCGTTTGAG   6050

Sna-1EIF4G      6584 AGAAAAGACAAAGAAAGGGATCTCTTAGCAAAGCTCCTTGTGAACCTCGT   6633
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6051 AGAAAAGACAAAGAAAGGGATCTCTTAGCAAAGCTCCTTGTGAACCTCGT   6100

Sna-1EIF4G      6634 GAAATCTGCTGACAACGCCTTAAACGAAGTCCAGCTAGTGAAAGGGTAAG   6683
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6101 GAAATCTGCTGACAACGCCTTAAACGAAGTCCAGCTAGTGAAAGGGTAAG   6150

Sna-1EIF4G      6684 TAAACAAAGCCCATAAATCTTGAAGCCCTAGTCTAATCAAACGTCTGTTG   6733
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6151 TAAACAAAGCCCATAAATCTTGAAGCCCTAGTCTAATCAAACGTCTGTTG   6200

Sna-1EIF4G      6734 CTTAAATCTTTTACTGCTTTTTTTTTCACAGGTTTGAATCGGTTTTGAAA   6783
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6201 CTTAAATCTTTTACTGCTTTTTTTTTCACAGGTTTGAATCGGTTTTGAAA   6250

Sna-1EIF4G      6784 ACCCTGGAGGATGCAGTAAATGATGCTCCAAAAGCAGCAGAGTTTCTTGG   6833
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6251 ACCCTGGAGGATGCAGTAAATGATGCTCCAAAAGCAGCAGAGTTTCTTGG   6300

Sna-1EIF4G      6834 TAGAATATTTGGGAAAAGTGTGACAGAGAAAGTAGTGACATTGACAGAGA   6883
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6301 TAGAATATTTGGGAAAAGTGTGACAGAGAAAGTAGTGACATTGACAGAGA   6350

Sna-1EIF4G      6884 TTGGTCGGTTAATCCAGGAAGGAGGAGAAGAACCAGGAAGTCTGATAGAG   6933
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6351 TTGGTCGGTTAATCCAGGAAGGAGGAGAAGAACCAGGAAGTCTGATAGAG   6400

Sna-1EIF4G      6934 TTTGGATTAGGCGGCGATGTTCTTGGGAGTGTTTTGGAGATGATAAAAAC   6983
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6401 TTTGGATTAGGCGGCGATGTTCTTGGGAGTGTTTTGGAGATGATAAAAAC   6450
```

```
Sna-1EIF4G      6984 AGAAGCTGGAGAAGAAACGTTGGTTGAGATTCGCCGGAGCTCAGGTCTGA   7033
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6451 AGAAGCTGGAGAAGAAACGTTGGTTGAGATTCGCCGGAGCTCAGGTCTGA   6500

Sna-1EIF4G      7034 GGATTGAAAATTTCAAACCTCATGCACCTAACCGGTCTAAGATATTAGAG   7083
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6501 GGATTGAAAATTTCAAACCTCATGCACCTAACCGGTCTAAGATATTAGAG   6550

Sna-1EIF4G      7084 AAATTTACTTAGGAAAAAAAAATGGAACCATCTTTTGGGTTCCTTTCTTC   7133
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6551 AAATTTACTTAGGAAAAAAAAATGGAACCATCTTTTGGGTTCCTTTCTTC   6600

Sna-1EIF4G      7134 TTCTCTTTTTTTGTTTCTCTCTTAAAAGTCTTTTCTCTTTTCAAGTGCTT   7183
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6601 TTCTCTTTTTTTGTTTCTCTCTTAAAAGTCTTTTCTCTTTTCAAGTGCTT   6650

Sna-1EIF4G      7184 CAAACAAAACTAATTTGTTATAAAGGGAGTTTCTCTATTTTATTATATAG   7233
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6651 CAAACAAAACTAATTTGTTATAAAGGGAGTTTCTCTATTTTATTATATAG   6700

Sna-1EIF4G      7234 CAAAAAACTTCCAAAATTTCTCATTTCTGTGTTTAACCTTTTCGTACATC   7283
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6701 CAAAAAACTTCCAAAATTTCTCATTTCTGTGTTTAACCTTTTCGTACATC   6750

Sna-1EIF4G      7284 AGTTTTAAAGCACAGAGAGCTCAATGTTCTTCCAATATCGTTATTAATAA   7333
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6751 AGTTTTAAAGCACAGAGAGCTCAATGTTCTTCCAATATCGTTATTAATAA   6800

Sna-1EIF4G      7334 ATTTTGATTAAATTCAATCAAATCGGAGTTATATTACCACATGAGGTTAA   7383
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6801 ATTTTGATTAAATTCAATCAAATCGGAGTTATATTACCACATGAGGTTAA   6850

Sna-1EIF4G      7384 AGGGCCATATTAAAAAGTCTGCACTTCATATGAGCAACAAGGCTTTTATG   7433
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6851 AGGGCCATATTAAAAAGTCTGCACTTCATATGAGCAACAAGGCTTTTATG   6900

Sna-1EIF4G      7434 TCTTTATGGTTGATTTGATGGCCCATATATGATAGTTCAAAGGCCCATAT   7483
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0EIF4G      6901 TCTTTATGGTTGATTTGATGGCCCATATATGATAGTTCAAAGGCCCATAT   6950

Sna-1EIF4G      7484 TAAAAAATGCCCTAAC   7499
                     ||||||||||||||||
Col-0EIF4G      6951 TAAAAAATGCCCTAAC   6966
```

F. Sna-1 eIF4G predicted amino acid sequence

MSYNQSRPDRSETQYRRTGRSTGNQQQQQQHRSSSAAGYGKGAGAPGSAPAPSTYPD
NSSLSSNRSFKKPGNAQGGGQPRVNLPPVNHPNNHNNGPNAHSRSQVTGEPGVGGPT
NPTESFNRNTGPIPKAPTSQSTVMSSKINETPNTAKVAASGDASQAFPLQFGSLGPD
LMVPARTTSAPPNMDDQKRAQMQQSSLRTASNVPASVPKKDSSNKGADNQLMRKEGH
NPSSEKADIQVPHIAPPSQTQKSPITNIRMPSVQTPYQHTQVPHPVHFGGPNMHMQP
PVTATSFQMPMPMALSMGNTPQIPPQVFYQGHPPHPMHHQGMMHQAQGHGFATPMGA
QIHPQLGHVGVGLSPQYPQQQGGKYGGARKTTPVKITHPDTHEELRLDRRGDPYSEG
DSTALKPHSNPPPRSQPVSSFAPRPVNLVQPSYNSNTMIYPPVSVPLNNGPMSSAQA
PRYHYPVIDGSQRVQLINQPAHTAPQLIRPAAPAHLSSDSTSSVKARNAQNVMSSAL
PVNAKVSVKPAGVSEKLGSPKDRSHGEVNISLSQKNVEACSLSSSQQPKPSFVSGVP
NSSAPPAKSPVETVPLAKSSVETVPPVKSSVETAPVTTTEIRRAEMVSESISVEDQT
CKVEPPHNLTENRGQTMPDSLVSDPETATVAAKENLSLPATNGFRKQLLKVSTTSDA
PTSDSVDTSIDKSTEGSSHASSEISGSSPQEKDLKCDNRTASDKLDERSVISDAKHE
TLSGVLEKAQNEVDGATDVCPVSEKLAVTDDTSSDLPHSTHVLSSTVPLGHSETHKS
AVETNTRRNTSTKGKKKIKEILQKADAAGTTSDLYMAYKGPEEKKESSNVVHDVSNQ
NLLPAIPQAVEAIVDTEPVKNEPEDWEDAADVSTPKLETADNSVNAKRGSSDEVSDN
CINTEKKYSRDFLLKFADLCTALPEGFDVSPDIANALIVAYMGASHHEHDSYPTPGK
VMDRQASGARLDRRPSNVAGDDRWTKNQGSLPAGYGGNVGFRPGQGGNSGVLRNPRM
QGPIISRPMQPVGPMGGMGRNTPDLERWQRGSNFQQKGLFPSPHTPMQVMHKAERKY
QVGTIADEEQAKQRQLKSILNKLTPQNFEKLFEQVKSVNIDNAVTLSGVISQIFDKA
LMEPTFCEMYADFCFHLSGALPDFNENGEKITFKRLLLNKCQEEFERGEKEEEEASR
VAEEGQVEQTEEEREEKRLQVRRRMLGNIRLIGELYKKRMLTEKIMHACIQKLLGYN
QDPHEENIEALCKLMSTIGVMIDHNKAKFQMDGYFEKMKMLSCKQELSSRVRFMLIN
AIDLRKNKWQERMKVEGPKKIEEVHRDAAQERQTQANRLSRGPSMNSSGRRGHMEFS
SPRGGGGMLSPPAAQMGSYHGPPQGRGFSNQDIRFDDRPSYEPRMVPMPQRSVCEEP
ITLGPQGGLGQGMSIRRPAVASNTYQSDATQAGGGDSRRPAGGLNGFGSHRPASPVT
HGRSSFQERGTAYVHREFASLSRASDLSPEVSSARQVLQGPSATVNSPRENALSEEQ
LENLSLSAIKEYYSARDENEIGMCMKDMNSPAYHPTMISLWVTDSFERKDKERDLLA
KLLVNLVKSADNALNEVQLVKG

G. Sna-1 and Col-0 eIF4G predicted amino acid alignment

```
Sna-1eIF4G         1 MSYNQSRPDRSETQYRRTGRSTGNQQQQQQHRSSSAAGYGKGAGAPGSAP     50
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G         1 MSYNQSRPDRSETQYRRTGRSTGNQQQQQQHRSSSAAGYGKGAGAPGSAP     50

Sna-1eIF4G        51 APSTYPDNSSLSSNRSFKKPGNAQGGGQPRVNLPPVNHPNNHNNGPNAHS    100
                     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G        51 APSTYPDNSSLSSNRSFKKPGNAQGGGQPRVNLPPVNHPNNHNNGPNAHS    100

Sna-1eIF4G       101 RSQVTGEPGVGGPTNPTESFNRNTGPIPKAPTSQSTVMSSKINETPNTAK    150
                     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G       101 RSQVTGEPGVGGPTNPTESFNRNTGPIPKAPTSQSTVMSSKINETPNTAK    150

Sna-1eIF4G       151 VAASGDASQAFPLQFGSLGPDLMVPARTTSAPPNMDDQKRAQMQQSSLRT    200
                     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G       151 VAASGDASQAFPLQFGSLGPDLMVPARTTSAPPNMDDQKRAQMQQSSLRT    200

Sna-1eIF4G       201 ASNVPASVPKKDSSNKGADNQLMRKEGHNPSSEKADIQVPHIAPPSQTQK    250
                     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G       201 ASNVPASVPKKDSSNKGADNQLMRKEGHNPSSEKADIQVPHIAPPSQTQK    250

Sna-1eIF4G       251 SPITNIRMPSVQTPYQHTQVPHPVHFGGPNMHMQPPVTATSFQMPMPMAL    300
                     |||||||||||||||||||||||||||||||||||.|||||||||||||||
Col-0eIF4G       251 SPITNIRMPSVQTPYQHTQVPHPVHFGGPNMHMQTPVTATSFQMPMPMAL    300

Sna-1eIF4G       301 SMGNTPQIPPQVFYQGHPPHPMHHQGMMHQAQGHGFATPMGAQIHPQLGH    350
                     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G       301 SMGNTPQIPPQVFYQGHPPHPMHHQGMMHQAQGHGFATPMGAQIHPQLGH    350

Sna-1eIF4G       351 VGVGLSPQYPQQQGGKYGGARKTTPVKITHPDTHEELRLDRRGDPYSEGD    400
                     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G       351 VGVGLSPQYPQQQGGKYGGARKTTPVKITHPDTHEELRLDRRGDPYSEGD    400

Sna-1eIF4G       401 STALKPHSNPPPRSQPVSSFAPRPVNLVQPSYNSNTMIYPPVSVPLNNGP    450
                     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G       401 STALKPHSNPPPRSQPVSSFAPRPVNLVQPSYNSNTMIYPPVSVPLNNGP    450

Sna-1eIF4G       451 MSSAQAPRYHYPVIDGSQRVQLINQPAHTAPQLIRPAAPAHLSSDSTSSV    500
                     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G       451 MSSAQAPRYHYPVIDGSQRVQLINQPAHTAPQLIRPAAPAHLSSDSTSSV    500

Sna-1eIF4G       501 KARNAQNVMSSALPVNAKVSVKPAGVSEKLGSPKDRSHGEVNISLSQKNV    550
                     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G       501 KARNAQNVMSSALPVNAKVSVKPAGVSEKLGSPKDRSHGEVNISLSQKNV    550

Sna-1eIF4G       551 EACSLSSSQQPKPSFVSGVPNSSAPPAKSPVETVPLAKSSVETVPPVKSS    600
                     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G       551 EACSLSSSQQPKPSFVSGVPNSSAPPAKSPVETVPLAKSSVETVPPVKSS    600

Sna-1eIF4G       601 VETAPVTTTEIRRAEMVSESISVEDQTCKVEPPHNLTENRGQTMPDSLVS    650
                     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G       601 VETAPVTTTEIRRAEMVSESISVEDQTCKVEPPHNLTENRGQTMPDSLVS    650

Sna-1eIF4G       651 DPETATVAAKENLSLPATNGFRKQLLKVSTTSDAPTSDSVDTSIDKSTEG    700
                     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G       651 DPETATVAAKENLSLPATNGFRKQLLKVSTTSDAPTSDSVDTSIDKSTEG    700

Sna-1eIF4G       701 SSHASSEISGSSPQEKDLKCDNRTASDKLDERSVISDAKHETLSGVLEKA    750
                     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G       701 SSHASSEISGSSPQEKDLKCDNRTASDKLDERSVISDAKHETLSGVLEKA    750

Sna-1eIF4G       751 QNEVDGATDVCPVSEKLAVTDDTSSDLPHSTHVLSSTVPLGHSETHKSAV    800
                     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G       751 QNEVDGATDVCPVSEKLAVTDDTSSDLPHSTHVLSSTVPLGHSETHKSAV    800
```

```
Sna-1eIF4G       801 ETNTRRNTSTKGKKKIKEILQKADAAGTTSDLYMAYKGPEEKKESSNVVH    850
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G       801 ETNTRRNTSTKGKKKIKEILQKADAAGTTSDLYMAYKGPEEKKESSNVVH    850

Sna-1eIF4G       851 DVSNQNLLPAIPQAVEAIVDTEPVKNEPEDWEDAADVSTPKLETADNSVN    900
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G       851 DVSNQNLLPAIPQAVEAIVDTEPVKNEPEDWEDAADVSTPKLETADNSVN    900

Sna-1eIF4G       901 AKRGSSDEVSDNCINTEKKYSRDFLLKFADLCTALPEGFDVSPDIANALI    950
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G       901 AKRGSSDEVSDNCINTEKKYSRDFLLKFADLCTALPEGFDVSPDIANALI    950

Sna-1eIF4G       951 VAYMGASHHEHDSYPTPGKVMDRQASGARLDRRPSNVAGDDRWTKNQGSL   1000
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G       951 VAYMGASHHEHDSYPTPGKVMDRQASGARLDRRPSNVAGDDRWTKNQGSL   1000

Sna-1eIF4G      1001 PAGYGGNVGFRPGQGGNSGVLRNPRMQGPIISRPMQPVGPMGGMGRNTPD   1050
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G      1001 PAGYGGNVGFRPGQGGNSGVLRNPRMQGPIISRPMQPVGPMGGMGRNTPD   1050

Sna-1eIF4G      1051 LERWQRGSNFQQKGLFPSPHTPMQVMHKAERKYQVGTIADEEQAKQRQLK   1100
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G      1051 LERWQRGSNFQQKGLFPSPHTPMQVMHKAERKYQVGTIADEEQAKQRQLK   1100

Sna-1eIF4G      1101 SILNKLTPQNFEKLFEQVKSVNIDNAVTLSGVISQIFDKALMEPTFCEMY   1150
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G      1101 SILNKLTPQNFEKLFEQVKSVNIDNAVTLSGVISQIFDKALMEPTFCEMY   1150

Sna-1eIF4G      1151 ADFCFHLSGALPDFNENGEKITFKRLLLNKCQEEFERGEKEEEEASRVAE   1200
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G      1151 ADFCFHLSGALPDFNENGEKITFKRLLLNKCQEEFERGEKEEEEASRVAE   1200

Sna-1eIF4G      1201 EGQVEQTEEEREEKRLQVRRRMLGNIRLIGELYKKRMLTEKIMHACIQKL   1250
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G      1201 EGQVEQTEEEREEKRLQVRRRMLGNIRLIGELYKKRMLTEKIMHACIQKL   1250

Sna-1eIF4G      1251 LGYNQDPHEENIEALCKLMSTIGVMIDHNKAKFQMDGYFEKMKMLSCKQE   1300
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G      1251 LGYNQDPHEENIEALCKLMSTIGVMIDHNKAKFQMDGYFEKMKMLSCKQE   1300

Sna-1eIF4G      1301 LSSRVRFMLINAIDLRKNKWQERMKVEGPKKIEEVHRDAAQERQTQANRL   1350
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G      1301 LSSRVRFMLINAIDLRKNKWQERMKVEGPKKIEEVHRDAAQERQTQANRL   1350

Sna-1eIF4G      1351 SRGPSMNSSGRRGHMEFSSPRGGGGMLSPPAAQMGSYHGPPQGRGFSNQD   1400
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G      1351 SRGPSMNSSGRRGHMEFSSPRGGGGMLSPPAAQMGSYHGPPQGRGFSNQD   1400

Sna-1eIF4G      1401 IRFDDRPSYEPRMVPMPQRSVCEEPITLGPQGGLGQGMSIRRPAVASNTY   1450
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G      1401 IRFDDRPSYEPRMVPMPQRSVCEEPITLGPQGGLGQGMSIRRPAVASNTY   1450

Sna-1eIF4G      1451 QSDATQAGGGDSRRPAGGLNGFGSHRPASPVTHGRSSFQERGTAYVHREF   1500
                     ||||||||||||||||||||||||||||||||||||.||||||||||||||
Col-0eIF4G      1451 QSDATQAGGGDSRRPAGGLNGFGSHRPASPVTHGRSSPQERGTAYVHREF   1500

Sna-1eIF4G      1501 ASLSRASDLSPEVSSARQVLQGPSATVNSPRENALSEEQLENLSLSAIKE   1550
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G      1501 ASLSRASDLSPEVSSARQVLQGPSATVNSPRENALSEEQLENLSLSAIKE   1550

Sna-1eIF4G      1551 YYSARDENEIGMCMKDMNSPAYHPTMISLWVTDSFERKDKERDLLAKLLV   1600
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Col-0eIF4G      1551 YYSARDENEIGMCMKDMNSPAYHPTMISLWVTDSFERKDKERDLLAKLLV   1600

Sna-1eIF4G      1601 NLVKSADNALNEVQLVKG   1618
                     ||||||||||||||||||
Col-0eIF4G      1601 NLVKSADNALNEVQLVKG   1618
```

Abbreviations

Ω	Ohms
µF	Microfarad
aa	Amino acid
ABRC	Arabidopsis Biological Resource Centre
AFLP	Amplified fragment length polymorphism
AIS	Arabidopsis Information Service
Avr	Avirulence
BChV	*Beet chlorosis virus*
BLAST	Basic Local Alignment Search Tool
BMYV	*Beet mild yellowing virus*
bp	Base pair
BWYV	*Beet western yellows virus*
BYV	*Beet yellows virus*
BYDV	*Barley yellow dwarf virus*
BYSV	*Beet yellow stunt virus*
CAPS	Cleaved Amplified Polymorphic Sequences
CMV	*Cucumber mosaic virus*
Col-0	Columbia
CP	Coat Protein
dNTP	Deoxynuleoside triphosphate
dsDNA	Double stranded DNA
eIF	Eukaryotic translation initiation
EBI	European Bioinformatics Institute
ETI	Effector triggered immunity
ETS	Effector triggered susceptibility
F_1	First filial generation
F_2	Second filial generation
FS cDNA	First strand complementary DNA
HR	Hypersensitive response
IgG	Immunoglobulin G

IRES	Internal ribosomal entry site
Kb	Kilobase pair
kDa	Kilodaltons
KO	Knock-out
LB	Left border
Ler	Landsberg
LLR	Leucine-rich repeat
NASC	Nottingham Arabidopsis Stock Centre
NBS	Nucleotide binding site
NCBI	National Centre for Biotechnology Information
NDR	Non-race specific disease resistance
nm	nanometers
nos	Nopaline synthase
NPC	Nuclear pore channel
OD	Optical density
ORF	Open reading frame
p35S	*Cauliflower mosaic virus* 35S promotor
PABP	Poly-A binding protein
PAMP	Pathogen associated molecular pattern
PCR	Polymerase chain reaction
PLRV	*Potato leafroll virus*
PRR	Pathogen recognition receptor
PTGS	Post-transcriptional gene silencing
PTI	PAMP triggered immunity
R	Resistance gene
RB	Right boarder
RdRp	RNA dependent RNA polymererase
RFLP	Restriction Fragment Length Polymorphism
RIL	Recombinant inbred line
RISC	RNA-induced silencing complex
ROI	Reactive oxygen intermediate
RT	Reverse transcription

RTD	Read-through domain
ssDNA	Single stranded DNA
SE	Standard Error
ssRNA	Single stranded RNA
T4SS	Type 4 secretion system
TAIR	The Arabidopsis Information Resource
TAS-ELISA	Triple antibody sandwich enzyme-linked immunosorbant assay
T-DNA	Transfer DNA
Ti	Tumor inducing
TIR	Toll-interleukin receptor
TMV	*Turnip mosaic virus*
TuYV	*Turnips yellows virus*
UTR	Untranslated region
V	Voltage
v/v	Volume/volume
vir	Virulence
VPg	Viral genome linked protein
w/v	Weight/volume

References

Abu-Arish, A., Frenkiel-Krispin, D., Fricke, T., Tzfira, T., Citovsky, V., Wolf, S.G., Elbaum, M., 2004. Three-dimensional reconstruction of Agrobacterium VirE2 protein with single-stranded DNA. The Journal of Biological Chemistry 279, 25359–63.

AGI, 2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408, 796–815.

Alam, S.L., Atkins, J.F., Gesteland, R.F., 1999. Programmed ribosomal frameshifting: much ado about knotting! Proceedings of the National Academy of Sciences of the United States of America 96, 14177–9.

Albar, L., Bangratz-Reyser, M., Hébrard, E., Ndjiondjop, M.-N., Jones, M., Ghesquière, A., 2006. Mutations in the eIF(iso)4G translation initiation factor confer high resistance of rice to Rice yellow mottle virus. The Plant Journal : for cell and molecular biology 47, 417–26.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C., Ecker, J.R., 2003. Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science (New York, N.Y.) 301, 653–7.

Ashrafi, K., Chang, F.Y., Watts, J.L., Fraser, A.G., Kamath, R.S., Ahringer, J., Ruvkun, G., 2003. Genome-wide RNAi analysis of Caenorhabditis elegans fat regulatory genes. Nature 421, 268–72.

Azpiroz-Leehan, R., Feldmann, K.A., 1997. T-DNA insertion mutagenesis in Arabidopsis: going back and forth. Trends in Genetics 13, 152–156.

Barton, K., Binns, N., Matzke, J., Chilton, M.D., 1983. Regeneration of intact tobacco plants containing full length copies of genetically engineered T-DNA, and transmission of T-DNA to R1 progeny. Cell 32, 209–213.

Baumberger, N., Tsai, C.-H., Lie, M., Havecker, E., Baulcombe, D.C., 2007. The Polerovirus silencing suppressor P0 targets ARGONAUTE proteins for degradation. Current Biology 17, 1609–1614.

Beale, H.P., 1931. Specificity of the precipitin reaction in tobacco moasic disease. Contrib. Boyce Thompson Inst. 3, 463–473.

Beauchemin, C., Boutet, N., 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, 775–82.

Bechtold, N., Jaudeau, B., Jolivet, S., Maba, B., Vezon, D., Voisin, R., Pelletier, G., 2000. The maternal chromosome set is the target of the T-DNA in the in planta transformation of Arabidopsis thaliana. Genetics 155, 1875–87.

Bechtold, N., Pelletier, G., 1998. In Planta Agroacterium-Mediated Transformation of Adult Arabidopsis thaliana Plants by Vacuum Infiltration. Methods in Molecular Biology 82.

Beijerinck, M.W., 1898. Concerning a contagium vivum fluidum as cause of the spot disease of tobacco leaves. Phytopathological Classics.

Belkhadir, Y., Subramaniam, R., Dangl, J.L., 2004. Plant disease resistance protein signaling: NBS-LRR proteins and their partners. Current Opinion in Plant Biology 7, 391–399.

Bell, J.M., 1995. Meal and By-product Utilisation in Animal Nutrition, in: Kimber, D.S., McGregor, D.I. (Eds.), Brassica Oilseeds - Production and Utilisation. CAB International, Saskatoon, pp. 301–337.

Bent, A., 2006. Arabidopsis thaliana floral dip transformation method., in: Methods in Molecular Biology (Clifton, N.J.). pp. 87–103.

Beuve, M., Stevens, M., Hsing-Yeh, L., Wintermantel, W.M., Hauser, S., Lemaire, O., 2008. Biological and Molecular Characterization of an American Sugar Beet-Infecting Beet western yellows virus Isolate. Plant Disease 92, 51–60.

Bevan, M., Lane, M., 1984. Binary Agrobactenum vectors for plant transfomation. Nucleic Acids Research 12.

Bortolamiol, D., Pazhouhandeh, M., Marrocco, K., Genschik, P., Ziegler-Graff, V., 2007. The Polerovirus F box protein P0 targets ARGONAUTE1 to suppress RNA silencing. Current Biology 17, 1615–1621.

Bouchez, D., Hofte, H., 1998. Functional Genomics in Plants. Plant Physiology 118, 725–732.

Brault, V., Herrbach, E., Rodriguez-Medina, C., 2011. Luteoviruses. Encyclopedia of life sciences 1–11.

Browning, K.S., 2004. Plant translation initiation factors: it is not easy to be green. Biochemical Society Transactions 32, 589–91.

Bruun-Rasmussen, M., Møller, I.S., Tulinius, G., Hansen, J.K.R., Lund, O.S., Johansen, I.E., 2007. The same allele of translation initiation factor 4E mediates resistance against two Potyvirus spp. in Pisum sativum. Molecular plant-microbe interactions : MPMI 20, 1075–82.

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

Chilton, M.-D., Drummond, M.H., Merlo, D.J., Sciaky, D., 1978. Highly conserved DNA of Ti plasmids overlaps T-DNA maintained in plant tumours. Nature 275, 147–149.

Chory, J., Ecker, J.R., Briggs, S., Caboche, M., Coruzzi, G.M., Cook, D., Dangl, J., Grant, S., Guerinot, M. Lou, Henikoff, S., Martienssen, R., Okada, K., Raikhel, N. V, Somerville, C.R., Weigel, D., 2000. National Science Foundation-Sponsored Workshop Report : " The 2010 Project " Functional Genomics and the Virtual Plant . A Blueprint for Understanding How Plants Are Built and How to Improve Them. Plant Physiology 123, 423–425.

Christie, P.J., 2004. Type IV secretion: the Agrobacterium VirB/D4 and related conjugation systems. Biochimica et biophysica acta 1694, 219–34.

Christou, P., 1996. Transformation Technology. Trends in plant science reviews 1.

Citovsky, V., Wong, M.L., Zambryski, P., 1989. Cooperative interaction of Agrobacterium VirE2 protein with single-stranded DNA: implications for the T-DNA transfer process. Proceedings of the National Academy of Sciences of the United States of America 86, 1193–7.

Clough, S.J., Bent, A.F., 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. The Plant journal : for cell and molecular biology 16, 735–43.

Cooke, D.A., Scott, R.K., 1993. The Sugar Beet Crop, 1st ed. Chapman and Hall, Cambridge.

Coppinger, P., Repetti, P.P., Day, B., Dahlbeck, D., Mehlert, A., Staskawicz, B.J., 2004. Overexpression of the plasma membrane-localized NDR1 protein results in enhanced bacterial disease resistance in Arabidopsis thaliana. The Plant journal : for cell and molecular biology 40, 225–37.

Cornish, P. V, Stammler, S.N., Giedroc, D.P., 2006. The global structures of a wild-type and poorly functional plant luteoviral mRNA pseudoknot are essentially identical. RNA 12, 1959–1969.

Csorba, T., Lózsa, R., Hutvágner, G., Burgyán, J., 2010. Polerovirus protein P0 prevents the assembly of small RNA-containing RISC complexes and leads to degradation of ARGONAUTE1. The Plant J 62, 463–72.

Cutler, S.R., Ehrhardt, D.W., Griffitts, J.S., Somerville, C.R., 2000. Random GFP::cDNA fusions enable visualization of subcellular structures in cells of Arabidopsis at a high frequency. Proceedings of the National Academy of Sciences of the United States of America 97, 3718–23.

D'Arcy, C J, Torrance, L, Martin, R.., 1989. Discrimination Among Luteoviruses and Their Strains by Monoclonal Antibodies and Identification of Common Epitopes.pdf. The American Phytopathological Society 79, 869–873.

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

Desfeux, C., Clough, S.J., Bent, a F., 2000. Female reproductive tissues are the primary target of Agrobacterium-mediated transformation by the Arabidopsis floral-dip method. Plant physiology 123, 895–904.

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

Dickinson, M., 2005. Molecular Plant Pathology, 2nd ed. BIOS Scientific Publishers, London.

Domier, L.L., McCoppin, N.K., Larsen, R.C., D'Arcy, C.J., 2002. Nucleotide sequence shows that Bean leafroll virus has a Luteovirus-like genome organization. The Journal of general virology 83, 1791–1798.

Dreher, T., Miller, W., 2006. Translational control in positive strand RNA plant viruses. Virology 344, 185–197.

Duffus, J.E., Russell, G.., 1970. Serological and host range evidence for the occurence of beet western yellows virus in Europe. Phytopathology 60, 1199–1202.

Duprat, A., Caranta, C., Revers, F., Menand, B., Browning, K.S., 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, 927–34.

Ekbom, B., 1995. Insect Pests, in: Kimber, D.., McGregor, D.I. (Eds.), Brassica Oilseeds - Production and Utilisation. CAB International, Uppsala, pp. 141–152.

Elliot, M.C., Weston, G.., 1993. Biology and Plysiology of the Sugar beet plant, in: Cooke, D.., Scott, R.. (Eds.), The Sugar Beet Crop - Science into Practice. Chapman and Hall, pp. 37–66.

Eskelin, K., Hafrén, A., Rantalainen, K.I., Mäkinen, K., 2011. Potyviral VPg enhances viral RNA Translation and inhibits reporter mRNA translation in planta. Journal of virology 85, 9210–21.

Fauquat, C.M., Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A., 2005. Virus Taxonomy - Eighth report of the international committee on taxonomy of viruses, 8th ed. Elevier.

Filichkin, S.A., Gelvin, S.B., 1993. Formation of a putative relaxation intermediate during T-DNA processing directed by the Agrobacterium tumefaciens VirD1, D2 endonuclease. Molecular microbiology 8, 915–926.

Flor, H.H., 1971. Current status of the gene-for-gene concept. Annual review of phytopathologyof phytopathology 9, 275–296.

Fraser, R.S.., 1990. The Genetics Of Resistance To Plant Viruses. Annual Review of Phytopathology 28, 179–200.

Gallie, D.R., 1991. The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. Genes & development 5, 2108–16.

Gallie, D.R., 2001. Cap-Independent Translation Conferred by the 5' Leader of Tobacco Etch Virus Is Eukaryotic Initiation Factor 4G Dependent. Journal of Virology 75, 12141–12152.

Gallie, D.R., Browning, K.S., 2001. eIF4G functionally differs from eIFiso4G in promoting internal initiation, cap-independent translation, and translation of structured mRNAs. The Journal of biological chemistry 276, 36951–60.

Gao, Z., Johansen, E., Eyers, S., Thomas, C.L., Noel Ellis, T.H., 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, 376–85.

Garret, A., Kerlan, C., Thomas, D., 1996. Ultrastructural study of acquisition and retention of potato leafroll luteovirus in the alimentary canal of its aphid vector, Myzus persicae Sulz. Archives of Virology 141, 1279–1292.

Gladders, P., Ginsburg, D., Ritchie, F., Smith, J.A., Waterhouse, S., Tucker, C., Tonguc, L., 2006. The encyclopaedia of oilseed rape diseases. ADAS BASF.

Glazebrook, J., 2005. Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. Annual review of phytopathology 43, 205–27.

Godwin, I., Todd, G., Ford-Lloyd, B., Newbury, H.J., 1991. The effects of acetosyringone and pH on Agrobacterium-mediated transformation vary according to plant species. Plant Cell Reports 9, 671–675.

Gray, S., Gildow, F.E., 2003. Luteovirus-aphid interactions. Annual review of phytopathology 41, 539–66.

Gray, S.., Power, A.., Smith, D.., Seaman, A.., Altman, N.., 1991. Aphid transmission of barley yellow dwarf virus- inoculation access periods and epidemiologic implications. Phytopathology 81, 539–545.

Gray, S.., Smith, D., Sorrells, M., 1994. Reduction of disease incidence in small field plots by isolate specific resistance to barley yellow dwarf virus. Phytopathology 84, 713–718.

Guilley, H., Richards, K.E., Jonard, G., 1995. Nucleotide sequence of beet mild yellowing virus RNA. Archives of Virology 140, 1109–1118.

H.G. Smith, I. Barker, G. Brewer, M.S. and P.B.H., 1996. Production and evaluation of monoclonal antibodies for the detection of beet mild yellowing luteovirus and related strains. European Journal of Plant Pathology 102, 163–169.

Hagiwara, Y., Komoda, K., Yamanaka, T., Tamai, A., Meshi, T., Funada, R., Tsuchiya, T., Naito, S., Ishikawa, M., 2003. Subcellular localization of host and viral proteins associated with tobamovirus RNA replication. The EMBO journal 22, 344–53.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence editor and analysis program for windows 95/98/NT. Nucleic Acids Symposium Series 41, 95–98.

Hartland, J.., 1993. By-Products, in: Cooke, D.A., Scott, R.K. (Eds.), The Sugar Beet Crop - Science into Practice. Chapman and Hall, Cambridge, pp. 619–647.

Haseloff, J., Amos, B., 1995. GFP in plants. Trends in Genetics 11, 328–329.

Hauser, S., Weber, C., Vetter, G., Stevens, M., Beuve, M., Lemaire, O., 2000. Improved detection and differentiation of poleroviruses infecting beet or rape by multiplex RT-PCR. Journal of virological methods 89, 11–21.

Hébrard, E., Bessin, Y., Michon, T., Longhi, S., Uversky, V.N., Delalande, F., Van Dorsselaer, A., Romero, P., Walter, J., Declerck, N., Fargette, D., 2009. Intrinsic disorder in Viral Proteins Genome-Linked: experimental and predictive analyses. Virology journal 6, 23.

Herrera-Estrella, L., Depicker, A., Van Montagu, M., Schell, J., 1983. Expression of chimaeric genes transferred into plant cells using a Ti-plasmi-derived vector. Nature 303, 209–213.

Hoekema, A., Roelvink, P.W., Hooykaas, P.J.J., Schilperoort, R.A., 1984. Delivery of T-DNA from the Agrobacterium tumefaciens chromosome into plant cells. The EMBO journal 3, 2485–90.

Hull, R., 2009. Comparative Plant Virology. Roger Hull.

Jaag, H.M., Kawchuk, L., Rohde, W., Fischer, R., Emans, N., Prüfer, D., 2003. An unusual internal ribosomal entry site of inverted symmetry directs expression of a potato leafroll polerovirus replication-associated protein. Proceedings of the National Academy of Sciences of the United States of America 100, 8939–44.

Jaggard, K.W., Dewar, A.M., Pidgeon, J.D., 1998. The relative effects of drought stress and virus yellows on the yield of sugarbeet in the UK, 1980-95. The Journal of Agricultural Science 130, 337–343.

Jay, C.N., Rossall, S., Smith, H.G., 1999. Effects of beet western yellows virus on growth and yield of oilseed rape (Brassica napus). Journal of Agricultural Science 133, 131–139.

Jia, Y., McAdams, S. a, Bryan, G.T., Hershey, H.P., Valent, B., 2000. Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. The EMBO journal 19, 4004–4014.

Jones, J.D.G., Dangl, J.L., 2006. The plant immune system. Nature 444, 323–9.

Jones, R.A.C., Coutts, B.A., Hawkes, J., 2007. Yield-limiting potential of Beet western yellows virus in Brassica napus. Australian Journal of Agricultural Research 58, 788–801.

Juergens, M., Paetsch, C., Kramer, I., Zahn, M., Rabenstein, F., Schondelmaier, J., Schliephake, E., Snowdon, R., Friedt, W., Ordon, F., 2010. Genetic analyses of the host-pathogen system Turnip yellows virus (TuYV)—rapeseed (Brassica napus L.) and development of molecular markers for TuYV-resistance. Theoretical and Applied Genetics 120, 735–744.

Juszczuk, M., Paczkowska, E., Sadowy, E., Zagorski, W., Hulanicka, D., 2000. Effect of genomic and subgenomic leader sequences of potato leafroll virus on gene expression. FEBS Letters 484, 33–36.

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

Kanyuka, K., Druka, A., Caldwell, D.., Tymon, A., McCallum, N., Waugh, R., Adams, M.., 2005. Evidence that the recessive bymovirus resistance

locus rym 4 in barley corresponds to the eukaryotic translation initiation factor 4E gene. Molecular Plant Pathology 6, 449–458.

Kawaguchi, R, Bailey-Serres, J., 2002. Regulation of translational initiation in plants. Current Opinion in Plant Biology 5, 460–465.

Kimber, D.S., McGregor, D.I., 1995. Brassica Oilseeds - Production and Utilisation, 1st ed. CAB International, Wallingford.

Kneller, E., Rakotondrafara, A., Miller, W., 2006. Cap-independent translation of plant viral RNAs. Virus Research 119, 63–75.

Kole, C., 2007. Genome Mapping and Molecular Breeding in Plants - Oilseeds, 1st ed. Springer, Berlin.

Koornneef, M., Alonso-Blanco, C., Vreugdenhil, D., 2004. Naturally occurring genetic variation in Arabidopsis thaliana. Annual review of plant biology 55, 141–72.

Koornneef, M., Meinke, D., 2010. The development of Arabidopsis as a model plant. The Plant journal : for cell and molecular biology 61, 909–21.

Koornneef, M., Van der Veen, J., 1983. Trisomics in Arabidopsis thaliana and the location of linkage groups. Genetica 61, 41–46.

Korbitz, W., 1995. Utilisation of Oil as a Biodiesel Fuel, in: Kimber, D.S., McGregor, D.I. (Eds.), Brassica Oilseeds - Production and Utilisation. CAB International, Vienna, pp. 353–373.

Kornberg, A., 2005. DNA Replication, 2nd ed. University Science Books, USA.

Labra, M., Vannini, C., Grassi, F., Bracale, M., Balsemin, M., Basso, B., Sala, F., 2004. Genomic stability in Arabidopsis thaliana transgenic plants obtained by floral dip. Theoretical and applied genetics 109, 1512–8.

Lacroix, B., Citovsky, V., 2013. Agrobacterium, in: Brenner's Encyclopedia of Genetics, 2nd Edition, Volume 1. pp. 23–25.

Larkin, P.J., Scowcroft, W.R., 1981. Somaclonal variation — a novel source of variability from cell cultures for plant improvement. Theoretical and Applied Genetics 60, 197–214.

Le Gall, O., Aranda, A., Caranta, C., 2011. Plant resistance to viruses mediated by translation intiation factors, in: Caranta, C., Aranda, M., Tepfer, M., Lopez-Moya, J. (Eds.), Recent Advances in Plant Virology. Caister Academic Press, Caister, pp. 177–194.

Lee, L.-Y., Gelvin, S.B., 2008. T-DNA binary vectors and systems. Plant physiology 146, 325–32.

Leiser, R.M., Ziegler-Graff, V., Reutenauer, A., Herrbach, E., Lemaire, O., Guilley, H., Richards, K., Jonard, G., 1992. Agroinfection as an alternative to insects for infecting plants with beet western yellows luteovirus. Proceedings of the National Academy of Sciences of the United States of America 89, 9136–40.

Lellis, A.D., Kasschau, K.D., Whitham, S. a, 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, 1046–51.

Lemaire, O., Herrbach, E., Stevens, M., Bouchery, Y., Smith, H.G., 1995. Detection of sugar beet-infecting beet mild yellowing luteovirus isolates with a specific RNA probe. The American Phytopathological Society 85, 1513–1518.

Léonard, S., Plante, D., Wittmann, S., Daigneault, N., Fortin, M., 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, X., Ryan, M.D., Lamb, J.W., 2000. Potato leafroll virus protein P1 contains a serine proteinase domain. The Journal of general virology 81, 1857–1864.

Lloyd, a M., Barnason, a R., Rogers, S.G., Byrne, M.C., Fraley, R.T., Horsch, R.B., 1986. Transformation of Arabidopsis thaliana with Agrobacterium tumefaciens. Science (New York, N.Y.) 234, 464–6.

Marggraf, A.S., 1749. Experiences chymiques, faites dans le dessein de tirer un veritable sucre de diverses plantes qui crossent dans nos contrees. Histoire de l'Academie Royale des Sciences et BElles Lettres 79–90.

Martienssen, R. a, 1998. Functional genomics: probing plant gene function and expression with transposons. Proceedings of the National Academy of Sciences of the United States of America 95, 2021–6.

Martin, G.B., Bogdanove, A.J., Sessa, G., 2003. Understanding the functions of plant disease resistance proteins. Annual review of plant biology 54, 23–61.

Martin, R.R., Keese, P.K., Young, M.J., Waterhouse, P.M., Gerlach, W.L., 1990. Evolution and molecular biology of Luteoviruses. Annual Review of Phytopathology 28, 341–363.

Matthews, R.E.., 1991. Plant Virology, 3rd ed. Academic Press inc., London.

Maule, A.J., Caranta, C., Boulton, M., 2007. Sources of natural resistance to plant viruses : status and prospects. Molecular Plant Pathology 8, 223–231.

Mayo, M.A., 2002. ICTV at the Paris ICV: Results of the Plenary Session and the Binomial Ballot. Archives of virology 147, 2254–2260.

McDonald, B.E., 1995. Oil Properties of Importance in Human Nutrition, in: Kimber, D., McGregor, D. (Eds.), Brassica Oilseeds - Production and Utilisation. CAB International, Winnipeg, p. 291299.

McElver, J., Tzafrir, I., Aux, G., Rogers, R., Ashby, C., Smith, K., Thomas, C., Schetter, a, Zhou, Q., Cushman, M. a, Tossberg, J., Nickle, T., Levin, J.Z., Law, M., Meinke, D., Patton, D., 2001. Insertional mutagenesis of genes required for seed development in Arabidopsis thaliana. Genetics 159, 1751–63.

Meinke, D., Sweeney, C., Muralla, R., 2009. Integrating the genetic and physical maps of Arabidopsis thaliana: identification of mapped alleles of cloned essential (EMB) genes. PloS one 4, e7386.

Meyerowitz, E.M., 2001. Prehistory and history of Arabidopsis research. Plant physiology 125, 15–9.

Michelmore, R.W., 2003. The impact zone: genomics and breeding for durable disease resistance. Current Opinion in Plant Biology 6, 397–404.

Michelmore, R.W., Paran, I., Kesseli, R. V, 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proceedings of the National Academy of Sciences of the United States of America 88, 9828–32.

Mittelbach, M., 1996. Diesel fuel derived from vegetable oils, VI; Specifications and quality control of Biodiesel. Bioresource Technology 56, 7–11.

Morel, J.B., Dangl, J.L., 1997. The hypersensitive response and the induction of cell death in plants. Cell death and differentiation 4, 671–83.

Murakami, R., Nakashima, N., Hinomoto, N., Kawano, S., Toyosato, T., 2011. The genome sequence of pepper vein yellows virus (family Luteoviridae, genus Polerovirus). Archives of virology 156, 921–3.

Murphy, J.F., Rychlik, W., Rhoads, R.E., Hunt, a G., 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, 511–3.

Neeleman, L., Olsthoorn, R.C., Linthorst, H.J., Bol, J.F., 2001. Translation of a nonpolyadenylated viral RNA is enhanced by binding of viral coat protein or polyadenylation of the RNA. Proceedings of the National Academy of Sciences of the United States of America 98, 14286–91.

Nicaise, C., German-retana, S., Sanjua, R., Dubrana, M.-P., Mazier, M., Maisonneuve, B., Candresse, T., Caranta, C., LeGall, O., 2003. The Eukaryotic Translation Initiation Factor 4E Controls Lettuce Susceptibility to the Potyvirus Lettuce mosaic virus 1. Plant Physiology 132, 1272–1282.

Nicaise, V., Gallois, J.-L., Chafiai, F., Allen, L.M., Schurdi-Levraud, V., Browning, K.S., Candresse, T., Caranta, C., Le Gall, O., German-Retana, S., 2007. Coordinated and selective recruitment of eIF4E and eIF4G factors for potyvirus infection in Arabidopsis thaliana. FEBS letters 581, 1041–6.

Nieto, C., Morales, M., Orjeda, G., Clepet, C., Monfort, A., Sturbois, B., Puigdomènech, P., Pitrat, M., Caboche, M., Dogimont, C., Garcia-Mas, J., Aranda, M. a, Bendahmane, A., 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, 452–62.

Nieto, C., Rodríguez-Moreno, L., Rodríguez-Hernández, A.M., Aranda, M. a, Truniger, V., 2011. Nicotiana benthamiana resistance to non-adapted Melon necrotic spot virus results from an incompatible interaction between virus RNA and translation initiation factor 4E. The Plant journal : for cell and molecular biology 66, 492–501.

Nixon, P.L., Cornish, P. V, Suram, S. V, Giedroc, D.P., 2002. Thermodynamic analysis of conserved loop-stem interactions in P1-P2 frameshifting RNA pseudoknots from plant Luteoviridae. Biochemistry 41, 10665–10674.

Orson, J., 1995. Weeds and Their Control, in: Kimber, D.., McGregor, D.I. (Eds.), Brassica Oilseeds - Production and Utilisation. CAB International, Cambridge, pp. 93–109.

Pazhouhandeh, M., Dieterle, M., Marrocco, K., Lechner, E., Berry, B., Brault, V., Hemmer, O., Kretsch, T., Richards, K.E., Genschik, P., Ziegler-Graff, V., 2006. F-box-like domain in the polerovirus protein P0 is required for silencing suppressor function. Proceedings of the National Academy of Sciences of the United States of America 103, 1994–1999.

Percival-Alwyn, L., 2010. A Molecular Genetic Analysis of Beet Polerovirus Infection. Advances in cancer research. University of East Anglia.

Peter, K.A., Gildow, F., Palukaitis, P., Gray, S.M., 2009. The C terminus of the polerovirus p5 readthrough domain limits virus infection to the phloem. Journal of Virology 83, 5419–29.

Pfeffer, S., Dunoyer, P., Heim, F., Richards, K.E., Jonard, G., Ziegler-Graff, V., 2002. P0 of Beet Western Yellows Virus Is a Suppressor of Posttranscriptional Gene Silencing. Journal of Virology 76, 6815–6824.

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

Rédei, G.P., 1975. Arabidopsis as a genetic tool. Annual Review of Genetics 9, 111–127.

Réidei, G., 1992. A heuristic glance at the past of Arabidopsis gentics, in: Koncz, C., Chua, N., Schell, J. (Eds.), Methods in Arabidopsis Research. World Scientific Publishing Co., Singapore, pp. 1–15.

Reinbold, C., Lacombe, S., Ziegler-Graff, V., Scheidecker, D., Wiss, L., Beuve, M., Caranta, C., Brault, V., 2013. Closely related poleroviruses depend on distinct translation initiation factors to infect Arabidopsis thaliana. Molecular plant-microbe interactions : MPMI 26, 257–65.

Reutenauer, A., Ziegler-Graff, V., Lot, H., Scheidecker, D., Guilley, H., Richards, K., Jonard, G., 1993. Identification of beet western yellows luteovirus genes implicated in viral replication and particle morphogenesis. Virology 195, 692–699.

Rimmer, S.R., Buchwaldt, L., 1995. Diseases, in: McGregor, D., Kimber, D.. (Eds.), Brassica Oilseeds - Production and Utilisation. CAB International, Manitoba, pp. 111–140.

Ritzenthaler, C., 2005. Resistance to plant viruses: old issue, news answers? Current opinion in biotechnology 16, 118–22.

Robaglia, C., Caranta, C., 2006. Translation initiation factors: a weak link in plant RNA virus infection. Trends in plant science 11, 40–5.

Rodriguez, C.M., Freire, M. a, Camilleri, C., 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 : for cell and molecular biology 13, 465–73.

Roossinck, M.J., 1997. Mechanisms of Plant Virus Evolution. Annu. Rev. of Phytopathol. 35, 191–209.

Roudet-Tavert, G., Michon, T., Walter, J., Delaunay, T., Redondo, E., Le Gall, O., 2007. Central domain of a potyvirus VPg is involved in the interaction with the host translation initiation factor eIF4E and the viral protein HcPro. The Journal of general virology 88, 1029–33.

Rousseau, D., Kaspar, R., Rosenwald, I., Gehrke, L., Sonenberg, N., 1996. Translation initiation of ornithine decarboxylase and nucleocytoplasmic transport of cyclin D1 mRNA are increased in cells overexpressing eukaryotic initiation factor 4E. Proceedings of the National Academy of Sciences of the United States of America 93, 1065–70.

Ruffel, S., Caranta, C., Palloix, A., Lefebvre, V., Caboche, M., Bendahmane, A., 2004. Structural analysis of the eukaryotic initiation factor 4E gene controlling potyvirus resistance in pepper: exploitation of a BAC library. Gene 338, 209–16.

Ruffel, S., Dussault, M.-H., Palloix, A., Moury, B., Bendahmane, A., Robaglia, C., Caranta, C., 2002. A natural recessive resistance gene against potato virus Y in pepper corresponds to the eukaryotic initiation factor 4E (eIF4E). The Plant journal : for cell and molecular biology 32, 1067–75.

Ruffel, S., Gallois, J.-L., Moury, B., Robaglia, C., Palloix, A., Caranta, C., 2006. Simultaneous mutations in translation initiation factors eIF4E and eIF(iso)4E are required to prevent pepper veinal mottle virus infection of pepper. The Journal of general virology 87, 2089–98.

Ruffel, S., Gallois, J.L., Lesage, M.L., 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, 346–53.

Russel, G.., 1960. Sugar beet yellows: further study of the distribution and interrelationship of viruses and virus strains found in East Anglia. Annals of Applied Biology 46, 393–398.

Russell, G.E., 1965. The host range of some English isolates of beet yellowing viruses. Annals of Applied Biology 55, 245–252.

Sadowy, E., Maasen, A., Juszczuk, M., David, C., Zagoroski-Ostoja, W., Gronenborn, B., Hulanicka, M.D., 2001. The ORF0 product of Potato leafroll virus is indispensable for. Journal of General Virology 82, 1529–1532.

Sato, M., Nakahara, K., Yoshii, M., Ishikawa, M., 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, 1167–71.

Schaad, M.C., Anderberg, R.J., 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, 300–6.

Schliephake, E., Graichen, K., Rabenstein, F., 2000. Investigations on the vector transmission of the Beet mild yellowing virus (BMYV) and the Turnip yellows virus (TuYV). Zeitschrift fur Pflanzenjrajenheiten und Pflanzenschultz 107, 81–87.

Sessions, A., Burke, E., Presting, G., Aux, G., Mcelver, J., Patton, D., Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J.D., Cotton, D., Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmerly, B., Mitzel, T., Katagiri, F.,

Glazebrook, J., Law, M., Goff, S.A., 2002. A High-Throughput Arabidopsis Reverse Genetics System. The Plant Cell 14, 2985–2994.

SHARON A. HALL, REBECCA L. ALLEN, RACHEL E. BAUMBER, LAURA A. BAXTER, KATE FISHER, PETER D. BITTNER-EDDY, LAURA E. ROSE, ERIC B. HOLUB, BEYNON, A.J.L., 2009. Maintenance of genetic variation in plants and pathogens involves complex networks of gene-for-gene interactions. Molecular Plant Pathology 10, 449–457.

Skaf, J.S., Schultz, M.H., Hirata, H., de Zoeten, G. a, 2000. Mutational evidence that the VPg is involved in the replication and not the movement of Pea enation mosaic virus-1. The Journal of general virology 81, 1103–1109.

Smith, E.F., 1894. Ueber die Mosaikkrankheit des Tabaks by Adolf Mayer Review by: Erwin F. Smith. The Journal of Mycology 7, 382–385.

Smith, H.G., Hinckes, J. a., 1985. Studies on beet western yellows virus in oilseed rape (Brassica napus ssp. oleifera) and sugar beet (Beta vulgaris). Annals of Applied Biology 107, 473–484.

Somerville, C., Koornneef, M., 2002. A fortunate choice: the history of Arabidopsis as a model plant. Nature Reviews Genetics 2, 883–889.

Sonntag, N.O.., 1995. Industrial Utilisation of Long-Chain Fatty Acids and their Derivatives, in: Kimber, D.S., McGregor, D.I. (Eds.), Brassica Oilseeds - Production and Utilisation. CAB International, Ovilla, pp. 339–352.

Stephan, D., Maiss, E., 2006. Biological properties of Beet mild yellowing virus derived from a full-length cDNA clone. The Journal of general virology 87, 445–9.

Stevens, M., Freeman, B., Hsing-Yeh, L., Herrbach, E., Lemaire, O., 2005. Beet poleroviruses : close friends or distant relatives? Molecular Plant Pathology 6, 1–9.

Stevens, M., Hallsworth, P.B., Smith, H.G., 2004. The effects of Beet mild yellowing virus and Beet chlorosis virus on the yield of UK field-grown sugar beet in 1997 , 1999 and 2000. Annals of Applied Biology 144, 113–119.

Stevens, M., Hsing-Yeh, L., Lemaire, O., 2006. Virus Diseases, in: Dreycott, A.. (Ed.), Sugar Beet. Blackwell Publishing, Oxford, pp. 256–278.

Stevens, M., Mcgrann, G., Clark, B., 2008. Turnip yellows virus (syn Beet western yellows virus): an emerging threat to European oilseed rape production? HGCA 0–36.

Stevens, M., Smith, H.G., Hallsworth, P.B., 1995. Detection of the luteoviruses, beet mild yellowing virus and beet western yellows virus, in

aphids caught in sugar-beet and oilseed rape crops, 1990–1993. Annals of Applied Biology 127, 309–320.

Swarbreck, D., Wilks, C., Lamesch, P., Berardini, T.Z., Garcia-Hernandez, M., Foerster, H., Li, D., Meyer, T., Muller, R., Ploetz, L., Radenbaugh, A., Singh, S., Swing, V., Tissier, C., Zhang, P., Huala, E., 2008. The Arabidopsis Information Resource (TAIR): gene structure and function annotation. Nucleic acids research 36, D1009–14.

Talbot, N.J., 2004. Plant-Pathogen Interactions, 1st ed. Blackwell Publishing, Oxford.

Thivierge, K., Nicaise, V., Dufresne, P., Cotton, S., Laliberté, J.F., Le Gall, O., Fortin, M., 2005. Plant Virus RNAs . Coordinated Recruitment of Conserved Host Functions by (+) ssRNA Viruses during Early Infection Events. Plant physiology 138, 1822–1827.

Thomas, C.M., Vos, P., Zabeau, M., Jones, D. a, Norcott, K. a, Chadwick, B.P., Jones, J.D., 1995. Identification of amplified restriction fragment polymorphism (AFLP) markers tightly linked to the tomato Cf-9 gene for resistance to Cladosporium fulvum. The Plant journal : for cell and molecular biology.

Ting, J.P.-Y., Davis, B.K., 2005. CATERPILLER: a novel gene family important in immunity, cell death, and diseases. Annual review of immunology 23, 387–414.

Topisirovic, I., Culjkovic, B., Cohen, N., Perez, J.M., Skrabanek, L., Borden, K.L.B., 2003. The proline-rich homeodomain protein, PRH, is a tissue-specific inhibitor of eIF4E-dependent cyclin D1 mRNA transport and growth. The EMBO journal 22, 689–703.

Torres, M.W., Corrêa, R.L., Schrago, C.G., 2005. Analysis of differential selective forces acting on the coat protein (P3) of the plant virus family Luteoviridae. Genetics and molecular research GMR 4, 790–802.

Trontin, C., Tisné, S., Bach, L., Loudet, O., 2011. What does Arabidopsis natural variation teach us (and does not teach us) about adaptation in plants? Current opinion in plant biology 14, 225–31.

Tsujimoto, Y., Numaga, T., Ohshima, K., Yano, M.-A., Ohsawa, R., Goto, D.B., Naito, S., Ishikawa, M., 2003. Arabidopsis TOBAMOVIRUS MULTIPLICATION (TOM) 2 locus encodes a transmembrane protein that interacts with TOM1. The EMBO journal 22, 335–43.

Turk, S.C., van Lange, R.P., Regensburg-Tuïnk, T.J., Hooykaas, P.J., 1994. Localization of the VirA domain involved in acetosyringone-mediated vir gene induction in Agrobacterium tumefaciens. Plant molecular biology 25, 899–907.

Tzfira, T., Citovsky, V., 2006. Agrobacterium-mediated genetic transformation of plants: biology and biotechnology. Current opinion in biotechnology 17, 147–54.

Uma, B., Rani, T.S., Podile, A.R., 2011. Warriors at the gate that never sleep: non-host resistance in plants. Journal of plant physiology 168, 2141–52.

Valvekens, D., Montagu, M. V, Van Lijsebettens, M., 1988. Agrobacterium tumefaciens-mediated transformation of Arabidopsis thaliana root explants by using kanamycin selection. Proceedings of the National Academy of Sciences of the United States of America 85, 5536–40.

Van den Heuvel, J.F., Hogenhout, S.A., van der Wilk, F., 1999. Recognition and receptors in virus transmission by arthropods. Trends in microbiology 7, 71–76.

Van der Want, J.P.H., Dijkstra, J., 2006. A history of plant virology. Archives of virology 151, 1467–98.

Van der Wilk, F., Verbeek, M., Dullemans, A.M., Van Den Heuvel, J.F., 1997. The genome-linked protein of potato leafroll virus is located downstream of the putative protease domain of the ORF1 product. Virology 234, 300–303.

Van Slogteren, E., Van Slogteren, D.H., 1957. Serological identification of plant viruses and serological diagnosis of virus diseases of plants. Annual review of microbiology 11, 149–64.

Veidt, I., Bouzoubaa, S.E., Leiser, R.M., Ziegler-Graff, V., Guilley, H., Richards, K., Jonard, G., 1992. Synthesis of full-length transcripts of beet western yellows virus RNA: messenger properties and biological activity in protoplasts. Virology 186, 192–200.

Watson, B., Currier, T.C., Gordon, M.P., Nester, E.W., 1975. Plasmid Required for Virulence of Agrobacterium tumefaciens. Journal of Bacteriology 123, 255–264.

Watson, M. a., 1940. Studies on the Transmission of Sugar-Beet Yellows Virus by the Aphis, Myzus persicae (Sulz.). Proceedings of the Royal Society B: Biological Sciences 128, 535–552.

Weigel, D., Glazebrook, J., 2002. How to study gene expression, in: Arabidopsis. Cold Spring Harbor Laboratory Press, Cold Spring Harbour.

Wells, S.E., Hillner, P.E., Vale, R.D., Sachs, a B., 1998. Circularization of mRNA by eukaryotic translation initiation factors. Molecular cell 2, 135–40.

Wesley, S. V, Helliwell, C. a, Smith, N. a, Wang, M.B., Rouse, D.T., Liu, Q., Gooding, P.S., Singh, S.P., Abbott, D., Stoutjesdijk, P. a, Robinson, S.P.,

Gleave, a P., Green, a G., Waterhouse, P.M., 2001. Construct design for efficient , effective and high-throughput gene silencing in plants. The Plant journal : for cell and molecular biology 27, 581–90.

Wimmer, E., 1982. Genome-linked proteins of viruses. Cell 28, 199–201.

Winner, C., 1993. History of the Crop, in: Cooke, D.A., Scott, R.K. (Eds.), The Sugar Beet Crop - Science into Practice. Chapman and Hall, pp. 1–32.

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

Xia, Z.-L., Wang, M.-P., Liu, Q.-J., Wang, D.-W., 2007. [Expression and subcellular localization of the ORF4 gene of barley yellow dwarf virus GAV strain in baculovirus-insect cell expression system]. Sheng wu gong cheng xue bao = Chinese journal of biotechnology 23, 1107–11.

Yamanaka, T., Ohta, T., Takahashi, M., Meshi, T., Schmidt, R., Dean, C., Naito, S., Ishikawa, M., 2000. TOM1, an Arabidopsis gene required for efficient multiplication of a tobamovirus, encodes a putative transmembrane protein. Proceedings of the National Academy of Sciences of the United States of America 97, 10107–12.

Ye, G.N., Stone, D., Pang, S.Z., Creely, W., Gonzalez, K., Hinchee, M., 1999. Arabidopsis ovule is the target for Agrobacterium in planta vacuum infiltration transformation. The Plant journal : for cell and molecular biology 19, 249–57.

Yoshii, M., Nishikiori, M., Tomita, K., Yoshioka, N., Kozuka, R., Naito, S., Ishikawa, M., 2004. The Arabidopsis cucumovirus multiplication 1 and 2 loci encode translation initiation factors 4E and 4G. Journal of Virology 78, 6102–6111.

Ziegler-Graff, V., Brault, V., Mutterer, J.D., Simonis, M., Herrbach, E., Guilley, H., Richards, K.E., Jonard, G., 1996. The coat protein of Beet western yellows luteovirus is essential for systemic infection but the viral gene products p29 and p19 are dispensable for systemic infection and aphid transmission. The American Phytopathological Society 9, 501–510.

Zipfel, C., Felix, G., 2005. Plants and animals: a different taste for microbes? Current opinion in plant biology 8, 353–60.