TRANSGENERATIONAL CHANGES IN PROGENY OF COMPATIBLE PATHOGEN INFECTED PLANTS

PALAK KATHIRIA
MSc (Plant Biotechnology), University of Lethbridge, 2006
BSc (Agricultural sciences), Gujarat Agricultural University, 2003

A Thesis
Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfilment of the
Requirements for the Degree

PHD: MOLECULAR AND CELLULAR BIOLOGY

Department of Biological Sciences
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

© Palak Kathiria, 2010

Dedicated to

My parents Keshubhai and Jayaben Kathiria

for unconditional love and support

Acknowledgement:

I am thankful wholeheartedly to my Guru and mentor Dr. Igor Kovalchuk, for supporting my quest during the program. I learned many things from him in and outside the field of science.

Special thanks to my committee members Dr. Theresa Burg and Dr. Gerlinde Metz for their valuable support and advice. Also, I am thankful to Dr. Denise Gaudet and Dr. Robert Sutherland for the help during thesis defence.

I am also thankful to all the friends in Kovalchuk labs, who have helped me during the PhD program and made the journey a very pleasant one.

Table of Contents

1.0 Introduction 1
1.1 References: 4
2.0 Literature Review 5
2.1 Introduction 5
2.2 Different pathogen resistance mechanisms in plants 6
2.2.1 Non-host resistance 7
2.2.2 The mechanism of active resistance in plants 8
2.2.2.1 Incompatible interactions 8
2.2.2.2 Systemic acquired resistance (SAR) 11
2.2.2.3 The role of *PR* genes 12
2.2.2.4 The role of callose induction following pathogen infection 14
2.2.2.5 The role of polyphenols and secondary metabolites in pathogen resistance 15
2.2.2.6 Conserved regions in the disease resistance genes 17
2.2.2.7 The function of disease resistance genes 18
2.2.2.8 The Evolution of disease resistance genes 21
2.3 Homologous recombination as a double-strand break repair mechanism 23
2.3.1 The balance between genome stability and evolution 24
2.3.2 The role of homologous recombination in plant-pathogen interactions 25
2.3.3 Changes in progeny of pathogen-infected plants 26
2.3.4 Different designs of recombination constructs to study the rate of recombination 27
2.4 *TMV* and its interaction with tobacco 29
2.5 Reporter genes and their uses 30
2.6 References: 32
3.0 Tobacco mosaic virus infection results in an increase in recombination frequency and resistance to viral, bacterial and fungal pathogens in the progeny of infected tobacco plants 46
3.1 Abstract 46
3.2 Introduction 46
3.3 Materials and Methods 49
3.3.1 Generation and use of transgenic tobacco plants carrying a homologous recombination substrate 49
3.3.2 Infection of SR1 plants with *TMV* and generation of infected plant's progeny 49
3.3.3 Analysis of plant resistance to *TMV* infection. Plant inoculation, monitoring infection symptoms and analysis of the viral titer 50
3.3.4 Black Shank Disease Assays 51
3.3.5 Analysis of plant resistance to *Pseudomonas syringae* 52
3.3.6 Generating the second progeny of infected plants 53
3.3.7 Visualization of luciferase activity 53
3.3.8 Exposure to methyl methane sulfonate (MMS) 53
3.3.9 Exposure to NaCl and rose Bengal 54
3.3.10 Analysis of the phenolic content 54
3.3.11 Analysis of callose deposition 54
3.3.12 Analysis of global genome methylation using a cytosine extension assay and using the anti-5-methyl cytosine antibody 55
3.3.13 Analysis of PR1 expression in tobacco infected with *TMV* or *Pseudomonas syringae* 56

3.4 Results 57
3.4.1 Generating the progeny of virus-infected plants 57
3.4.2 PI_1 plants showed higher spontaneous and induced somatic HRF as compared to PC_1 plants 57
3.4.3 The progeny of infected plants exhibit delay in the onset of the symptoms of viral infection 58
3.4.4 The progeny of infected plants have lower bacterial titer upon infection with *Pseudomonas syringae* 59
3.4.5 The progeny of infected plants have higher resistance to the infection with *Phytophthora nicotianae* 60
3.4.6 The second generation of plants exposed to virus also showed higher spontaneous HRF 60
3.4.7 PI_2 plants also showed lower viral titers 61
3.4.8 PI_1, PI_2 and PI_1_C1 plants were more tolerant to methyl methane sulfonate 62
3.4.9 The progeny of infected plants exhibited global genome hypermethylation 62
3.4.10 Infection with *TMV* and *P. syringae* results in a stronger induction of the PR1 gene in PI_1 plants 63
3.5 Discussion 64
3.5.1 Local infection of SR1 tobacco plants with *TMV* leads to heritable changes in genome stability 64
3.5.2 The progeny of infected plants exhibit higher resistance to *TMV* infection and MMS 65
3.5.3 Higher resistance to stress could be a result of enhanced innate immunity 66
3.5.4 Possible mechanisms of transgenerational changes in the progeny of stressed plants 68
3.6 References: 70
4.0 Progeny of tobacco mosaic virus-infected *Nicotiana tabacum* plants exhibit trans-generational changes in metabolic profiles 96
4.1 Abstract 96
4.2. Introduction 97
4.3 Experimental procedures 99
4.3.1. Infection of SR1 plants with *TMV* and generation of the progeny of infected plants 99
4.3.2. Infection of the progeny of infected and progeny of control plants with *TMV* 100
4.3.3 Metabolite extraction 101
4.3.4. NMR sample preparation and spectroscopy 101
4.3.5 NMR compound identification and quantification 102
4.3.6 GC-MS compound identification and quantification 103
4.4 Results 105
4.4.1 Experimental set-up and generation of samples 105
4.4.2 NMR Identification and Quantification 106
4.4.3 GC-MS Identification and Quantification 106
4.4.4 The progeny of infected plants exhibited substantial differences in metabolite levels 108
4.5 Discussion 109
4.6 References: 115
5.0 Development of novel reporter gene-based constructs for the study of sequence- specific recombination rate 129
5.1 Abstract 129
5.2 Introduction 130
5.3 Materials and methods 135

5.3.1 Plant growth conditions 135
5.3.2 Plant screening for antibiotic resistance 135
5.3.3 Construct design..... 136
5.3.4 Selection of sequences 136
5.3.5 Artificial synthesis 137
5.3.6 The Gateway-based cloning system..... 138
5.3.7 Plant transformation by the Agrobacterium-mediated gene delivery system 139
5.3.8 CCD analysis 140
5.3.9 Direct whole-mount immunohistochemistry 141
5.3.10 Western blot analysis 142
5.3.11 PCR confirmation of recombination events in somatic tissues 143
5.4 Results..... 144
5.4.1 Synthesis of binary vectors 144
5.4.2 Transient expression analysis 145
5.4.3 Stable expression analysis..... 146
5.4.4 Reporter gene analysis 146
5.4.5 Analysis of recombination event in the transgene 147
5.5 Discussion 147
5.6 References:..... 154
6.0 Conclusion 170
6.1 References:..... 175

List of Figures:

Figure 3.1: Schematic representation of the experimental set-up 77
Figure 3.2: Increased frequency of recombination in the first progeny of infected plants 79
Figure 3.3: PI_1 plants show a higher increase in HRF upon exposure to abiotic stress 80
Figure 3.4: PI_1 and PI_2 plants exhibit delayed symptoms and a lower viral titer upon *TMV* infection 82
Figure 3.5: PI_1 plants exhibit higher resistance to P. syringae and P. nicotianae 84
Figure 3.6: The Progeny of virus-infected plants exhibit higher tolerance to MMS 86
Figure 3.7: PI plants exhibit methylation changes 87
Figure 3.8: PI_1 plants show a higher level of PR1 expression and higher callose deposition 89
Figure 3.9: PI_1 plants show a higher callose deposition 92
Supplementary Figure 3.1: Reporter gene based recombination construct for analysis of sequence independent recombination rate. 93
Supplementary Figure 3.2: Analysis of polyphenols in progeny of infected lines. 94
Supplementary Figure 3.3: Total secondary metabolite content in PI and PC plants 95
Figure 4.1: Typical 500 MHz 1H-NMR spectrum of a Tobacco leaf sample. 121
Figure 4.2: Total metabolite levels, as measured by NMR 122
Figure 4.3: Comparison of the average metabolite concentrations in non-infected PC, PI #9 and PI #19 plants, as measured by NMR and GC-MS 123
Figure 4.4: Comparison of fold difference in the amount of metabolites as measured by NMR and GC-MS 124
Figure 4.5: the concentrations of amino acids, as measured by NMR 125
Figure 4.6: the concentrations of sugars and metabolites involved in sugar metabolism in PC and PI plants, as measured by NMR or GC-MS 126
Figure 4.7: The concentrations of other metabolites in PC and PI plants, as measured by NMR or GC_MS 127
Figure 4.8: The concentrations of other metabolites in PC and PI plants, as measured by NMR or GC-MS 128
Figure 5.1: A schematic diagram representing the design of a sequence-specific recombination construct. 158
Figure 5.2: An overview of various cloning stages. 159
Figure 5.3: Transient expression analysis using a CCD camera. 160
Figure 5.4: Transient expression analysis using immunohistochemistry 162
Figure 5.5 Stable expression of recombination constructs in Arabidopsis 164
Figure 5.6: A luciferase expression pattern in homozygous transgenic plants 165
Figure 5.7: Western blot analysis of the HA-tag and luciferase proteins 166
Figure 5.8: Schematic representation of a PCR approach for identification of recombination events in the transgene 167
Figure 5.9: PCR confirmation of recombination events in the transgenes 168
Supplementary figure 5.1: Restriction analysis of binary vectors obtained after LR reactions 169

List of tables:

Table 3.1: Black shank disease assay shows higher resistance of PI_1 plans to *P. nicotianae*.....76
Table 4.1: Unique metabolites detected by NMR and GC-MS.....120

List of Abbreviations:

5-MeC	5 Methyl cytosine
ANOVA	Analysis of variance
ATP	Adenosine triphosphate
Avr	Avirulent
BHT	Butyrated hydroxytoluene
BSA	Bovine serum albumin
CAM	Crassulacean acid metabolism
CaMV	Cauliflower mosaic virus
CC	Coiled coil
CCD	Charged couple device
CDPK	Calcium dependent protein kinases
DAPI	4', 6-diamidino-2-phenylindole
Dpi	Days post infection
DSB	Double strand break
DSS	Disodium-2, 2-dimethyl-2-silapentane-5-sulphonate
EI	Electron impact
GM	Genetically modified
CP	Coat protein
GC-MS	Gas chromatography - mass spectroscopy
GTP	Guanosine triphosphate
GUS	β-Glucuronidase
HA-tag	Hemagglutinin tag
Hpi	Hours post infection
HR	Homologous recombination
HRF	Homologous recombination frequency
IHC	Immunohistochemistry
IM	Inoculation media
LRR	Leucine rich repeat region
LUC	Luciferase
MMS	Methyl methane sulfonate
MP	Movement protein
MSTFA	N-Methyl-N-trifluoroacetamide
MUG	4-methyl umbelliferyl glucuronide
NBS	Nucleotide binding site region
NHEJ	Non homologous end joining
NIST	National Institute of Standards and Technology
NMR	Nuclear magnetic resonance
NO	Nitric oxide
ORF	Open reading frame
PBS	Phosphate buffered saline
PC	Progeny of control
PCD	Programmed cell death
PCR	Polymerase chain reaction

PFA	Paraformaldehyde
PFG	Pulsified field gradient
PI	Progeny of infected
Ppm	Particulate per million
PR gene	Pathogenesis related gene
Pst	*Pseudomonas syringae* pv *tomato*
R gene	Resistance gene
RB	Rose Bengal
RFLP	Restricted fragment length polymorphism
ROS	Reactive oxygen species
SA	Salicylic acid
SAR	Systemic acquired resistance
SqMV	Squash mosaic virus
TCA	Tricarboxylic acid
TIR	Toll/interleukin-1 receptor
TMCS	Trimethylchlorosilane
TMV	Tobacco mosaic virus
TYMV	Turnip yellow mosaic virus
X-gluc	5-bromo-4-chloro-3-indolyl glucuronide
YEP	Yeast extract peptone
ZYMV	Zucchini yellow mosaic virus

1.0 Introduction

Plants being unable to evade stresses just by moving away from them have to respond to stressful conditions in unique ways. One of the major stresses faced by plants during their lifespan is biotic stress, and it is in the form of pathogens (Chen et al., 2008). Two types of plant pathogen interactions are possible depending on the presence or absence of plant resistance (*R*) genes. If an *R* gene from the pathogen is able to recognize the invading pathogen, it will lead to programmed cell death in localized tissues. This will result in localization of the pathogen, and such interaction is known as an incompatible plant-pathogen interaction (Shirasu and Schulze-Lefert, 2000). Contrarily, if there is no *R* gene against a given pathogen in a plant, this pathogen will be successful in infecting the plant and causing a disease. Such interaction is known as a compatible plant-pathogen interaction (Hammond-Kosack and Jones, 1997; Keen, 1990).

The R gene consists of several conserved regions, including the nucleotide binding site (NBS) region responsible for signal transduction and the leucine rich repeats (LRR) region responsible for specific recognition of the pathogen (Dangl and Jones, 2001; Traut, 1994). The NBS regions are evolutionarily well-conserved regions, while the LRR regions are highly variable in their DNA sequence.

Incompatible interactions represent a more explored area of plant pathogen interactions, while less knowledge is available about compatible interactions. During one such study of compatible interactions, a novel systemic signal was identified (Kovalchuk et al., 2003). The

signal travels faster than the pathogen and leads to an increase in somatic and meiotic recombination frequencies in the plant (Kovalchuk et al., 2003). Subsequent studies revealed many other changes triggered by this still enigmatic signal. In the progeny of pathogen-infected plants, the level of global genome DNA methylation increases. Also, the LRR regions of *R* genes were shown to be hypomethylated and more prone to rearrangements in the progeny of these infected tobacco plants (Boyko et al., 2007). It can be hypothesized that these rearrangements might be due to the upregulated levels of somatic and meiotic recombination in plants. Such increased frequency of recombination would result in the occurrence of novel LRR regions, which in turn would increase the probability of converting an incompatible interaction into a compatible one. Hence, it was important to study the rate of recombination in LRR regions following compatible pathogen infection. It was also necessary to analyze what other changes can be found in the genome and in the physiology of the progeny of infected plants. It was also of great importance to find out whether the progeny of infected plants would respond to pathogen infection in a different manner.

To seek answers to all the above mentioned questions, we conducted various studies involving generation of the signal via a compatible plant pathogen interaction followed by subsequent analysis of changes in the progeny. Tobacco plants cultivars SR1 that do not carry a resistance gene against tobacco mosaic virus were infected with *TMV*, and the progeny of infected tobacco plants was collected. The rate of *TMV* propagation in the progeny of infected and control lines was analyzed. The progeny of infected plants showed a lower viral titer and slower progression of viral infection. Also, significant changes in the expression of pathogenesis-related (PR) genes, in the callose deposition process as well as in the total polyphenolics content were observed. Hence, the signal generated by compatible

infection leads to transgenerational changes which alter the response of the progeny to pathogen infection.

The second direction of research involved analysis of the level of various secondary metabolites in the progeny of infected and control tobacco plants. The results showed that the progeny of infected plants contains a significantly higher amount of metabolites as compared to that of controls. The higher amount of secondary metabolites may enhance the plant's ability to cope with pathogen infection.

In the last part of the study, the efforts were initiated to analyze changes in the recombination rate in the LRR regions. To analyze changes in specific genomic regions, we designed constructs that carry the LRR regions from various resistance genes. Since these regions were cloned in duplicates, they allowed recombination events to occur. Control constructs were made of the NBS regions of these resistance genes as well as unspecific regions of house-keeping genes. Transgenic Arabidopsis plants carrying the constructs have been produced, and the functionality of these constructs has been confirmed.

1.1 References:

Boyko, A., Kathiria, P., Zemp, F.J., Yao, Y., Pogribny, I., and Kovalchuk, I. (2007). Transgenerational changes in the genome stability and methylation in pathogen-infected plants: (virus-induced plant genome instability). Nucleic Acids Res *35*, 1714-1725.

Chen, H.L., Wang, Y.Q., Chu, C.C., and Li, P. (2008). Plant non-host resistance: current progress and future prospect. Yi Chuan *30*, 977-982.

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

Hammond-Kosack, K.E., and Jones, J.D. (1997). Plant Disease Resistance Genes. Annu Rev Plant Physiol Plant Mol Biol *48*, 575-607.

Keen, N.T. (1990). Gene-for-gene complementarity in plant-pathogen interactions. Annu Rev Genet *24*, 447-463.

Kovalchuk, I., Kovalchuk, O., Kalck, V., Boyko, V., Filkowski, J., Heinlein, M., and Hohn, B. (2003). Pathogen-induced systemic plant signal triggers DNA rearrangements. Nature *423*, 760-762.

Shirasu, K., and Schulze-Lefert, P. (2000). Regulators of cell death in disease resistance. Plant Mol Biol *44*, 371-385.

Traut, T.W. (1994). The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites. European Journal of Biochemistry *222*, 9-19.

2.0 Literature Review

2.1 Introduction

With augmentation of molecular biology in twenty-first century, our understanding of biology has furthered like never before. Now we have increased understanding of life forms including processes like seed germination, plant development and aging. One of the interesting phenomena is how organisms respond to stress. The field of stress biology addresses the interaction between different environmental stimuli of abiotic and biotic nature and internal factors of an organism. Plants present interestingly and distinct patterns of stress response since they are restricted in stress avoidance due to immobility.

The major stresses that plants encounter are in the form of pathogens. As a food source, plants play an important role in present day economy. It is estimated that 13% of food crops are lost annually to preharvest and postharvest diseases, with additional reductions in the monetary value (Montesinos et al., 2002). Approximately 11,000 plant diseases have been characterized. There are 120 genera of fungi, 30 types of viruses, and 8 genera of bacteria which are the causal organisms of these diseases. It has been suggested that fungi are responsible for the majority of plant diseases as opposed to animal diseases (Montesinos et al., 2002). In addition, there are many organisms that establish symbiotic interactions with plants. Such organisms include nitrogen fixing bacteria and specialized soil fungi, mycorrhizae.

One of the major goals of plant research nowadays is to develop plants which are resistant to attack by biotic factors. Various genetically modified (GM) and non-GM based approaches

are being developed. Such approaches provide promise for future development of pathogen-resistant plant varieties. To achieve the desired goals, thorough understanding the details of plant pathogen interactions is of the utmost importance.

2.2 Different pathogen resistance mechanisms in plants

Interactions between plants and other organisms can be classified into three types: beneficial, mutualistic or harmful. Harmful interactions may comprise of those with fungi, bacteria and viruses. Most mycoplasma are also considered to be pathogenic parasites.

Resistance to a pathogen is defined as an ability of a plant to deter the initiation of pathogen infection, multiplication or systemic movement. As each pathogen requires certain growth conditions provided by the host plant, most plants, being not a permissive host, are resistant to the potential pathogens. Such resistance is termed as non-host resistance or basic resistance (Chen et al., 2008). Resistance imparted by induction of specific plant processes upon microbe recognition is a part of active resistance.

If, in the process of plant pathogen interaction, a plant is able to prevent pathogen infection or spread, such interaction is termed incompatible. Conversely, if a pathogen is able to infect, multiply and spread in a host plant, the interaction is termed compatible (Hammond-Kosack and Jones, 1997).

2.2.1 Non-host resistance

There are many barriers a pathogen encounters during plant invasion and pathogenesis. The very first barrier is in form of the structural and chemical composition of the plant cell wall. For example, in rust fungi the surface topography and wax composition of the cell wall determine the ability to induce growth and subsequent multiplication of appressorium (Hoch et al., 1987; Tsuba et al., 2002). It has been established quite early that the cell wall composition and the presence of specific chemicals can hinder the growth of pathogens (Papadopoulou et al., 1999). Avenacin A-1, an antimicrobial saponin, is present in the roots of oats. *Gauemannomyces graminis* var. *tritici* can infect wheat but is unable to attack oats as it cannot detoxify the saponin; hence, *G. graminis* var. *tritici* is a non-host pathogen for oats (Papadopoulou et al., 1999).

The process of non-host resistance also includes induction of general barrier mechanisms following non-specific pathogen recognition by plants. This process is based on recognition of general elicitors produced by pathogenic fungi, bacteria or viruses. As a result, a cell can recognize the presence of fungus, bacterium or virus but not the specific strain of pathogen. Hence, such recognition is termed non-specific recognition. A good example of it is papillae formation at the site of fungal penetration. Plant cells produce a structural barrier in the form of papillae on the inner surface of the cell wall upon the recognition of general elicitors from fungal pathogens (Wei et al., 1998).

Non-host resistance is efficient in preventing infections by a large number of pathogens which are non-host pathogens for particular plants. It also decreases infection caused by host

pathogens up to a certain level. However, efficient resistance against host pathogens is provided by specific gene-to-gene interaction or active immunity.

2.2.2 The mechanism of active resistance in plants

Once a pathogen overcomes the initial structural and chemical resistance mechanisms and succeeds in entering a plant cell, plant resistance depends on the specific identification of the pathogen by the infected plant cell. Upon pathogen recognition, counter mechanisms to stop its growth can be implemented. This process relies on the interaction between gene products from the pathogen and the host.

Pathogens produce certain proteins known as virulence factors, products of avirulent (*Avr)* genes that are unique for each pathogen. They are required for penetration into host cells, and they are also necessary to colonize the host (Van Der Biezen and Jones, 1998). Products of *Avr* genes are identified by receptors in the host cell that are products of resistance (*R*) genes. Hence, this form of immunity entirely depends on the presence of both, *Avr* and *R* genes (Keen, 2000).

2.2.2.1 Incompatible interactions

The most efficient way for a plant to evade pathogen infection is via a gene-for- gene interaction. This kind of interaction is based on *Avr* gene products from pathogens and *R* genes from plants. Following pathogen recognition, there are many changes that take place

in the host cell. In most of the plants, *R* gene-mediated resistance leads to the appearance of necrotic lesions produced upon massive production of free radicals leading to the death of infected and neighboring non-infected cells. This phenomenon is known as a programmed cell death (PCD) or hypersensitive response (Shirasu and Schulze-Lefert, 2000). Such a response is believed to deter the growth of pathogens in infected tissue. However, there are few incompatible interactions where the hypersensitive response is not present. They include the interaction between *Erysiphe graminis* sp *hordei* and the *mlo* gene in barley (Knogge, 1996). Another example is *Rx* gene-mediated resistance to potato virus X in potato (Kohm et al., 1993).

During incompatible interactions, there are other responses associated with the hypersensitive response. They include the generation of reactive oxygen species (ROS), changes in ion flux, the activation of signaling pathways, the alteration in transcription profile, and nitric oxide (NO) production (Gozzo, 2003).

The generation of ROS has been suggested as a trigger required for the initiation of hypersensitive response (Doke et al., 1996). Most of the superoxide ions (O_2^-) produced are quickly converted into H_2O_2 either non-enzymatically or with the help of specific enzymes (like superoxide dismutase). Hence, a large amount of H_2O_2 is detected during the onset of incompatible interactions (Levine et al., 1994; Nurnberger et al., 1994). H_2O_2 molecules are able to pass through membranes and can be converted into other molecules which have the potential to cause membrane damage (Halliwell and Gutteridge, 1990). The ROS burst is also associated with the induction of defense-related genes and is able to increase cross-linking in the cell wall (Vranova et al., 2002). The function of NO is similar to that of ROS

and is required in PCD just as the induction of defense-related genes (del Río et al., 2004). A calcium ion is one of the important signaling molecules in cells. Changes in ion flux result in the activation of many signaling pathways. The calcium-binding activity can trigger phosphorylation of many calcium-dependent protein kinases (CDPK) which are involved in pathogen response (Rudd and Franklin-Tong, 2001).

In addition, the production of phytoalexins around the infected tissue has also been noted. Phytoalexins are low-molecular-weight lipophilic antimicrobial compounds (Smith, 1996). Phytoalexin induction is associated with the activation of many enzymes and hence requires highly coordinated signaling pathways. The antimicrobial function of phytoalexins has long been established *in vitro*, yet their role during the initial phases of incompatible interactions is still not clear (Mauch-Mani and Slusarenko, 1996).

At present, it is predicted that the activation of preexisting protein kinases, phosphatases and G proteins is the very first step after pathogen recognition (Staskawicz et al., 1995). The next step is changes in calcium and other ion fluxes and phosphorylation/dephophorylation of various cellular proteins (Low and Merida, 1996). The induction of these initial processes leads to the development of hypersensitive response and various other phenomena.

Hence, an incompatible interactions result in a complex cross-talk between different pathways. However, the end result of any incompatible interaction is the same: the exclusion of the invading pathogens and removal of already infected plant tissues.

2.2.2.2 Systemic acquired resistance (SAR)

Early studies of plant pathogen interactions revealed that following PCD induced by an incompatible interaction, distal uninfected tissues of plants become more resistant to secondary infection. Such induced resistance is marked by upregulation of many disease-related genes. This type of secondary resistance following the incompatible interaction which spreads from a local site of infection to systemic plant tissues is termed systemic acquired resistance (SAR) (Ross, 1961). It helps plants evade pathogens quicker and can last for the rest of the plant's life. Secondary resistance also provides broad-spectrum and efficient resistance against a large number of fungal, viral and bacterial pathogens (Ryals et al., 1996; Sticher et al., 1997).

The onset signature of SAR is induced expression of the pathogenesis-related (*PR*) gene family. The genes encoding the PR proteins are induced in both the local area of infection and in distant tissues. In early experiments, infection of tobacco plants with tobacco mosaic virus (*TMV*) led to the detection of various novel proteins which revealed the existence of *PR* genes (van Loon, 1975; Van Loon and Van Strien, 1999). The functions of all classes of PR proteins are still unknown and need to be characterized. It is thought that resistance due to the SAR mechanism is not an outcome of a single PR protein but is the result of a combined effect of genes encoding many PR proteins.

In addition to pathogens, SAR can be induced by external application of salicylic acid (SA), aspirin, and benzoic acid. These observations lead to the development of a hypothesis that SA is a systemic signal which spreads SAR from the infected site to the rest of the plant

(White, 1979). This hypothesis was supported by the finding that the concentration of endogenous SA increases in local and systemic tissues as well as in the phloem in response to pathogens (Malamy et al., 1990; Metraux et al., 1990). The infected leaf produces 70% of SA present in the systemic tissue (Shulaev et al., 1995). These experiments indicate that SA plays a central role in SAR induction. However, the validity of the function of SA as a systemic signaling molecule is still debated. Experiments were carried out using the bacterial *nahG* gene which codes for enzyme salicylate hydroxylase that converts SA to catechol; hence reducing the level of SA in the plant (Delaney et al., 1994; Gaffney et al., 1993). The plants unable to accumulate SA were fully capable to produce and translocate a signal to other tissues (Vernooij et al., 1994). Recent experiments have added to the controversy about SA. It has been proposed that the signal might be lipid-based. Plants with a mutation in *Dir1*, which has sequence similarity to lipid transfer proteins, are not able to develop SAR or induce PR proteins (Maldonado et al., 2002). While the signal molecule for SAR is under review, the function of SAR in resistance is well understood.

2.2.2.3 The role of *PR* genes

Following an incompatible interaction, the development of SAR is marked by induction of *PR* genes. These genes were first identified in tobacco during incompatible interactions. At the time these proteins can only be associated with a pathogen attack, and hence they were named *PR* genes (Antoniw et al., 1983). PR proteins are induced systemically in infected plants. This induction can be due to infection with bacteria, fungi or viruses (van Loon, 1985). In addition to pathogens, PR proteins are also induced by application of synthetic

chemicals like polyacrylic acid and aspirin (Bol et al., 2003; Dumas et al., 1987). Also, phytohormones, like auxins, cytokinins and ethylene, are able to induce the expression of *PR* genes (Hughes and Dickerson, 1991; Memelink et al., 1990; van Loon, 1983). However, only a pathogen is able to induce the majority of the *PR* genes, while any abiotic or another biotic factor induces either single or few *PR* genes.

There are five major classes of *PR* genes. The first proteins identified belong to the PR-1 group:, PR-1a, PR-1b and PR-1c, have been the subject of the majority of initial work in the field (Jamet and Fritig, 1986). More than 90% homology has been found between these proteins in tobacco and PR-1 proteins from other plant species (Nasser et al., 1988; Nassuth and Sänger, 1986). Constitutive expression of the PR-1 group genes leads to higher resistance to pathogens, especially *TMV* (Stintzi et al., 1993). The PR-1 proteins are antifungal in nature. In various studies, they have been shown to provide plants with efficient protection against *Phytophthora infestans* and *Peronospora tabacina* (Niderman et al., 1995; Ryals et al., 1996).

Other groups of PR proteins have been identified relatively recently. The second group, *PR-2* genes, includes members that possess glucanase activity. They are able to impart protection against pathogens like *Phytophthora megasperma* (Boller et al., 1983; Pegg and Young, 1982). In contrast, the third PR-3 group is characterized by chitinase activity. Hence, PR-3 proteins are able to impart resistance against both insects and fungi (Legrand et al., 1987). The PR-4 group is one of the least understood groups. There are just four members in it, and they all are present in the extracellular spaces (Friedrich et al., 1991). These proteins are wound-inducible, but their biological function is yet to be discovered. The PR-5 group

consists of antifungal proteins with varied functions. They are able to penetrate fungal membrane and act to stop the growth of hyphae (Hejgaard et al., 1991; Woloshuk et al., 1991). Further details of their functions are still under study.

2.2.2.4 The role of callose induction following pathogen infection

During a pathogen attack, one of the resistance mechanisms is to mount physical and chemical barriers against invading pathogens in the form of a papilla. This barrier is in form of papilla formation. Such papilla consists of callose and antimicrobial compounds (Schmelzer, 2002; Sisler, 1986).

Callose is a β- 1, 3 glucan involved in a wide range of processes in plant cells (Chen and Kim, 2009). It is involved in multiple stages of pollen development and pollen tube formation (McCormick, 1993). It is also deposited at the cell plate during the process of cytokinesis (Hong et al., 2001; Samuels et al., 1995). In terms of pathogen attack, the most relevant function is that it is deposited at plasmodesmata and is involved in the control of cell-to-cell movement (Bucher et al., 2001; Iglesias and Meins, 2000). Callose is synthesized on the outer plasma membrane of the cell and then translocated to plasmodesmata or papillae (Bhat et al., 2005). In addition to callose, other compounds are present in papillae, such as phytoalexins, phenolics, H_2O_2, as well as specific proteins including enzyme inhibitors (An et al., 2006).

Besides its participation in basal defense, callose accumulation has been discovered to play a role in hypersensitive response against viruses, fungi and bacteria. Prior to lesion appearance, callose accumulation can be found in and around infected tissue. This process results in containment of the pathogen in infected tissue (Atabekov and Dorokhov Yu, 1984; Shimomura and Dijkstra, 1975). In addition to providing pathogen protection, such cellular isolation leads to hindering of cellular function and ultimately cell death.

Recently, glucans have been found to have antimicrobial activity as well (Bucher et al., 2001). This indicates towards yet another role of callose in pathogen resistance.

2.2.2.5 The role of polyphenols and secondary metabolites in pathogen resistance

Polyphenols, being common secondary metabolites present in plants, are widely known for their antioxidant properties (Manach et al., 2004). According to structure, they are classified into many different types. Phenylpropanoids, flavonoids, coumarins, phloroglucinols and xanthones are good examples of them such types (Yazaki et al.). Antibacterial and antiviral activities, the production of anti-oxidants, anti-tyrosinase, anti-nitric oxide, estrogen sulfotransferase inhibition – these are some examples of these functions (Appendino et al., 2008; Dong et al., 2007; Kapche et al., 2009; Lee et al., 2007; Mesia-Vela et al., 2001; Miranda et al., 2000).

Out of all the known polyphenols, flavonoids constitute the best characterized group. The chemical structure of flavonoids consists of two aromatic cycles and one heterocycle. Based

on properties of heterocycles, they are further classified into flavonols, anthocynins and flavan-3-ols. Modifications of the aromatic rings of flavonoid compounds produce a diverse array of secondary metabolites, such as proanthocyanidins also known as condensed tannins (Dixon et al., 2005; Marles et al., 2003).

Flavonoids are very essential for many plant-environment interactions. For example, they are efficient in ultraviolet protection. They have antimicrobial properties and help plants avoid herbivory (Dixon et al., 2005; Marles et al., 2003). The properties of flavonoids are derived from their cytotoxic effects and ability to physically interact with cellular enzymes.

Some polyphenols are also well known for their antimicrobial role properties. For example, quinones are very powerful antibiotics and also repel plant feeding insects by causing starvation (Felton et al., 1992; Schweigert et al., 2001). Avenacins are effective against *Gaeumannomyces graminis* var. *tritici* in oat, and a similar compound α-tomatine from tomato imparts efficient protection against *Fusarium solani* (Bednarek and Osbourn, 2009). Resvitols from tobacco and 2, 4-dihydroxy-7-methoxy-1, 4-benzoxazin-3-one from maize provide a wide spectrum of pathogen protection (Frey et al., 1997; Hain et al., 1993).

Hence, polyphenols from plants are vital for various plant-pathogen interactions that result in resistance development in plants.

2.2.2.6 Conserved regions in the disease resistance genes

In addition to a very high number in the plant genome, the *R* genes are characterized by a very high structural diversity. Although there is much variation in the origin and function, they share a significant homology within DNA and protein sequences. Based on DNA sequences, plant *R* genes are divided into five major classes (Hammond-Kosack and Jones, 1997).

The first group is serine/threonine protein kinases which are cytoplasmic proteins. One of the most studied and characterized *R* gene in this group is *Pto* from tomato. The *Pto* gene is responsible for imparting resistance against the *AvrPto* gene from *Pseudomonas syringae* pv. *tomato* (Keen, 1990; Martin et al., 1993).

The second major type of *R* proteins consists of N-terminal coiled-coil (CC), nucleotide-binding site (NBS) and C terminal leucine-rich repeat (LRR) domains. The Rx protein from potato which confers resistance against potato virus X belongs to this class of intracellular proteins (Bendahmane et al., 1999).

NBS and LRR domains are also present in the third class of R proteins, but unlike N-terminal CC domains, they consist of a Toll/interleukin-1 receptor (TIR) which is similar to that found in Drosophila and mammals respectively. This class of *R* genes includes the *N* gene from tobacco which confers resistance against tobacco mosaic virus (Whitham et al., 1994; Whitham et al., 1996). The third class of *R* proteins is predicted to be cytoplasmic proteins.

The fourth class of *R* proteins possesses an extracellular domain. Cf protein from tomato has an extracellular N terminal LRR region which is attached to a transmembrane region followed by the short C-terminal cytoplasmic tail. The Cf protein provides resistance against *Cladosporium fulvum* (Dixon et al., 1996; Van den Ackerveken et al., 1992).

The last class of *R* gene products contains an extracellular N-terminal LRR region, a transmembrane region and the intracellular C-terminal serine threonine kinase domain. *Xa21* from rice is the best representative of this class of R genes (Liu et al., 2002).

2.2.2.7 The function of disease resistance genes

The overall function of disease resistance genes is to evade pathogen infection. They do so by inherent ability of specific pathogen recognition. Furthermore, they need to trigger a rapid resistance response. Based on evidence from different experiments, there are currently two hypothesized models of how R genes are able to function.

The very first model put forward in the 1970s was based on receptor-ligand relations between R proteins and *Avr* factors. This model implies that there are gene-on-gene interactions between these two factors. Also, it suggests that the *R* gene product directly interacts with the *Avr* gene product, and no immediate mediator genes are required for recognition (Flor, 1971). Although, at the time experiments supporting the receptor-ligand model were observed, there appeared studies indicating that such phenomena were rather

rare (Schneider, 2002; Van der Hoorn et al., 2002). Van der Biezen (1998) proposed a new model, known as the guard hypothesis, as an alternative model. According to this model, R proteins monitor some of the cellular proteins ("guardees") that are affected by the presence of pathogen. The interaction with pathogens leads to a change in conformation of these "guardees". Any change in conformation of a guardee is recognized by *R* genes and results in recognition of a pathogen. Hence, R proteins are not directly associated with the pathogen, they rather monitor other cellular components that are a substrate for pathogen virulence factors or those involved in early pathogen response (Van Der Biezen and Jones, 1998).

Each *R* gene consists of specific conserved regions in their DNA sequence. Each such domain is responsible for a specific role during its function. Most of the R proteins consist of NBS and LRR domains. NBS domains share sequence homology with mammalian cell death proteins. They are involved in the binding and hydrolysis of adenosine triphosphate (ATP) and guanosine triphosphate (GTP); hence, these domains are able to activate kinases (Traut, 1994). Because of these properties, NBS regions are believed to be part of specific signal transduction pathways during pathogen interaction.

The LRR domain is the most variable domain of R-proteins. As supported by various studies, it is believed to be a key player in determining specificity against avirulence factors (Dangl and Jones, 2001). Various alleles of the flax resistance gene *L* contain differences only in their LRR regions. These alleles give/confer resistance to different strains of the flax rust pathogen *Melampsora lini* (Ellis et al., 1999). Additionally, the analysis of mutations in the *Arabidopsis R* gene, *RPS2* (CC-NBS-LRR class) in the LRR region has revealed that these mutations lead to suppression of the function of multiple *R* genes. This suggests that

the LRR region is involved not only in signal perception but also in downstream signal transduction (Warren et al., 1998).

A functional role of the CC region has also been predicted. It has been shown to undergo homodimerization or heterodimerization. Hence, it is able to physically interact with other cellular proteins. R proteins containing the CC domain can be initially present as single molecules that later undergo dimerization upon pathogen challenge (Landschulz et al., 1988). The TIR domain could play a role in initiating a defense response and may be required for the production of ROS (Hammond-Kosack and Jones, 1996). The serine/threonine kinase domain of many R proteins can function in down-stream signaling. The kinase domain of the *Pto R* gene product is shown to induce phosphorylation of many other genes. This in turn activates the pathogen resistance response (Sessa and Martin, 2000).

Upon pathogen recognition by an R protein, activation of many downstream responses can occur. Programmed cell death (PCD), as depicted by localized death of infected cells or hypersensitive response-related death, is the most common response to pathogen infection (Greenberg and Vinatzer, 2003; Nimchuk et al., 2003). The death of infected and surrounding cells results in the formation of necrotic lesions. This prevents the further spread of the pathogen into the host.

2.2.2.8 The Evolution of disease resistance genes

There is a large number of *R* genes present in any plant genome. Strikingly, each *R* gene has recognition specificity towards very specific pathogens. Yet new *R* genes are being discovered in the well-established genomes like the genomes of Arabidopsis. This raises the question of how these different kinds of *R* genes are developed. To answer this question, evolutionary studies of *R* genes are required.

In any particular plant genome, the organization of *R* genes is very complex. One typical characteristic of *R* genes is that they are generally found in clusters, and genes within a cluster often belong to the same gene family. For example, *Xa21*, an *R* gene from rice, is a member of the *Xa* gene family which consists of at least eight other members (Song et al., 1995). The majority of the *Xa21* gene family members are present at the same locus on chromosome 11 of rice within a region of ~230 kb. At this locus, nine other *R* genes are also present which are not related to the *Xa* gene family (Ronald et al., 1992; Williams et al., 1996). In some cases, members of the same gene family may be distributed at different but structurally analogous loci. Two members of the *Cf-x R* gene family from tomato, *Cf-9* and *Cf-4,* are present on the short arm of chromosome 1, while two other closely related members, *Cf-2* and *Cf-5*, are present on chromosome 6 (Dixon et al., 1996). Hence, it has been interpreted that the majority of *R* genes are present at linked loci of related or unrelated *R* genes.

Another hypothesis has been developed regarding the rapid evolution of R genes. It has been proposed that followed by sequence diversification, duplication events may be the reason

behind the evolution of *R* gene families (Lawrence et al., 1995; Whitham et al., 1994). A molecular analysis of the *Cf-x* gene family revealed that gene duplication by unequal crossing-over played an important role in their evolution. Due to the use of unequally aligned sequences for recombination, the number and cluster composition of gene family members can change, which results in an increase in the variation present in the population (Parniske et al., 1997).

Patterns of nucleotide substitution have been studied for different domains of *R* genes. A nucleotide substitution which leads to a change in the amino acid code is known as a nonsynonymous mutation; whereas if there is no change in the primary structure of proteins caused by a nucleotide substitution, it is termed a synonymous substitution. In any given sequence, if the ratio of nonsynonymous to synonymous mutations is higher than one, it provides evidence for positive selection at that locus (Stahl and Bishop, 2000). The nonsynonymous to synonymous ratio has been found to be high in the LRR domain of *Cf* family *R* genes from tomato which codes for an extracellular LRR domain. In contrast, in the same *R* genes, the ratio was less than one in the sequences that do not code for the LRR domain (Parniske et al., 1997). A comparison of LRR regions between the family members *Xa21* and *Xa21D* has revealed that the rate of nonsynonymous substitution is significantly higher than that of synonymous substitutions. The ratio was less than one when substitution rates were calculated for the entire *R* gene (Wang et al., 1998). This suggests that the rate of LRR domain evolution of an R protein could be higher than that of other R protein domains.

2.3 Homologous recombination as a double-strand break repair mechanism

Organisms encounter various genotoxic threats from the environment, and double-strand breaks are a major concern because even a single unrepaired break can result in cell death (Bennett et al., 1996; Bennett et al., 1993). All organisms have evolved survival strategies to meet the needs of double strand break repair. Two conserved pathways are involved in double strand break processing: non-homologous end joining (NHEJ) and homologous recombination (HR). The NHEJ pathway is simple as it requires little or no homology for repair. Inherently, such repair is error-prone and can result in the deletion or insertion of DNA nucleotides (Krogh and Symington, 2004; Lees-Miller and Meek, 2003; Meek et al., 2004). On the other hand, homologous recombination is based on homology of sequences under repair. Hence, such repair has fewer mutations in repaired sequences, and is a more complex and accurate pathway for repair. In HR, the sequence information is given a higher priority (Puchta, 2005). However, the use of HR pathways for double strand break repair has potential for unequal crossing over and gene duplications. Thus, repair via both NHEJ and HR pathways can result in increased genetic variation. Furthermore, HR can recognize homologous sequences from adjacent genes, which may result in the formation of new recombinant genes.

Homologous recombination also plays an important role in meiotic recombination (Schuermann et al., 2005). It is responsible for the alignment and crossing over of homologous sequences as a result of synapsis. One of the requirements for the initiation of meiotic recombination is the creation of a double-strand break. Subsequent to double strand breakage, HR is responsible for the alignment and crossing over of homologous sequences

as a result of synapsis. Recombination in meiotic cells is an important factor leading to increased genetic diversity present in a population (Schuermann et al., 2005). Hence, HR serves two functions: in somatic cells, it stabilizes the genome via damage repair; in germline cells, it results in the formation of new allelic combinations and creation of additional genetic diversity.

2.3.1 The balance between genome stability and evolution

Maintaining genome stability is essential to any organism. Changes in the genome may lead to alterations in genes which are required for normal cell function. Such changes may also lead to phenotypic alterations which can be beneficial, neutral, or detrimental. If the alterations are detrimental, the organism has to avoid permanent changes as they can be lethal. On the other hand, such changes may result in developing characters that can ensure better survival of the organism under adverse conditions. It can be suggested that every organism has to be able to balance the integrity and flexibility of the genome in order to adjust to environmental conditions.

Adaptation to a constantly changing environment requires constant micro- and macro-evolutionary processes in plants. New phenotypes may arise as a result of mutations, including insertions, deletions and nucleotide substitutions. Duplication and rearrangements of entire genes caused by recombination events are only some of the factors producing major phenotypic changes. Many abiotic and biotic factors such as radiation, pathogens and mutagenic chemicals (Kovalchuk et al., 2003a; Kovalchuk et al., 2003b) can lead to

rearrangements in the genome. Increased rates of rearrangements can result in the creation of high phenotypic variability. This process can be mediated by HR and can lead to increased probability of individuals having higher resistance to stresses as a result of expression of novel stress-related genes.

2.3.2 The role of homologous recombination in plant-pathogen interactions

The effect of pathogens on plants is apparent in the form of either symptom development or resistance. However, the effect of pathogen infection on the host genome remains an interesting question. Lucht et al. (2002) suggested a direct relationship between pathogen stress and HR as infection with the *Peronospora parasitica* resulted in an increase in the frequency of HR (Lucht et al., 2002).

Kovalchuk *et al.,* (2003b) supported a link between pathogen infection and HR. In their experiments, they discovered a signal which spreads ahead of the pathogen in systemic plant tissues. This specific signal was termed a systemic recombination signal (SRS). SRS increases the frequency of HR in uninfected distal tissues and results in increased somatic recombination frequency (Kovalchuk et al., 2003b). This signal may direct higher rate of rearrangements in *R*-gene clusters, which can lead to the creation of novel *R*-gene recognition specificities of resistance to new pathogens. Hence, it can be hypothesized that plants can increase the frequency of rearrangements through the HR mechanism, which leads to the evolution of new *R*-gene specificities.

2.3.3 Changes in progeny of pathogen-infected plants

Compatible pathogen infection leads to an increase in the frequency of HR in uninfected tissues. Further studies related to this phenomenon were carried out by Kovalchuk et al (Boyko et al., 2007; Kovalchuk et al., 2003b). In their studies, transmission of changes in HR was checked in the next generation. It was observed that an increase in the frequency of recombination in somatic tissues of exposed plants was also noted in somatic tissues of the progeny of exposed plants suggesting in also affected germline tissue (Boyko et al., 2007). For the experiment, a transgenic line containing a reporter-based recombination construct was used. When a recombination event occurs within the transgene, it leads to expression of reporter gene. The rate of somatic recombination was calculated by counting the number of spots of reporter gene expression (Boyko et al., 2007).

The study also investigated changes in DNA methylation levels in the progeny of both infected and control plants. In the progeny of infected plants, the increased level of whole-genome DNA methylation was observed, as compared to the progeny of control plants. While studying DNA methylation in specific sequences, it was discovered that many sequences with high levels of homology to LRR sequences of *R* genes appeared to be hypomethylated. In contrast, housekeeping genes, such as actin, were either not methylated or underwent hypermethylation (Boyko et al., 2007).

Changes in methylation in a sequence-specific manner can be a precursor to changes in the rate of rearrangements in that sequence. Hence, the rate of rearrangements in the LRR-like sequences was evaluated. Interestingly, it was observed that indeed the rate of

rearrangements was higher in the progeny of infected plants (Boyko et al., 2007). This study showed that infection with a compatible pathogen results in genomic changes in infected plants, and these changes are further passed onto the next generation.

However, a few questions remained unanswered. It is still unclear whether rearrangements can be caused which type of mutations. The role of HR in establishment of pathogen response needs to be investigated. Direct and indirect impacts of all above-mentioned changes on actual pathogen infection also need to be studied.

2.3.4 Different designs of recombination constructs to study the rate of recombination

Homologous recombination has been a central issue of many scientific studies. To make such studies more efficient, indirect ways of monitoring the rate of HR have been developed. Some of the initial studies made use of plant tissue pigmentation as a reference (Christianson, 1975; Kawata and Cheung, 1990). The rate of HR was calculated using the appearance of light or dark pigmentation.

The availability of various reporter genes, such as firefly luciferase (LUC) and β-Glucuronidase (*uidA*/GUS), has facilitated the development of reporter-based constructs which consist of two truncated but overlapping copies of a reporter gene. In the first copy, the 3’ region of the reporter gene is removed, while in the second copy, the 5’ region is removed. These copies provide a substrate for HR proteins. HR in these regions will lead to the formation of a functional copy of a reporter gene. This can be visualized as a small patch

of tissue expressing the reporter gene. Hence, the rate of HR can simply be studied by counting such spots (Kathiria and Kovalchuk, 2010).

The extent of overlap between the two sequences is a crucial parameter. The length of the overlap should be at least 200 bp and preferably more than 400 bp (Li et al., 2004; Lyznik et al., 1991; Puchta and Hohn, 1991).

The above-mentioned construct design will result in gene expression as a consequence of either interchromosomal or intrachromosomal recombination. There also exists another variation of the construct in which the position of 5' and 3' truncated copies is switched. Such construct will lead to reporter gene expression only in the case of interchromosomal recombination (Li et al., 2004).

All recombination constructs used till now are based on truncated reporter genes. Hence, although they are able to reveal the rate of HR, they do not provide any information about the rate of HR in specific sequences of interest. Such constructs have been widely used especially as part of stress experiments. An increase in HR following various biotic and abiotic stresses has been recorded with their aid (Boyko et al., 2006a; Boyko et al., 2006b; Kovalchuk et al., 2003b).

2.4 *TMV* and its interaction with tobacco

TMV is one of the best-studied plant viruses. The research on *TMV* started in the first decade of the 20th century, and at present it has shown to be an ideal system for understanding the structure and function of virus particles. *TMV* is a rod-shaped virus, which is composed of positive-sense single-stranded RNA, coated with proteins. In an intact virus particle, this RNA is 6.4 kb in size and is helically bound to 2160 subunits of coat proteins. The open reading frame 1 (ORF1) coding sequence codes for the 126 kDa replicase protein which can also produce the 183 kDa protein by readthrough transcription. The ORF2 coding sequence produces the 30 kDa-movement protein (MP), while the ORF3 coding sequence codes for the 17.5 kDa-coat protein (CP) (Lewandowski and Dawson, 2000).

The mode of entry of *TMV* into plant cells is thought to be passive, through damaged cells. The process of infection involves the attachment of virus particles, entry into the cell followed by immediate uncoating of its RNA genome (Shaw, 1999). Once *TMV* enters the cell, it rapidly disassembles and initiates the transcriptional machinery leading to viral replication (Wu and Shaw, 1997). These processes occur at specific sites in the cell, which makes researchers believe that a virus particle or its RNA alone enters either through ectodesma or via pinocytosis (Gaard and de Zoeten, 1979; Hills et al., 1987; Shaw, 1999). The disassembly of viral RNA results in the production of MP. The RNA can now form either a MP complex for cell-to-cell movement or a complex with CP for the movement though the phloem and systemic infection (Hilf and Dawson, 1993). Systemic infection of tobacco with *TMV* results in the appearance of mosaic-like symptoms characterized by light and dark blotches on tissue.

Resistance genes conferring resistance to *TMV* have been identified in some tobacco cultivars. The N gene product from tobacco recognizes the *TMV* replicase protein, while the N'-gene recognizes *TMV* CP. The N' gene has not been thoroughly studied, but due to its specificity to CP, it is predicted that the N' gene functions in the same way as the *R* gene (Erickson et al., 1999).

2.5 Reporter genes and their uses

Expression analysis of many genes requires the use of reporter genes. Gene sequences that are easily detected when expressed in an organism are called reporter genes. These genes should not be normally present in the organism being studied. Reporter genes have been utilized for many purposes, one of which is to use them as marker genes. In transformation experiments, marker gene expression can be utilized for the identification of transformation events. Another important use of reporter genes is for analysis of the activity of various promoters.

The *uidA* (GUS) gene which codes for the enzyme β-glucuronidase was isolated from *Escherichia coli*. The GUS gene is one of the most widely used reporter genes in plants, and it can act on two specific substrates: 5-bromo-4-chloro-3-indolyl glucuronide (X-gluc) and 4-methyl umbelliferyl glucuronide (MUG). β-glucuronidase converts X-gluc to a blue chromophore which can be used for histological localization. Alternatively, MUG is converted to a fluorescent product by the β-glucuronidase gene and is utilized for measuring

the specific activity of an enzyme (Jefferson et al., 1987). The GUS gene is widely used in plants due to the lack of background GUS activity frequently observed in microorganisms, vertebrates and invertebrates (Gilissen et al., 1998). Using histochemical localization, the expression of this gene can be detected at the subcellular level. The GUS gene is ideal for gene fusion; it was used for many gene-tagging experiments resulting in the study of many genetic elements and regulatory elements (Fobert et al., 1994; Foster et al., 1999). The only disadvantage of the GUS reporter system is that the assay used to analyze gene expression is destructive to plants. Hence, it cannot be utilized for *in vivo* reporter gene function analysis.

Luciferase (LUC) is present in many unicellular and multicellular organisms. Firefly (*Photinus pyrelis*) luciferase is the most widely used reporter gene. This enzyme acts through the ATP-dependent oxidative decarboxylation of the substrate luciferin which produces a reaction accompanied by photon emission. A sensitive apparatus capable of detecting these photons can be utilized to measure the LUC activity. This assay is very sensitive and can be utilized for *in vivo* expression analyses as it does not require destructive methods for analysis (Millar et al., 1992; Ow et al., 1986). This system proves to be very efficient for gene expression analysis during various developmental stages. The half-life of LUC *in vivo* is very short, and it does not accumulate over time. Hence, it is the most efficient and accurate reporter gene to assay an indirect transcriptional activity based on gene expression analysis (Millar et al., 1992; Van Leeuwen et al., 2000). Yet when compared, the spatial expression of GUS can be analyzed more precisely than that of LUC.

2.6 References:

An, Q., Hückelhoven, R., Kogel, K.-H., and Bel, A.J.E.v. (2006). Multivesicular bodies participate in a cell wall-associated defence response in barley leaves attacked by the pathogenic powdery mildew fungus. Cellular Microbiology *8*, 1009-1019.

Antoniw, J.F., Ooms, G., White, R.F., Wullems, G.J., and v. Vloten-Doting, L. (1983). Pathogenesis-related proteins in plants and tissues of *Nicotiana tabacum* transformed by *Agrobacterium tumefaciens*. Plant Molecular Biology *2*, 317-320.

Appendino, G., Gibbons, S., Giana, A., Pagani, A., Grassi, G., Stavri, M., Smith, E., and Rahman, M.M. (2008). Antibacterial Cannabinoids from *Cannabis sativa*: A Structure−Activity Study. Journal of Natural Products *71*, 1427-1430.

Atabekov, J.G., and Dorokhov Yu, L. (1984). Plant virus-specific transport function and resistance of plants to viruses. Adv Virus Res *29*, 313-364.

Bednarek, P., and Osbourn, A. (2009). Plant-microbe interactions: chemical diversity in plant defense. Science *324*, 746-748.

Bendahmane, A., Kanyuka, K., and Baulcombe, D.C. (1999). The *Rx* Gene from Potato Controls Separate Virus Resistance and Cell Death Responses. Plant Cell *11*, 781-792.

Bennett, C., Westmoreland, T., Snipe, J., and Resnick, M. (1996). A double-strand break within a yeast artificial chromosome (YAC) containing human DNA can result in YAC loss, deletion or cell lethality. Mol Cell Biol *16*, 4414-4425.

Bennett, C.B., Lewis, A.L., Baldwin, K.K., and Resnick, M.A. (1993). Lethality induced by a single site-specific double-strand break in a dispensable yeast plasmid. Proc Natl Acad Sci U S A *90*, 5613-5617.

Bhat, R.A., Miklis, M., Schmelzer, E., Schulze-Lefert, P., and Panstruga, R. (2005). Recruitment and interaction dynamics of plant penetration resistance components in a plasma membrane microdomain. Proc Natl Acad Sci U S A *102*, 3135-3140.

Bol, J.F., Linthorst, H.J.M., and Cornelissen, B.J.C. (2003). Plant Pathogenesis-Related Proteins Induced by Virus Infection. Annual Review of Phytopathology *28*, 113-138.

Boller, T., Gehri, A., Mauch, F., and Vögeli, U. (1983). Chitinase in bean leaves: induction by ethylene, purification, properties, and possible function. Planta *157*, 22-31.

Boyko, A., Kathiria, P., Zemp, F.J., Yao, Y., Pogribny, I., and Kovalchuk, I. (2007). Transgenerational changes in the genome stability and methylation in pathogen-infected plants: (virus-induced plant genome instability). Nucleic Acids Res *35*, 1714-1725.

Bucher, G.L., Tarina, C., Heinlein, M., Di Serio, F., Meins, F., Jr., and Iglesias, V.A. (2001). Local expression of enzymatically active class I beta-1, 3-glucanase enhances symptoms of *TMV* infection in tobacco. Plant J *28*, 361-369.

Chen, H.L., Wang, Y.Q., Chu, C.C., and Li, P. (2008). Plant non-host resistance: current progress and future prospect. Yi Chuan *30*, 977-982.

Chen, X.Y., and Kim, J.Y. (2009). Callose synthesis in higher plants. Plant Signal Behav *4*, 489-492.

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

del Río, L.A., Javier Corpas, F., and Barroso, J.B. (2004). Nitric oxide and nitric oxide synthase activity in plants. Phytochemistry *65*, 783-792.

Delaney, T.P., Uknes, S., Vernooij, B., Friedrich, L., Weymann, K., Negrotto, D., Gaffney, T., Gut-Rella, M., Kessmann, H., Ward, E., *et al.* (1994). A Central Role of Salicylic Acid in Plant Disease Resistance. Science *266*, 1247-1250.

Dixon, M.S., Jones, D.A., Keddie, J.S., Thomas, C.M., Harrison, K., and Jones, J.D. (1996). The tomato Cf-2 disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. Cell *84*, 451-459.

Dixon, R.A., Xie, D.Y., and Sharma, S.B. (2005). Proanthocyanidins--a final frontier in flavonoid research? New Phytol *165*, 9-28.

Doke, N., Miura, Y., Sanchez, L.M., Park, H.J., Noritake, T., Yoshioka, H., and Kawakita, K. (1996). The oxidative burst protects plants against pathogen attack: mechanism and role as an emergency signal for plant bio-defence--a review. Gene *179*, 45-51.

Dong, X., Fan, Y., Yu, L., and Hu, Y. (2007). Synthesis of four natural prenylflavonoids and their estrogen-like activities. Arch Pharm (Weinheim) *340*, 372-376.

Dumas, E., Gianinazzi, S., and Cornu, A. (1987). Genetic aspects of polyacrylic-acid-induced resistance to tobacco mosaic virus and tobacco necrosis virus in *Nicotiana* plants. Plant Pathology *36*, 544-550.

Ellis, J.G., Lawrence, G.J., Luck, J.E., and Dodds, P.N. (1999). Identification of regions in alleles of the flax rust resistance gene *L* that determine differences in gene-for-gene specificity. Plant Cell *11*, 495-506.

Erickson, F.L., Holzberg, S., Calderon-Urrea, A., Handley, V., Axtell, M., Corr, C., and Baker, B. (1999). The helicase domain of the *TMV* replicase proteins induces the *N*-mediated defence response in tobacco. Plant J *18*, 67-75.

Felton, G.W., Donato, K.K., Broadway, R.M., and Duffey, S.S. (1992). Impact of oxidized plant phenolics on the nutritional quality of dietar protein to a noctuid herbivore, *Spodoptera exigua*. Journal of Insect Physiology *38*, 277-285.

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

Friedrich, L., Moyer, M., Ward, E., and Ryals, J. (1991). Pathogenesis-related protein 4 is structurally homologous to the carboxy-terminal domains of hevein, Win-1 and Win-2. Molecular and General Genetics MGG *230*, 113-119.

Gaffney, T., Friedrich, L., Vernooij, B., Negrotto, D., Nye, G., Uknes, S., Ward, E., Kessmann, H., and Ryals, J. (1993). Requirement of Salicylic Acid for the Induction of Systemic Acquired Resistance. Science *261*, 754-756.

Gilissen, L.J., Metz, P.L., Stiekema, W.J., and Nap, J.P. (1998). Biosafety of *E. coli* beta-glucuronidase (GUS) in plants. Transgenic Res *7*, 157-163.

Gozzo, F. (2003). Systemic Acquired Resistance in Crop Protection: From Nature to a Chemical Approach. Journal of Agricultural and Food Chemistry *51*, 4487-4503.

Halliwell, B., and Gutteridge, J.M. (1990). Role of free radicals and catalytic metal ions in human disease: an overview. Methods Enzymol *186*, 1-85.

Hammond-Kosack, K.E., and Jones, J.D. (1996). Resistance gene-dependent plant defense responses. Plant Cell *8*, 1773-1791.

Hammond-Kosack, K.E., and Jones, J.D. (1997). Plant Disease Resistance Genes. Annu Rev Plant Physiol Plant Mol Biol *48*, 575-607.

Hilf, M.E., and Dawson, W.O. (1993). The tobamovirus capsid protein functions as a host-specific determinant of long-distance movement. Virology *193*, 106-114.

Hoch, H.C., Staples, R.C., Whitehead, B., Comeau, J., and Wolf, E.D. (1987). Signaling for Growth Orientation and Cell Differentiation by Surface Topography in Uromyces. Science *235*, 1659-1662.

Hong, Z., Delauney, A.J., and Verma, D.P.S. (2001). A Cell Plate-Specific Callose Synthase and Its Interaction with Phragmoplastin. Plant Cell *13*, 755-768.

Hughes, R.K., and Dickerson, A.G. (1991). Modulation of Elicitor-Induced Chitinase and β-1,3-Glucanase Activity by Hormones in *Phaseolus vulgaris*. Plant Cell Physiol *32*, 853-861.

Iglesias, V.A., and Meins, F., Jr. (2000). Movement of plant viruses is delayed in a beta-1,3-glucanase-deficient mutant showing a reduced plasmodesmatal size exclusion limit and enhanced callose deposition. Plant J *21*, 157-166.

Jamet, E., and Fritig, B. (1986). Purification and characterization of 8 of the pathogenesis-related proteins in tobacco leaves reacting hypersensitively to tobacco mosaic virus. Plant Molecular Biology *6*, 69-80.

Jefferson, R.A., Kavanagh, T.A., and Bevan, M.W. (1987). GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J *6*, 3901-3907.

Kapche, G.D.W.F., Fozing, C.D., Donfack, J.H., Fotso, G.W., Amadou, D., Tchana, A.N., Bezabih, M., Moundipa, P.F., Ngadjui, B.T., and Abegaz, B.M. (2009). Prenylated arylbenzofuran derivatives from *Morus mesozygia* with antioxidant activity. Phytochemistry *70*, 216-221.

Kathiria, P., and Kovalchuk, I. (2010). Reporter Gene-Based Recombination Lines for Studies of Genome Stability (Humana press).

Keen, N.T. (1990). Gene-for-gene complementarity in plant-pathogen interactions. Annu Rev Genet *24*, 447-463.

Keen, N.T. (2000). A century of plant pathology: A Retrospective View on Understanding Host-Parasite Interactions. Annu Rev Phytopathol *38*, 31-48.

Knogge, W. (1996). Fungal Infection of Plants. Plant Cell *8*, 1711-1722.

Kohm, B.A., Goulden, M.G., Gilbert, J.E., Kavanagh, T.A., and Baulcombe, D.C. (1993). A Potato Virus X Resistance Gene Mediates an Induced, Nonspecific Resistance in Protoplasts. Plant Cell *5*, 913-920.

Kovalchuk, I., Filkowski, J., Smith, K., and Kovalchuk, O. (2003a). Reactive oxygen species stimulate homologous recombination in plants. Plant, Cell & Environment *26*, 1531-1539.

Kovalchuk, I., Kovalchuk, O., Kalck, V., Boyko, V., Filkowski, J., Heinlein, M., and Hohn, B. (2003b). Pathogen-induced systemic plant signal triggers DNA rearrangements. Nature *423*, 760-762.

Krogh, B.O., and Symington, L.S. (2004). Recombination Proteins in yeast. Annual Review of Genetics *38*, 233-271.

Landschulz, W.H., Johnson, P.F., and McKnight, S.L. (1988). The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science *240*, 1759-1764.

Lawrence, G.J., Finnegan, E.J., Ayliffe, M.A., and Ellis, J.G. (1995). The *L6* gene for flax rust resistance is related to the Arabidopsis bacterial resistance gene *RPS2* and the tobacco viral resistance gene *N*. Plant Cell *7*, 1195-1206.

Lee, J.C., Kundu, J.K., Hwang, D.M., Na, H.K., and Surh, Y.J. (2007). Humulone inhibits phorbol ester-induced COX-2 expression in mouse skin by blocking activation of NF-kappaB and AP-1: IkappaB kinase and c-Jun-N-terminal kinase as respective potential upstream targets. Carcinogenesis *28*, 1491-1498.

Lees-Miller, S.P., and Meek, K. (2003). Repair of DNA double strand breaks by non-homologous end joining. Biochimie *85*, 1161-1173.

Legrand, M., Kauffmann, S., Geoffroy, P., and Fritig, B. (1987). Biological function of pathogenesis-related proteins: Four tobacco pathogenesis-related proteins are chitinases. Proc Natl Acad Sci U S A *84*, 6750-6754.

Levine, A., Tenhaken, R., Dixon, R., and Lamb, C. (1994). H_2O_2 from the oxidative burst orchestrates the plant hypersensitive disease resistance response. Cell *79*, 583-593.

Lewandowski, D.J., and Dawson, W.O. (2000). Functions of the 126- and 183-kDa Proteins of Tobacco Mosaic Virus. Virology *271*, 90-98.

Li, L., Santerre-Ayotte, S., Boivin, E.B., Jean, M., and Belzile, F. (2004). A novel reporter for intrachromosomal homoeologous recombination in *Arabidopsis thaliana*. Plant J *40*, 1007-1015.

Liu, G.Z., Pi, L.Y., Walker, J.C., Ronald, P.C., and Song, W.Y. (2002). Biochemical characterization of the kinase domain of the rice disease resistance receptor-like kinase XA21. J Biol Chem *277*, 20264-20269.

Low, P.S., and Merida, J.R. (1996). The oxidative burst in plant defense: Function and signal transduction. Physiologia Plantarum *96*, 533-542.

Lucht, J.M., Mauch-Mani, B., Steiner, H.Y., Metraux, J.P., Ryals, J., and Hohn, B. (2002). Pathogen stress increases somatic recombination frequency in Arabidopsis. Nat Genet *30*, 311-314.

Malamy, J., Carr, J.P., Klessig, D.F., and Raskin, I. (1990). Salicylic Acid: A Likely Endogenous Signal in the Resistance Response of Tobacco to Viral Infection. Science *250*, 1002-1004.

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

Manach, C., Scalbert, A., Morand, C., Remesy, C., and Jimenez, L. (2004). Polyphenols: food sources and bioavailability. Am J Clin Nutr *79*, 727-747.

Marles, M.A., Ray, H., and Gruber, M.Y. (2003). New perspectives on proanthocyanidin biochemistry and molecular regulation. Phytochemistry *64*, 367-383.

Martin, G.B., Brommonschenkel, S.H., Chunwongse, J., Frary, A., Ganal, M.W., Spivey, R., Wu, T., Earle, E.D., and Tanksley, S.D. (1993). Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science *262*, 1432-1436.

Mauch-Mani, B., and Slusarenko, A.J. (1996). Production of Salicylic Acid Precursors Is a Major Function of Phenylalanine Ammonia-Lyase in the Resistance of Arabidopsis to *Peronospora parasitica*. Plant Cell *8*, 203-212.

McCormick, S. (1993). Male Gametophyte Development. Plant Cell *5*, 1265-1275.

Meek, K., Gupta, S., Ramsden, D.A., and Lees-Miller, S.P. (2004). The DNA-dependent protein kinase: the director at the end. Immunological Reviews *200*, 132-141.

Memelink, J., Linthorst, H.J.M., Schilperoort, R.A., and Hoge, J.H.C. (1990). Tobacco genes encoding acidic and basic isoforms of pathogenesis-related proteins display different expression patterns. Plant Molecular Biology *14*, 119-126.

Mesia-Vela, S., Sanchez, R.I., Estrada-Muniz, E., Alavez-Solano, D., Torres-Sosa, C., Jimenez, M., Estrada, Reyes-Chilpa, R., and Kauffman, F.C. (2001). Natural products isolated from Mexican medicinal plants: novel inhibitors of sulfotransferases, SULT1A1 and SULT2A1. Phytomedicine *8*, 481-488.

Metraux, J.P., Signer, H., Ryals, J., Ward, E., Wyss-Benz, M., Gaudin, J., Raschdorf, K., Schmid, E., Blum, W., and Inverardi, B. (1990). Increase in Salicylic Acid at the Onset of Systemic Acquired Resistance in Cucumber. Science *250*, 1004-1006.

Millar, A., Short, S., Hiratsuka, K., Chua, N.-H., and Kay, S. (1992). Firefly luciferase as a reporter of regulated gene expression in higher plants. Plant Molecular Biology Reporter *10*, 324-337.

Miranda, C.L., Aponso, G.L., Stevens, J.F., Deinzer, M.L., and Buhler, D.R. (2000). Prenylated chalcones and flavanones as inducers of quinone reductase in mouse Hepa 1c1c7 cells. Cancer Lett *149*, 21-29.

Montesinos, E., Bonaterra, A., Badosa, E., Frances, J., Alemany, J., Llorente, I., and Moragrega, C. (2002). Plant-microbe interactions and the new biotechnological methods of plant disease control. Int Microbiol *5*, 169-175.

Nurnberger, T., Nennstiel, D., Jabs, T., Sacks, W.R., Hahlbrock, K., and Scheel, D. (1994). High affinity binding of a fungal oligopeptide elicitor to parsley plasma membranes triggers multiple defense responses. Cell *78*, 449-460.

Ow, D.W., de Wet, J.R., Helinski, D.R., Howell, S.H., Wood, K.V., and Deluca, M. (1986). Transient and Stable Expression of the *Firefly Luciferase* Gene in Plant Cells and Transgenic Plants. Science *234*, 856-859.

Papadopoulou, K., Melton, R.E., Leggett, M., Daniels, M.J., and Osbourn, A.E. (1999). Compromised disease resistance in saponin-deficient plants. Proc Natl Acad Sci U S A *96*, 12923-12928.

Parniske, M., Hammond-Kosack, K.E., Golstein, C., Thomas, C.M., Jones, D.A., Harrison, K., Wulff, B.B., and Jones, J.D. (1997). Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the Cf-4/9 locus of tomato. Cell *91*, 821-832.

Pegg, G.F., and Young, D.H. (1982). Purification and characterization of chitinase enzymes from healthy and *Verticillium albo-atrum*-infected tomato plants, and from *V. albo-atrum*. Physiological Plant Pathology *21*, 389-398, IN317, 399-409.

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

Ronald, P.C., Albano, B., Tabien, R., Abenes, L., Wu, K.-s., McCouch, S., and Tanksley, S.D. (1992). Genetic and physical analysis of the rice bacterial blight disease resistance locus, Xa21. Molecular and General Genetics MGG *236*, 113-120.

Ross, A.F. (1961). Systemic acquired resistance induced by localized virus infections in plants. Virology *14*, 340-358.

Rudd, J.J., and Franklin-Tong, V.E. (2001). Unravelling response-specificity in Ca^{2+} signalling pathways in plant cells. New Phytologist *151*, 7-33.

Ryals, J.A., Neuenschwander, U.H., Willits, M.G., Molina, A., Steiner, H.Y., and Hunt, M.D. (1996). Systemic Acquired Resistance. Plant Cell *8*, 1809-1819.

Samuels, A.L., Giddings, T.H., Jr., and Staehelin, L.A. (1995). Cytokinesis in tobacco BY-2 and root tip cells: a new model of cell plate formation in higher plants. J Cell Biol *130*, 1345-1357.

Schmelzer, E. (2002). Cell polarization, a crucial process in fungal defence. Trends in Plant Science *7*, 411-415.

Schneider, D.S. (2002). Plant Immunity and Film Noir: What Gumshoe Detectives Can Teach Us about Plant-Pathogen Interactions. Cell *109*, 537-540.

Schuermann, D., Molinier, J., Fritsch, O., and Hohn, B. (2005). The dual nature of homologous recombination in plants. Trends in Genetics *21*, 172-181.

Schweigert, N., Zehnder, A.J., and Eggen, R.I. (2001). Chemical properties of catechols and their molecular modes of toxic action in cells, from microorganisms to mammals. Environ Microbiol *3*, 81-91.

Sessa, G., and Martin, G.B. (2000). Signal recognition and transduction mediated by the tomato Pto kinase: a paradigm of innate immunity in plants. Microbes Infect *2*, 1591-1597.

Shaw, J.G. (1999). Tobacco mosaic virus and the study of early events in virus infections. Philos Trans R Soc Lond B Biol Sci *354*, 603-611.

Shimomura, T., and Dijkstra, J. (1975). The occurrence of callose during the process of local lesion formation. European Journal of Plant Pathology *81*, 107-121.

Shirasu, K., and Schulze-Lefert, P. (2000). Regulators of cell death in disease resistance. Plant Mol Biol *44*, 371-385.

Shulaev, V., Leon, J., and Raskin, I. (1995). Is Salicylic Acid a Translocated Signal of Systemic Acquired Resistance in Tobacco? Plant Cell *7*, 1691-1701.

Sisler, H.D. (1986). Control of fungal diseases by compounds acting as antipenetrants. Crop Protection *5*, 306-313.

Smith, C.J. (1996). Accumulation of phytoalexins: defence mechanism and stimulus response system. New Phytologist *132*, 1-45.

Song, W.Y., Wang, G.L., Chen, L.L., Kim, H.S., Pi, L.Y., Holsten, T., Gardner, J., Wang, B., Zhai, W.X., Zhu, L.H., *et al.* (1995). A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. Science *270*, 1804-1806.

Stahl, E.A., and Bishop, J.G. (2000). Plant-pathogen arms races at the molecular level. Curr Opin Plant Biol *3*, 299-304.

Staskawicz, B.J., Ausubel, F.M., Baker, B.J., Ellis, J.G., and Jones, J.D. (1995). Molecular genetics of plant disease resistance. Science *268*, 661-667.

Sticher, L., Mauch-Mani, B., Métraux, and JP (1997). Systemic Acquired Resistance. Annual Review of Phytopathology *35*, 235-270.

Stintzi, A., Heitz, T., Prasad, V., Wiedemann-Merdinoglu, S., Kauffmann, S., Geoffroy, P., Legrand, M., and Fritig, B. (1993). Plant 'pathogenesis-related' proteins and their role in defense against pathogens. Biochimie *75*, 687-706.

Traut, T.W. (1994). The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites. European Journal of Biochemistry *222*, 9-19.

Tsuba, M., Katagiri, C., Takeuchi, Y., Takada, Y., and Yamaoka, N. (2002). Chemical factors of the leaf surface involved in the morphogenesis of *Blumeria graminis*. Physiological and Molecular Plant Pathology *60*, 51-57.

Van den Ackerveken, G.F., Van Kan, J.A., and De Wit, P.J. (1992). Molecular analysis of the avirulence gene *Avr9* of the fungal tomato pathogen *Cladosporium fulvum* fully supports the gene-for-gene hypothesis. Plant J *2*, 359-366.

Van Der Biezen, E.A., and Jones, J.D.G. (1998). Plant disease-resistance proteins and the gene-for-gene concept. Trends in Biochemical Sciences *23*, 454-456.

Van der Hoorn, R.A.L., De Wit, P.J.G.M., and Joosten, M.H.A.J. (2002). Balancing selection favors guarding resistance proteins. Trends in Plant Science *7*, 67-71.

Van Leeuwen, W., Hagendoorn, M., Ruttink, T., Van Poecke, R., Van Der Plas, L., and Van Der Krol, A. (2000). The use of the luciferase reporter system for in planta gene expression studies. Plant Molecular Biology Reporter *18*, 143-144.

van Loon, L. (1983). The induction of pathogenesis-related proteins by pathogens and specific chemicals. European Journal of Plant Pathology *89*, 265-273.

van Loon, L.C. (1975). Polyacrylamide disk electrophoresis of the soluble leaf proteins from *Nicotiana tabacum* var. "Samsun" and "Samsun NN" III. Influence of temperature and virus strain on changes induced by tobacco mosaic virus. Physiological Plant Pathology *6*, 289-300.

van Loon, L.C. (1985). Pathogenesis related proteins. Plant Mol Biol *4*, 111-116.

Van Loon, L.C., and Van Strien, E.A. (1999). The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. Physiological and Molecular Plant Pathology *55*, 85-97.

Vernooij, B., Friedrich, L., Morse, A., Reist, R., Kolditz-Jawhar, R., Ward, E., Uknes, S., Kessmann, H., and Ryals, J. (1994). Salicylic Acid Is Not the Translocated Signal Responsible for Inducing Systemic Acquired Resistance but Is Required in Signal Transduction. Plant Cell *6*, 959-965.

Vranova, E., Inze, D., and Van Breusegem, F. (2002). Signal transduction during oxidative stress. J Exp Bot *53*, 1227-1236.

Wang, G.L., Ruan, D.L., Song, W.Y., Sideris, S., Chen, L., Pi, L.Y., Zhang, S., Zhang, Z., Fauquet, C., Gaut, B.S., *et al.* (1998). *Xa21D* encodes a receptor-like molecule with a leucine-rich repeat domain that determines race-specific recognition and is subject to adaptive evolution. Plant Cell *10*, 765-779.

Warren, R.F., Henk, A., Mowery, P., Holub, E., and Innes, R.W. (1998). A mutation within the leucine-rich repeat domain of the Arabidopsis disease resistance gene *RPS5* partially suppresses multiple bacterial and downy mildew resistance genes. Plant Cell *10*, 1439-1452.

Wei, Y., Zhang, Z., Andersen, C.H., Schmelzer, E., Gregersen, P.L., Collinge, D.B., Smedegaard-Petersen, V., and Thordal-Christensen, H. (1998). An epidermis/papilla-specific oxalate oxidase-like protein in the defence response of barley attacked by the powdery mildew fungus. Plant Mol Biol *36*, 101-112.

White, R.F. (1979). Acetylsalicylic acid (aspirin) induces resistance to tobacco mosaic virus in tobacco. Virology *99*, 410-412.

Whitham, S., Dinesh-Kumar, S.P., Choi, D., Hehl, R., Corr, C., and Baker, B. (1994). The product of the tobacco mosaic virus resistance gene *N*: similarity to toll and the interleukin-1 receptor. Cell *78*, 1101-1115.

Whitham, S., McCormick, S., and Baker, B. (1996). The *N* gene of tobacco confers resistance to tobacco mosaic virus in transgenic tomato. Proc Natl Acad Sci U S A *93*, 8776-8781.

Williams, C.E., Wang, B., Holsten, T.E., Scambray, J., Silva, F., and Ronald, P.C. (1996). Markers for selection of the rice *Xa21* disease resistance gene. TAG Theoretical and Applied Genetics *93*, 1119-1122.

Wu, X., and Shaw, J.G. (1997). Evidence That a Viral Replicase Protein Is Involved in the Disassembly of Tobacco Mosaic Virus Particlesin Vivo. Virology *239*, 426-434.

Yazaki, K., Sasaki, K., and Tsurumaru, Y. Prenylation of aromatic compounds, a key diversification of plant secondary metabolites. Phytochemistry *70*, 1739-1745.

3.0 Tobacco mosaic virus infection results in an increase in recombination frequency and resistance to viral, bacterial and fungal pathogens in the progeny of infected tobacco plants

3.1 Abstract

Infection of *Nicotiana tabacum* plants with tobacco mosaic virus (*TMV*) leads to an increase in homologous recombination frequency (HRF). The progeny of infected plants also had an increased rate of rearrangements at resistance gene-like loci. Here, we report that tobacco plants infected with *TMV* exhibited an increase in HRF in two consecutive generations. Analysis of global genome methylation showed the hypermethylated genome in both generations of plants, whereas analysis of methylation via 5-MeC antibodies showed both hypomethylation and hypermethylation. Analysis of the response of the progeny of infected plants to *TMV*, *Pseudomonas syringae*, or *Phytophtora nicotianae* revealed a significant delay in symptom development. Infection of these plants with *TMV* or *Pseudomonas syringae* showed higher levels of induction of *PR1* gene expression and higher level of callose deposition. Our experiments suggest that viral infection triggers specific changes in the progeny that promote higher HRF and higher resistance to stress.

3.2 Introduction

Continuous exposure to stress leads to the evolutionary selection of adaptive traits beneficial in a particular environment. It is believed that such selection of the individuals fittest of a

population of plants grown under certain environmental conditions requires a long time. It is known, however, that plants also possess the ability to acclimate on a much shorter time scale. A modification of homeostasis, also termed acclimatization, is a well-documented process that is used for adjusting metabolism to a new environment (Lichtenthaler, 1998; Mullineaux and Emlyn-Jones, 2005).

Pathogens represent one of a variety of stresses that plants are constantly exposed to. In nature, the evolution of plant resistance to a particular pathogen, such as a virus, bacterium or fungus, has been the result of constant interactions with said pathogen (Friedman and Baker, 2007; McHale et al., 2006) leading to a constant plant-pathogen arms race (Ingle et al., 2006).

Plants are able to tolerate or resist pathogens in a variety of ways, which could be broadly attributed to mechanisms of innate immunity and actual gene-for-gene-based resistance. The latter one depends on direct or indirect recognition of pathogen avirulence gene products by plant resistance gene products (Durrant and Dong, 2004; Whitham et al., 1994). Pathogen recognition in this so-called incompatible interaction triggers complex events, including the local hypersensitive response that manifests itself as boost of radical production, activation of the salicylic acid-dependent pathway and necrotic lesions, which working together restrict pathogen spread. It also results in a plant-wide systemic acquired resistance response that provides protection or tolerance to future pathogen attacks (Durrant and Dong, 2004; Park et al., 2007; Vlot et al., 2008).

The interaction between a plant and a pathogen in the case when a functional pathogen resistance gene is absent (compatible interaction) is less understood phenomena. How do plants which lack a resistance gene respond to infection? We have previously reported that the compatible interaction between tobacco mosaic virus (*TMV*) and *Nicotiana tabacum*

(tobacco) cultivar SR1 plants that lack the *TMV* resistance *N*-gene results in the production of a systemic signal. The signal leads to the increase in the somatic homologous recombination frequency (HRF) (Kovalchuk et al., 2003a). Based on these observations, we initially hypothesized that these genomic changes could be inherited. Indeed, we found that the progeny of infected SR1 tobacco plants exhibited a higher frequency of restriction fragment length polymorphisms at the loci with more than 60% similarity to the LRR region of the *N*-gene (Boyko et al., 2007).

Although several reports showed an increase in genome instability in plants exposed to pathogens and pathogen elicitors (Boyko et al., 2007; Kovalchuk et al., 2003a; Lucht et al., 2002; Molinier et al., 2006), many questions still remained unanswered. What is the mechanism of occurrence of a pathogen-induced systemic increase in HRF? What is the mechanism of inheritance of high HRF? Is an increased HRF maintained throughout generations? What other changes occur in the progeny of infected plants?

Here, we attempted to answer the abovementioned questions by analyzing two consecutive progenies of *TMV*-infected tobacco cultivar SR1 plants. Both progenies of infected plants showed higher somatic HRF, higher resistance to *TMV* infection and tolerance to methyl methane sulfonate, an increase in callose deposition as well as a higher steady state PR1 RNA level. Analysis of methylation patterns revealed global genome hypermethylation in both progenies paralleled by hypomethylation in euchromatic areas.

3.3 Materials and Methods

3.3.1 Generation and use of transgenic tobacco plants carrying a homologous recombination substrate

The generation of transgenic tobacco cultivar SR1 plants line #A carrying in the genome a single copy of the luciferase (LUC) recombination substrate was described previously (Kovalchuk et al., 2003a).

3.3.2 Infection of SR1 plants with *TMV* and generation of infected plant's progeny

In these experiments, single leaves of 20 *Nicotiana tabacum* cultivar SR1 plants containing the recombination substrate (line #A) were infected with 200 ng (100 ng/μl in phosphate buffer) of *TMV* U1 strain, and twenty plants were mock-inoculated with phosphate buffer only. Carborundum was used as an abrasive. The experiments were done in two independent rounds, with 10 plants from each group in each round. Infected and mock-inoculated plants were kept at 32°C. At 24h after infection, upper non-infected leaves (as checked by PCR, see (Kovalchuk et al., 2003a)) were grafted to non-infected six-week-old plants (Kovalchuk et al., 2003a). The apical meristem of these grafted plants was removed to promote the development of lateral buds. Seeds from these plants were collected individually and named "the progeny of infected" (20 different lines, PI_1 #1, PI_1 #2 etc.) or "PI_1" and "the progeny of control" (20 lines, PC_1 #1, PC_1 #2) or "PC_1" (Figure 3.1). All these progenies had a comparable number and weight of seeds produced (data are not shown). Plants germinated from these seeds never developed any viral infection, unless infected.

Similarly, the progeny of infected plants never showed any presence of viral RNA (data not shown).

3.3.3 Analysis of plant resistance to *TMV* infection. Plant inoculation, monitoring infection symptoms and analysis of the viral titer

For analyzing resistance to *TMV*, single leaves of four hundred five-week-old PI_1 plants (twenty plants per each of 20 lines) and one hundred PC_1 plants (10 plants from 10 different lines) were infected with 200 ng (100 ng/μl) of *TMV* U1 particles. Infection was carried out by rub inoculation with Carborundum powder as abrasive. The experiments were done in two rounds, 10 plants per each PI group and 5 plants per each PC group for each experimental round. After infection, plants were grown at 16/8 h day/night at 32°C. PC plants were propagated at 32°C without infection. Symptom appearance (light coloration between veins in young leaves, followed up by leaf darkening and curling) was monitored daily. For the analysis of the viral titer, tissue samples of non-inoculated leaves from each infected plant (20 plants per each line), regardless of whether plants showed symptoms or not, were taken at 6, 9 and 12 days post infection (dpi). Viral RNA was extracted, and virus concentration was analyzed as previously published (Asurmendi et al., 2007). In brief, the plant tissue was homogenized and aliquoted to 500 mg per sample. 400 μl of 0.5M phosphate-extraction pH 7.0 buffer was added to 500 mg tissue homogenate (4.1 g Na_2HPO4, 2.5 g NaH_2PO_4, 100 μl β-mercaptoethanol per 100 ml). Samples were purified using equal volumes (400 μl) of chroloform and 1-butanol. Viral particles were precipitated using separate volumes of 50 μl of 40% polyethyleneglycol 6000 and a 10% NaCl solution.

The pellet was resuspended in 20 µl of a 1:50 dilution of phosphate buffer that lacked β-mercaptoethanol. The viral titer was estimated by measuring OD_{260} using a spectrophotometer. For *TMV*, the extinction coefficient is assumed to be three. The titer is expressed in µg of *TMV* per 10 mg fresh weight.

3.3.4 Black Shank Disease Assays

Phytophthora nicotianae Breda de Haan (synonym Phytophthora parasitica Dastur) assays were performed with an aggressive isolate (LA0921) on 6 week old plants of line PC_1, PI_1 #9 and PI_1 #16 grown in 1:1:1 soil/peat moss/vermiculite mixture (sterilized by autoclaving) in 4 inch pots. About eighty plants per each experimental group were used. Plants were watered, allowed to drain, and inoculated by applying 10 ml of a zoospore and sporangia suspension containing 1000 propagules per ml to the soil. Inoculated plants were incubated in plexiglass trays at 22°C and 100 % relative humidity. A wilt index was calculated for the assay at 5 days post inoculation as follows: 0 = no symptoms, 1 = some signs of wilting and stunting with reduced turgidity, 2 = advanced wilting and stunting but no chlorosis or necrosis, 3 = advanced wilting with stunting and chlorosis but no necrosis, 4 = severe wilting with necrosis, 5 = advanced necrosis with plants near death or dead. Average incidence was calculated by relating the number of plants with signs of wilting (any scale from 1 to 5) to the total number of plants used. All screening was blind with plants assayed in a random design.

3.3.5 Analysis of plant resistance to *Pseudomonas syringae*

Pseodomonas syringae pv *tomato* (Pst) DC3000 avirulent (Pst-avr) strain was used for the experiment. Bacteria were cultured at 28°C on King's medium (King et al., 1954) containing 25 mg/L rifampicin and 50 mg/L kanamycin (Pst-avr). Bacterial growth in the leaves of tobacco plants from PC_1, PI_1 #9 and PI_1 #16 lines was determined by syringe infiltrating the bacterial suspension *in vivo* into leaves as described (Whalen et al., 1991). Briefly, the fresh overnight bacterial culture was centrifuged at 3200 RCF for 5 min and washed twice with 10 mM $MgCl_2$. The OD_{600} of the suspension was determined by a spectrometry and further diluted to OD_{600}=0.05 with 10 mM $MgCl_2$.

Four week old plants were incubated at 30°C for 24 hours prior to infection. Plants were infiltrated with above mentioned *Pseudomonas* culture and the infiltrated area was marked. Tissues were harvested at 0, 24, 48 and 72 hours post infection. At each time point five samples were taken from five different plants from each line. Each of these samples consisted of three leaf discs (3 mm diameter) from independent infiltrations from a single plant. The samples were macerated in 10 mM $MgCl_2$ and the content was diluted 3000 and 5000 times. From each sample, two plates containing MS media were inoculated using 50 µl extract. Number of colonies formed was counted after two days. The experiment was repeated twice. Each experimental datum point is an average of two experiments, each consisting of ten individual data points (two plates from each of five samples).

3.3.6 Generating the second progeny of infected plants

PI_1 plants from lines 9 and 19 were used to produce the second generation of plants exposed to virus. At the age of five weeks, twenty PI_1 plants from line9 and twenty PI_1 plants from line19 were either infected with 200 ng (100 ng/μl) of *TMV* U1 particles or mock-inoculated. Next, ten plants (those that showed an increase in HRF and looked healthy) from each group were propagated to seeds. For propagating PC_1 plants, we used lines 1 and 3. Similarly, ten plants from each PC_1 line were propagated to generate PC_2 plants.

3.3.7 Visualization of luciferase activity

Recombination events in transgenic plants were visualized on living plants with a charged couple device (CCD) camera, 2 hours after the cleavage substrate, luciferine, was applied (Supplementary Figure 3.1b).

3.3.8 Exposure to methyl methane sulfonate (MMS)

Progenies of tobacco plants were exposed to various concentrations of methyl methane sulfonate by growing plants on MS media supplemented with this chemical. First, in the pilot experiment, it was shown that a concentration of 120 particulate per million (ppm) led to a significant change in the plant phenotypic appearance (data not shown). That is why the same concentration was used to analyze resistance of PC_1, PI_1, PI_2 and PI_C1 plants.

3.3.9 Exposure to NaCl and rose Bengal

Plants from individual PI_1 and PC_1 lines were germinated and grown on sterile MS media. At the age of one week, they were moved to MS media supplemented with various amounts of NaCl or rose Bengal. Four NaCl concentrations (0, 50, 75 and 100 mM) were used, whereas the concentrations of rose Bengal used were 0, 4, 10 and 20 µM. Approximately 100 plants per each plate and each line were sampled. Recombination frequency was tested at three weeks post germination.

3.3.10 Analysis of the phenolic content

For the analysis of total polyphenolic compounds, fresh tissues of PI_1 plants from line 9 and 19 and PC_1 plants from line 1 and 3 were ground in liquid nitrogen. 500 mg of tissue powder of each sample was weighed and used for extraction using 500 µl PBS. After centrifugation of the extract, 1.0 ml of Folin-Ciocalteu reagent (Sigma-Aldrich) and 0.8 ml of 7.5% sodium carbonate were added to 20 µl of the extract in a tube. The solution was mixed and incubated for 30 minutes. Absorbance was measured at 765 nm wavelength using a spectrophotometer. Gallic acid was used as a standard, and values were expressed as relative arbitrary units.

3.3.11 Analysis of callose deposition

For the analysis of callose deposition, leaves were harvested at 0, 12, 24, 36 and 48 hours after *Pseudomonas* infiltration. 15 µm transverse sections were obtained from leaves using a

cryomicrotome. After post fixation and washings, the tissue was blocked in 5% goat serum. The primary anti-callose antibody (Biosupplies, Australia) was diluted 1:250 in blocking solution and incubated at 4°C overnight. After w ashings, anti-mouse Alexa 488 (Invitrogen, USA) was diluted 1:500 in blocking buffer and incubated for 3 hours at room temperature. DNA counterstaining was done using 4', 6-diamidino-2-phenylindole (DAPI; in blue). After washing, slides were mounted and analyzed using a confocal microscope (Nikon, Japan). Pictures were quantified using ImageJ. Each time point is the average of five different readings stemming from sections of 3 different leaves.

3.3.12 Analysis of global genome methylation using a cytosine extension assay and using the anti-5-methyl cytosine antibody

Global genome methylation was analyzed as previously described (Boyko et al., 2007).

For immuno-histochemical staining, plants were fixed in 4% paraformaldehyde (PFA), and leaf cross sections (15 µm thick) were obtained by cryosectioning. The slides were post-fixed in 4% PFA and washed in 1X phosphate buffered saline (PBS). Blocking was carried out in a 5% bovine serum albumin (BSA) solution in PBS, and the primary anti-5-methyl cytosine antibody (Eurogentec, USA) was applied for overnight. After subsequent washes, DAPI was applied as a counterstain and Laser scanning electron microscopy was carried out. The intensity of staining at the center of the nucleus (Figure 3.6c, d) of PI_1 and PC_1 plants was analyzed by comparing the average ratio of red-to-blue in 10 independent cells from each of 3 independent leaves taken from 3 independent plants.

3.3.13 Analysis of PR1 expression in tobacco infected with *TMV* or *Pseudomonas syringae*

Single leaves of one-month-old tobacco plantlets from PC_1 and PI_1 groups were rub-inoculated with either 200 ng *TMV* or *P. syringae* (OD600 – 0.05). The control group was mock-treated. Tissues were collected at 24 and 48 hours post infection. Northern blot analysis for *PR1* RNA was conducted using tissues from tobacco plants infected with either *TMV* or *P. syringae,* as previously described (Kwon and An, 2001). The PR1 probe was prepared using the polymerase chain reaction (PCR) DIG probe synthesis kit (Roche, Germany) with the following primers: a forward primer - 5' TAGTCATGGGATTTGTTCTCT 3', a reverse primer – 5' CTAGACCATCAACACATGATT 3'. Fragment intensity was quantified using "Image J" (NIH, www.rsb.info.nih.gov/ij).

3.3.14 Statistical treatment of data

In all cases, average and standard deviations in three independent experiments were calculated. The statistical significance of experiments was confirmed by performing either a Student's t-test (two-tailed paired or non-paired) or Single factor analysis of variance (ANOVA) or Two-Factor ANOVA tests. Statistical analyses were performed using the MS Excel software and Microcal Origin 6.0.

3.4 Results

3.4.1 Generating the progeny of virus-infected plants

For the current study, we used transgenic *Nicotiana tabacum* cultivar SR1 containing two non-functional truncated copies of the luciferase (*LUC*) gene to serve as a substrate for homologous recombination (Supplementary Figure 3.1a). By using a special luciferase camera, recombination events can be visualized as bright spots (Supplementary Figure 3.1b). We have previously shown that infection of these transgenic tobacco plants with *TMV* increased HRF in infected leaves as well as in distant non-infected tissues (Kovalchuk et al., 2003a). Mock inoculation also led to an increase in HRF but to a lesser degree (Kovalchuk et al., 2003a). For the experiments presented in this paper, we used seeds collected from 20 independent infected and 20 independent mock-inoculated plants. These plants were named PI_1 for "progeny of infected plants, generation one", and PC_1 for "progeny of control plants, generation one" (Figure 3.1). Mock-inoculated plants were used as control.

3.4.2 PI_1 plants showed higher spontaneous and induced somatic HRF as compared to PC_1 plants

It is possible that a previously observed (Boyko et al., 2007) increase in the frequency of rearrangements at *N*-gene-like *R* gene loci reflects similar changes in other genomic loci. Therefore, we attempted to analyze genome stability in the progeny plants by measuring HRF at a transgene (LU-UC) locus. We analyzed the average number of somatic recombination events per plant in PI_1 and PC_1 plants and found that HRF of 3.3±0.4 in PI_1 plants was significantly (*t*-test, $P<0.05$) different from that of 1.98±0.5 in PC_1 plants

(Figure 3.2). Since recombination is a general DNA repair mechanism, it was possible that these changes would also be paralleled by changes in response to DNA damaging agents and tolerance to stress.

To analyze whether stress exposure affects HRF in PI_1 plants differently as compared to PC_1 plants, we exposed tobacco plants to 50, 75 and 100 mM NaCl or 4, 10 and 20 μM rose Bengal, a photosensitizing dye that produces a single oxygen molecule (Ledford et al., 2007). Both chemicals induce radical production and were shown to increase HRF in plants (Filkowski et al., 2004; Puchta et al., 1995). Our experiment showed that PI_1 plants responded to stress with a higher increase in HRF ($P<0.05$, Two-Factor ANOVA) (Figure 3.3a, b). HRF decreased in plants exposed to higher doses of NaCl (100 mM) or rose Bengal (10 and 20 μM); it was possible that a high concentration of these compounds resulted in massive death of cells in which recombination events could have been observed.

3.4.3 The progeny of infected plants exhibit delay in the onset of the symptoms of viral infection

To analyze how the progeny of infected plants respond to *TMV*, we infected five-week old PI_1 and PC_1 plants with 200 ng of *TMV* U1 particles. At 5 days post infection (dpi), over half of the PC_1 plants had the first symptoms of infection (light coloration between veins in young leaves). Further analysis of the symptoms showed that PI_1 plants had a significantly lower percentage of plants with these symptoms at 5, 6, 7 and 9 dpi ($P<0.01$ in all cases, single factor ANOVA) (Figure 3.4a). At 12 and 15 dpi, the number of plants with the symptoms was statistically similar in PI_1 and PC_1 plants ($P>0.1$, single factor ANOVA)

(Figure 3.4a), although several PI_1 lines had plants that showed no symptoms up to 15 dpi. All PC_1 plants exhibited the symptoms at 7, 9, 12 and 15 dpi.

The appearance of symptoms could probably correspond with the presence of virus in systemic tissues. Based on the symptom appearance data, it was decided to check a virus titer at 6, 9 and 12 dpi. The analysis showed that at 6 dpi the average viral titer in PI_1 plants was significantly lower ($P<0.01$) than that in PC_1 plants (Figure 3.4b). A similar trend was observed at 9 dpi, although the difference was not significant ($P>0.1$) (Figure 3.4b). In most of the PI_1 lines, virus concentrations were higher at 12 dpi compared to the PC_1 lines. An increase, peak and decline in the viral titer is a known phenomenon observed before in tomato plants infected with *TMV*; virus concentrations peaked at 7 dpi and declined steadily on later days (Balogun et al., 2002). Thus, these experiments showed that PI_1 plants were able to significantly delay virus replication and the progression of symptoms.

3.4.4 The progeny of infected plants have lower bacterial titer upon infection with *Pseudomonas syringae*

Next, we attempted to analyze the response of PC_1 and PI_1 plants to *Pseudomonas syringae*. Infection with *P. syringae* pv *tomato* (Pst) DC3000 avirulent strain showed that PI_1#9 and PI_1#16 had lower bacterial count. PI_1#9 plants had significantly lower bacterial counts at 24 and 72 hpi ($P<0.05$ for both) (Figure 3.5a). PI_1#16 plants had significantly lower bacterial counts at 48 hpi (Figure 3.5a).

3.4.5 The progeny of infected plants have higher resistance to the infection with *Phytophthora nicotianae*

To analyze whether the PI_1 plants also have different response to fungal pathogen, we infected PC_1 plants and plants from two PI_1 lines, PI_1#9 and PI_1#16 with *Phytophthora nicotianae* Breda de Haan. Analysis of symptom appearance showed dramatic differences between PC_1 plants and either of two PI_1 plants (Figure 3.5b). To analyze the infection progression, a wilt index was calculated (0 to 5, with 5 being most severe; see Materials and methods for details). The assay showed that PC_1 plants had an average severity of disease of 3.4, whereas PI_1#9 and PI_1#16 had only 0.3 and 0.5, respectively (Table 3.1). Also, PI_1 plants had much lower number of plants with symptoms as compared to PC_1 plants; whereas 87% of PC_1 plants showed infection symptoms, only 26% and 33% of PI_1#9 and PI_1#16 plants had any infection symptoms (Table 3.1). Table 3.1 shows the number of plants with wilt symptoms of different severity, where 0 – is no symptoms, and 5 – is the most severe symptoms. “Average severity” shows the average symptom appearance for the population of infected PC_1 or PI_1 plants. “Incidence” shows the percent of infected plants. The difference between PC_1 and either PI_1 line is statistically significant ($P<0.01$ in all cases).

3.4.6 The second generation of plants exposed to virus also showed higher spontaneous HRF

Previously, Molinier et al. (2006) found that plants exposed to UVC maintain a high frequency of recombination for several generations. Thus, the question arises as to whether

PI_1 plants will exhibit a different HRF if they are propagated for further generations with or without exposure to virus. To obtain the second generation, PI_1 plants of lines 9 and 19 were used. Plants of both lines showed significant changes in methylation patterns and in the frequency of rearrangements at the *N*-gene-like resistance gene loci (Boyko et al., 2007). To obtain the second generation of control plants, PC_1 plants of lines 1 and 3 were used (Figure 3.2). PI_1 plants were either mock-inoculated or infected with *TMV* and PC_1 plants were mock-inoculated only.

The analysis of spontaneous non-induced somatic recombination in the progeny of infected PI_1 plants (named PI_2) revealed a similar high HRF as compared to PI_1 plants (Figure 3.4c). The progeny of PI_1 propagated under control conditions (named PI_1_C1) also exhibited a high HRF, although lower than in PI_1 and PI_2. Recombination in PC_1 and PC_2 plants was comparably low. Therefore, it was shown that the propagation of plants in the presence of virus maintains a higher level of HRF, whereas the propagation without viral presence decreases the level of HRF albeit not to the level exhibited by control plants (Figure 3.4c).

3.4.7 PI_2 plants also showed lower viral titers

To analyze whether the second generation of exposed plants still maintains higher tolerance to viral infection, we infected PI_1, PC_1, PI_2, PI_1_C1 and PC_2 plants (twenty plants per treatment). The analysis of the viral titer showed that all PI_1, PI_2 and PI_1C1 plants had lower titers at 6 and 9 dpi, as compared to either PC_1 or PC_2 plants ($P<0.01$). Significant differences were observed in all lines at 6 dpi and in all but PI_1 – at 9 dpi. The

virus titers of PI_1, PI_2 and PI_1_C1 plants did not differ from each other (P>0.1) (Figure 3.4d). PC_1 and PC_2 plants also had a similar viral titer (P>0.1).

3.4.8 PI_1, PI_2 and PI_1_C1 plants were more tolerant to methyl methane sulfonate

Next, we analyzed whether the progeny of infected plants exhibit cross-tolerance to other stresses. PC_1, PI_1, PI_2 and PI_1_C1 plants were exposed to a DNA damaging agent, methyl methane sulfonate (MMS). A pilot experiment established the concentration of 120 ppm of MMS to be ideal for the observation of potential differences in stress tolerance (Figure 3.6a). We found that despite the fact that PC_1 plants had longer roots under normal conditions than PI_1 plants, PC_1 plants grown in the presence of MMS had significantly shorter roots as compared to PI_1, PI_2 and PI_1_C1 plants (P<0.05; Figure 3.6b). The analysis of plant biomass showed similar differences (data not shown).

3.4.9 The progeny of infected plants exhibited global genome hypermethylation

We assumed that epigenetic mechanisms might be responsible for an increase in stress tolerance. Heritable DNA methylation changes in the genome could allow differential gene expression in the progeny, which in its turn may provide more protection for plants (Kovalchuk et al., 2003b). Analysis of global genome methylation by a cytosine extension assay showed that all PI_1, PI_2 and PI_C1 plants had significant genome hypermethylation compared to PC_1 plants (P<0.05; Figure 3.7a) (data for PC_2 was similar).

In order to analyze methylation '*in situ*', we performed immunohistochemical staining of cross-sectioned tobacco leaves using an anti-5-Me-Cytosine antibody (Figure 3.7b-d). PI_1

plants showed higher methylation at the periphery (the brighter red intensity at the nuclear periphery) and lower methylation in the center of the nucleus (a low level of red coloring in the center of the nucleus) (Figure 3.7c-e). The analysis of the ratio of red (5-Me-C) to blue (DAPI) confirmed hypomethylation in the center of the nucleus (Figure 3.7c-e). A higher intensity of red coloring observed in PI_1 plants as compared to PC_1 plants could suggest hypermethylation in the nuclear envelope; however, it was difficult to measure it quantitatively (Figure 3.7c, d).

The abovementioned experiments showed that PI_1 and PC_1 plants differ not only in global genome methylation but possibly in the distribution of methylation across the genome.

3.4.10 Infection with *TMV* and *P. syringae* results in a stronger induction of the PR1 gene in PI_1 plants

Since we observed higher resistance of PI plants to viral infection, we hypothesized that these plants while exposed to a pathogen would also have higher expression of pathogenesis-related genes. Expression of the pathogenesis-related gene 1 (*PR1*) is an indicator of the attempted resistance response (Mitsuhara et al., 2008). We exposed plants of two PI_1 lines, PI_1#9 and PI_1#16 and PC_1 plants to *TMV* and *P. syringae* and analyzed steady-state RNA levels of the *PR1* gene at 0, 24 and 48 hours post infection (hpi). Northern blot analysis indeed showed that PI_1 plants responded with a higher increase in steady-state levels of PR1 RNA at 24 and 48 hpi (Figure 3.8a, b, c).

3.4.11 PI_1 plants showed a higher level of callose deposition

One of the mechanisms of pathogen resistance is the accumulation of callose in plasmodesmata. Callose is an amorphous β-1, 3-D-glucan found in numerous locations in higher plants and induced by abiotic and biotic stresses (Flors et al., 2008). Rapid deposition of callose at the point of attempted penetration by fungal pathogens has been observed (Ton and Mauch-Mani, 2004). To analyze callose deposition at plasmodesmata, we performed immunohistochemical staining with anti-callose antibodies (Figure 3.9a, b). We analyzed callose deposition at 12, 24, 36 and 48 hours post infection with *Pseudomonas syringae*. We found that PI_1 plants from line 9 and 16 had significantly more (P<0.01) callose depositions as compared to PC_1 plants (Figure 3.9a, b).

3.5 Discussion

In this paper, we show that the progeny of infected plants exhibit a higher level of spontaneous and stress-induced somatic HRF as well as higher resistance to viral, bacterial and fungal infection and tolerance to MMS exposure. We demonstrate that PI plants have higher induction of the PR1 gene and a higher level of callose deposition.

3.5.1 Local infection of SR1 tobacco plants with *TMV* leads to heritable changes in genome stability

Here, we found that the progeny of infected SR1 plants (PI_1) had higher spontaneous somatic HRF. The progeny of PI_1 plants also exhibited a higher level in HRF, although it was less pronounced when PI_1 plants were propagated under normal conditions. Previously, Molinier et al. (2006) reported five consecutive generations of UVC-exposed

plants that showed a significant increase in HRF. Unfortunately, the authors did not obtain similar results with flagellin-treated plants; they found an increase in HRF in the first progeny of flagellin-treated plants but did not profile further generations (Molinier et al., 2006).

A recent report by Pecinka et al. (2009) suggests that transgenerational changes in HRF in response to stress are rather stochastic and do not occur in response to all stresses. The authors showed that four out of ten stresses tested resulted in changes in HRF in the progeny (Pecinka et al., 2009). Unfortunately the authors did not test exposure to biotic stress.

3.5.2 The progeny of infected plants exhibit higher resistance to *TMV* infection and MMS

The fact that the progeny of infected plants have higher resistance to viral infection suggests that plants possess the mechanisms that allow certain adaptations to pathogen pressure. Single generation adaptations/acclimations have been reported in the past, although they have primarily been described for temperature stress.

It has been reported that warm temperatures during the development of maternal and embryonic tissues in *Picea abies* are associated with better performance of offspring grown at the same elevated temperature (Johnsen et al., 2005). Similar findings have recently been reported for *Arabidopsis* (Blodner et al., 2007). There has been demonstrated a memory of temperature treatments on a subsequent bud set and height growth in *Picea abies* during somatic embryogenesis, thereby eliminating the influence of maternal effects *per se* (Kvaalen and Johnsen, 2008).

Cross-protection against MMS in the progeny of infected tobacco plants also deserves some attention. It is known that exposure to abiotic and biotic stresses often trigger a response of a set of commonly regulated genes (Glombitza et al., 2004). It is common knowledge that plants with certain mycorrhizal associations are less sensitive to cadmium stress than non-mycorrhizal plants (Schutzendubel and Polle, 2002). Several recent reports have showed that plants exposed to UVC acquire protection against pathogen infection (Kunz et al., 2006; Kunz et al., 2008). Moreover, one of the reports showed that plants treated with salicylic acid exhibited higher tolerance to UVB and UVC (Mahdavian et al., 2007).

3.5.3 Higher resistance to stress could be a result of enhanced innate immunity

Higher resistance to pathogen stress could be a result of various events. Innate immunity is one of the types of broad-spectrum protection against pathogen infection. In this case, plant resistance is associated with plants' ability to slow down pathogen progression and/or to withstand a higher level of damage by a pathogen (de Wit, 2007). This is in part due to the function of secondary metabolites that could serve as signaling molecules, could be toxic to a pathogen, and could physically impair pathogen progression (Asselbergh et al., 2008; Jansen et al., 2001).

Indeed, we analyzed the total amount of phenolics and found much higher levels in non-stressed PI plants as compared to PC plants (Supplementary Figure 3.2). Many phenolic compounds were shown to be effective protectants against stress (Jansen et al., 2001; Kovacik and Klejdus, 2008). The total phenolic content of injured poplar and maple tree seedlings and non-injured seedlings cohabiting with them was significantly higher than that in sequestered controls (Baldwin and Schultz, 1983). Since we observed an elevated level of

phenolics in PI_1 plants, we assumed that these plants would also have changes in other metabolites. Using NMR spectroscopy, we analyzed metabolites in PI_1 and PC_1 plants and found that PI_1 plants had an increase of about 80% in the total amount of metabolites, as compared to PC_1 plants (Supplementary Figure 3.3).

In our experiments, the progeny of infected plants also showed a high level of callose deposition. This plant polysaccharide is known to accumulate in response to exposure to various abiotic and biotic stresses (Jones et al., 2006). In fact, *TMV* movement through plasmodesmata was substantially delayed in plants with a high level of callose (Iglesias and Meins, 2000).

Induced PR1 gene expression is one of the hallmarks of a hypersensitive response and SAR (Chamnongpol et al., 1998). High expression of *PR1* genes in rice is in good correlation with higher tolerance to *Xanthomonas oryzae* (Ponciano et al., 2006). Plants that had the spontaneously activated PR1 gene displayed higher tolerance to *Pseudomonas syringae* pv *tomato* DC3000 (Morita-Yamamuro et al., 2005). In our experiments we also found the PI_1 plants to have lower bacterial count upon infection with this pathogen.

High resistance of PI_1 plants to *P. nicotianae* deserves separate attention. Although the resistance of PI plants to viral (*TMV*; $p>0.1$ to $p<0.01$) and bacterial (*P. syringae*; $p<0.05$) was rather marginal, the response to fungal pathogen *P. nicotianae* was quite drastic ($p<0.01$). Whereas 87% of all infected PC_1 plants developed infection symptoms, only about a quarter of PI_1 plants did (Table 3.1). One of the possible explanations of such effect could be higher spontaneous and induced level of *PR1* gene. It was previously shown that tobacco plants overexpressing *PR1* gene exhibited higher tolerance to aforementioned fungal pathogen (Alexander et al., 1993). Curiously, another report by Shin et al. (2002) showed that transgenic pepper plants overexpressing tobacco stress induced gene 1 (Tsi1),

that is involved in regulation of *PR1* gene, exhibited higher resistance to viral, bacterial and fungal pathogens (Shin et al., 2002). The fact that PI plants also acquired higher resistance to viral, bacterial and fungal pathogens may suggest that plants with higher resistance can be obtained by local transient infection with *TMV*; only single leaves of parental generation plants were infected with *TMV* and infected leaves were removed 24 hpi. It remains to be shown whether other pathogens can trigger the same response and whether PI plants acquire similar resistance to other pathogens.

3.5.4 Possible mechanisms of transgenerational changes in the progeny of stressed plants

Our previous work showed a systemic increase in HRF upon local infection (Kovalchuk et al., 2003a). We hypothesized that upon local infection signalling molecules arrive at non-infected tissue faster than a virus. Love et al. (2005) showed that infection of *Arabidopsis* plants by the cauliflower mosaic virus (CaMV), a compatible pathogen of *Arabidopsis*, led to the activation of multiple defence signalling mechanisms (Love et al., 2005). This suggested the existence of signalling even in response to a compatible interaction in plants. It is possible that this signal reaches the gametes and establishes different methylation patterns, which possibly leads to differential gene expression, including expression of PR1, higher HRF, an increase in the level of phenolics, metabolites and thus in higher stress tolerance. Several recent publications suggest the existence of certain transgenerational memory in the response of plants to stress (Boyko et al.; Molinier et al., 2006; Pecinka et al., 2009). It is suggested that transgenerational response depends on transgenic plants used, the type, the intensity of stress, the mode of stress application and even conditions that plants are grown

at. More recent work suggests the involvement of siRNA biogenesis pathways in establishment of transgenerational response to abiotic stress (Boyko et al.). It remains to be shown whether this is applied to biotic stress.

Collectively, our findings suggest the existence of adaptive inheritance to *TMV* infection, possibly of an epigenetic nature. This indicates that biotic stress of viral origin can induce genetic- and epigenetic-related molecular changes, and suggests the possibility of multigenerational adaptive epigenetic phenomena in the progeny of infected plants.

3.6 References:

Alexander, D., Goodman, R.M., Gut-Rella, M., Glascock, C., Weymann, K., Friedrich, L., Maddox, D., Ahl-Goy, P., Luntz, T., Ward, E., *et al.* (1993). Increased tolerance to two oomycete pathogens in transgenic tobacco expressing pathogenesis-related protein 1a. Proc Natl Acad Sci U S A *90*, 7327-7331.

Asselbergh, B., De Vleesschauwer, D., and Hofte, M. (2008). Global switches and fine-tuning-ABA modulates plant pathogen defense. Mol Plant Microbe Interact *21*, 709-719.

Asurmendi, S., Berg, R.H., Smith, T.J., Bendahmane, M., and Beachy, R.N. (2007). Aggregation of *TMV* CP plays a role in CP functions and in coat-protein-mediated resistance. Virology *366*, 98-106.

Baldwin, I.T., and Schultz, J.C. (1983). Rapid Changes in Tree Leaf Chemistry Induced by Damage: Evidence for Communication Between Plants. Science *221*, 277-279.

Balogun, M.A., Ramsay, M.E., Hesketh, L.M., Andrews, N., Osborne, K.P., Gay, N.J., and Morgan-Capner, P. (2002). The prevalence of hepatitis C in England and Wales. J Infect *45*, 219-226.

Blodner, C., Goebel, C., Feussner, I., Gatz, C., and Polle, A. (2007). Warm and cold parental reproductive environments affect seed properties, fitness, and cold responsiveness in Arabidopsis thaliana progenies. Plant Cell Environ *30*, 165-175.

Boyko, A., Blevins, T., Yao, Y., Golubov, A., Bilichak, A., Ilnytskyy, Y., Hollander, J., Meins, F., Jr., and Kovalchuk, I. Transgenerational adaptation of Arabidopsis to stress requires DNA methylation and the function of Dicer-like proteins. PLoS ONE *5*, e9514.

Boyko, A., Kathiria, P., Zemp, F.J., Yao, Y., Pogribny, I., and Kovalchuk, I. (2007). Transgenerational changes in the genome stability and methylation in pathogen-infected plants: (virus-induced plant genome instability). Nucleic Acids Res *35*, 1714-1725.

Chamnongpol, S., Willekens, H., Moeder, W., Langebartels, C., Sandermann, H., Jr., Van Montagu, M., Inze, D., and Van Camp, W. (1998). Defense activation and enhanced pathogen tolerance induced by H2O2 in transgenic tobacco. Proc Natl Acad Sci U S A *95*, 5818-5823.

de Wit, P.J. (2007). How plants recognize pathogens and defend themselves. Cell Mol Life Sci *64*, 2726-2732.

Durrant, W.E., and Dong, X. (2004). Systemic acquired resistance. Annu Rev Phytopathol *42*, 185-209.

Filkowski, J., Yeoman, A., Kovalchuk, O., and Kovalchuk, I. (2004). Systemic plant signal triggers genome instability. Plant J *38*, 1-11.

Flors, V., Ton, J., van Doorn, R., Jakab, G., Garcia-Agustin, P., and Mauch-Mani, B. (2008). Interplay between JA, SA and ABA signalling during basal and induced resistance against Pseudomonas syringae and Alternaria brassicicola. Plant J *54*, 81-92.

Friedman, A.R., and Baker, B.J. (2007). The evolution of resistance genes in multi-protein plant resistance systems. Curr Opin Genet Dev *17*, 493-499.

Glombitza, S., Dubuis, P.H., Thulke, O., Welzl, G., Bovet, L., Gotz, M., Affenzeller, M., Geist, B., Hehn, A., Asnaghi, C., *et al.* (2004). Crosstalk and differential response to abiotic and biotic stressors reflected at the transcriptional level of effector genes from secondary metabolism. Plant Mol Biol *54*, 817-835.

Iglesias, V.A., and Meins, F., Jr. (2000). Movement of plant viruses is delayed in a beta-1,3-glucanase-deficient mutant showing a reduced plasmodesmatal size exclusion limit and enhanced callose deposition. Plant J *21*, 157-166.

Ingle, R.A., Carstens, M., and Denby, K.J. (2006). PAMP recognition and the plant-pathogen arms race. Bioessays *28*, 880-889.

Jansen, M.A., van den Noort, R.E., Tan, M.Y., Prinsen, E., Lagrimini, L.M., and Thorneley, R.N. (2001). Phenol-oxidizing peroxidases contribute to the protection of plants from ultraviolet radiation stress. Plant Physiol *126*, 1012-1023.

Johnsen, O., Daehlen, O.G., Ostreng, G., and Skroppa, T. (2005). Daylength and temperature during seed production interactively affect adaptive performance of Picea abies progenies. New Phytol *168*, 589-596.

Jones, D.L., Blancaflor, E.B., Kochian, L.V., and Gilroy, S. (2006). Spatial coordination of aluminium uptake, production of reactive oxygen species, callose production and wall rigidification in maize roots. Plant Cell Environ *29*, 1309-1318.

King, E.O., Ward, M.K., and Raney, D.E. (1954). Two simple media for the demonstration of pyocyanin and fluorescin. J Lab Clin Med *44*, 301-307.

Kovacik, J., and Klejdus, B. (2008). Dynamics of phenolic acids and lignin accumulation in metal-treated Matricaria chamomilla roots. Plant Cell Rep *27*, 605-615.

Kovalchuk, I., Kovalchuk, O., Kalck, V., Boyko, V., Filkowski, J., Heinlein, M., and Hohn, B. (2003a). Pathogen-induced systemic plant signal triggers DNA rearrangements. Nature *423*, 760-762.

Kovalchuk, O., Burke, P., Arkhipov, A., Kuchma, N., James, S.J., Kovalchuk, I., and Pogribny, I. (2003b). Genome hypermethylation in Pinus silvestris of Chernobyl--a mechanism for radiation adaptation? Mutat Res *529*, 13-20.

Kunz, B.A., Cahill, D.M., Mohr, P.G., Osmond, M.J., and Vonarx, E.J. (2006). Plant responses to UV radiation and links to pathogen resistance. Int Rev Cytol *255*, 1-40.

Kunz, B.A., Dando, P.K., Grice, D.M., Mohr, P.G., Schenk, P.M., and Cahill, D.M. (2008). UV-induced DNA damage promotes resistance to the biotrophic pathogen Hyaloperonospora parasitica in Arabidopsis. Plant Physiol *148*, 1021-1031.

Kvaalen, H., and Johnsen, O. (2008). Timing of bud set in Picea abies is regulated by a memory of temperature during zygotic and somatic embryogenesis. New Phytol *177*, 49-59.

Kwon, S.I., and An, C.S. (2001). Molecular cloning, characterization and expression analysis of a catalase cDNA from hot pepper (Capsicum annuum L.). Plant Sci *160*, 961-969.

Ledford, H.K., Chin, B.L., and Niyogi, K.K. (2007). Acclimation to singlet oxygen stress in Chlamydomonas reinhardtii. Eukaryot Cell *6*, 919-930.

Lichtenthaler, H.K. (1998). The stress concept in plants: an introduction. Ann N Y Acad Sci *851*, 187-198.

Love, A.J., Yun, B.W., Laval, V., Loake, G.J., and Milner, J.J. (2005). Cauliflower mosaic virus, a compatible pathogen of Arabidopsis, engages three distinct defense-signaling pathways and activates rapid systemic generation of reactive oxygen species. Plant Physiol *139*, 935-948.

Lucht, J.M., Mauch-Mani, B., Steiner, H.Y., Metraux, J.P., Ryals, J., and Hohn, B. (2002). Pathogen stress increases somatic recombination frequency in Arabidopsis. Nat Genet *30*, 311-314.

Mahdavian, K., Kalantari, K.M., and Ghorbanli, M. (2007). The effect of different concentrations of salicylic acid on protective enzyme activities of pepper (Capsicum annuum L.) plants. Pak J Biol Sci *10*, 3162-3165.

McHale, L., Tan, X., Koehl, P., and Michelmore, R.W. (2006). Plant NBS-LRR proteins: adaptable guards. Genome Biol *7*, 212.

Mitsuhara, I., Iwai, T., Seo, S., Yanagawa, Y., Kawahigasi, H., Hirose, S., Ohkawa, Y., and Ohashi, Y. (2008). Characteristic expression of twelve rice PR1 family genes in response to pathogen infection, wounding, and defense-related signal compounds (121/180). Mol Genet Genomics *279*, 415-427.

Molinier, J., Ries, G., Zipfel, C., and Hohn, B. (2006). Transgeneration memory of stress in plants. Nature *442*, 1046-1049.

Morita-Yamamuro, C., Tsutsui, T., Sato, M., Yoshioka, H., Tamaoki, M., Ogawa, D., Matsuura, H., Yoshihara, T., Ikeda, A., Uyeda, I., *et al.* (2005). The Arabidopsis gene CAD1 controls programmed cell death in the plant immune system and encodes a protein containing a MACPF domain. Plant Cell Physiol *46*, 902-912.

Mullineaux, C.W., and Emlyn-Jones, D. (2005). State transitions: an example of acclimation to low-light stress. J Exp Bot *56*, 389-393.

Park, S.W., Kaimoyo, E., Kumar, D., Mosher, S., and Klessig, D.F. (2007). Methyl salicylate is a critical mobile signal for plant systemic acquired resistance. Science *318*, 113-116.

Pecinka, A., Rosa, M., Schikora, A., Berlinger, M., Hirt, H., Luschnig, C., and Mittelsten Scheid, O. (2009). Transgenerational stress memory is not a general response in Arabidopsis. PLoS ONE *4*, e5202.

Ponciano, C.R., Martinez, R., Farenzena, L.S., Iza, P., da Silveira, E.F., Homem, M.G., Naves de Brito, A., and Wien, K. (2006). Electronic sputtering produced by fission fragments on condensed CO and CO2. J Am Soc Mass Spectrom *17*, 1120-1128.

Puchta, H., Swoboda, P., Gal, S., Blot, M., and Hohn, B. (1995). Somatic intrachromosomal homologous recombination events in populations of plant siblings. Plant Mol Biol *28*, 281-292.

Schutzendubel, A., and Polle, A. (2002). Plant responses to abiotic stresses: heavy metal-induced oxidative stress and protection by mycorrhization. J Exp Bot *53*, 1351-1365.

Shin, R., Park, J.M., An, J.M., and Paek, K.H. (2002). Ectopic expression of Tsi1 in transgenic hot pepper plants enhances host resistance to viral, bacterial, and oomycete pathogens. Mol Plant Microbe Interact *15*, 983-989.

Ton, J., and Mauch-Mani, B. (2004). Beta-amino-butyric acid-induced resistance against necrotrophic pathogens is based on ABA-dependent priming for callose. Plant J *38*, 119-130.

Vlot, A.C., Klessig, D.F., and Park, S.W. (2008). Systemic acquired resistance: the elusive signal(s). Curr Opin Plant Biol *11*, 436-442.

Whalen, M.C., Innes, R.W., Bent, A.F., and Staskawicz, B.J. (1991). Identification of Pseudomonas syringae pathogens of Arabidopsis and a bacterial locus determining avirulence on both Arabidopsis and soybean. Plant Cell *3*, 49-59.

Whitham, S., Dinesh-Kumar, S.P., Choi, D., Hehl, R., Corr, C., and Baker, B. (1994). The product of the tobacco mosaic virus resistance gene N: similarity to toll and the interleukin-1 receptor. Cell *78*, 1101-1115.

Table 3.1: Black shank disease assay shows higher resistance of PI_1 plans to P. nicotianae

	0	1	2	3	4	5	Average severity	Incidence
PC_1	0	0	0	57	33	0	3.4	87%
PI_1#9	66	21	3	0	0	0	0.3	26%
PI_1#16	57	24	3	3	3	0	0.5	33%

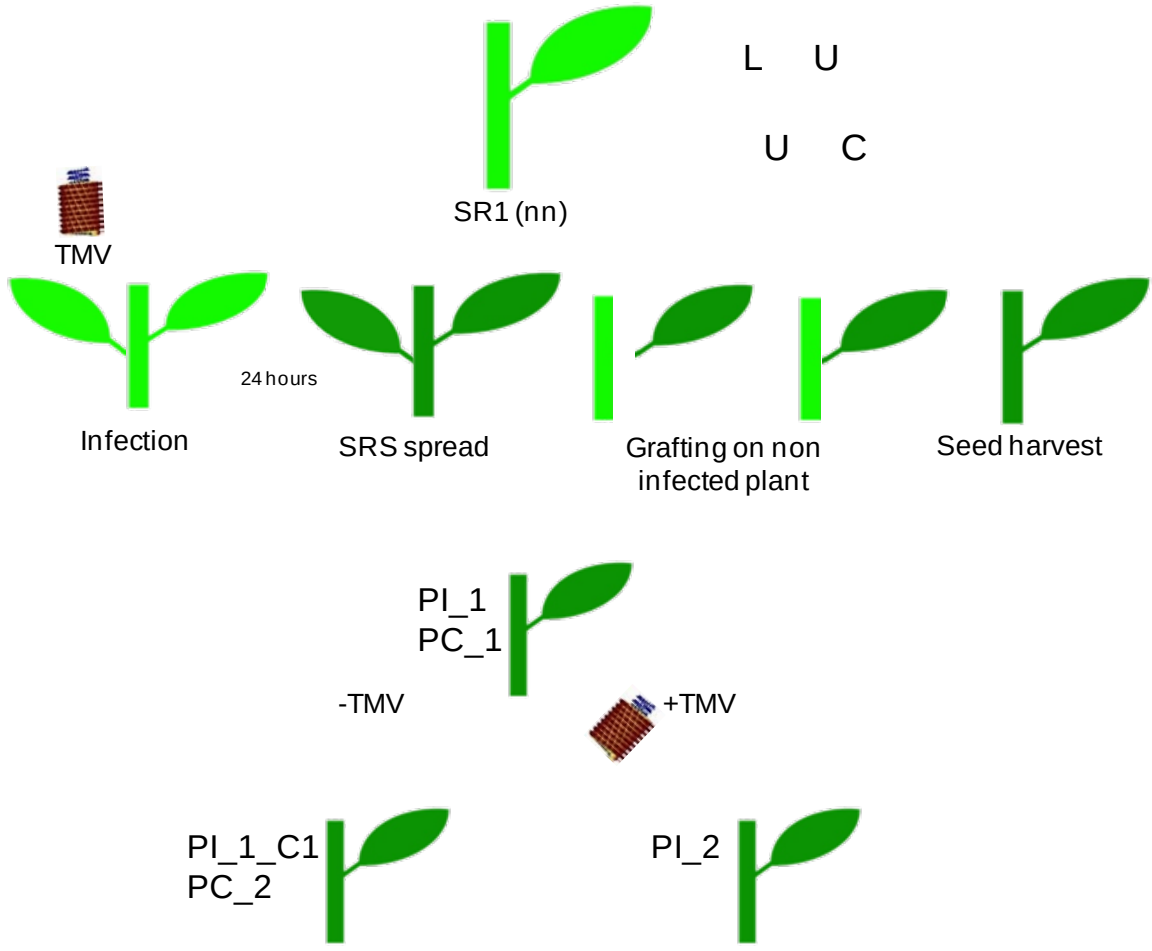


Figure 3.1: Schematic representation of the experimental set-up

Single leaves of 10-week-old SR1 tobacco plants were inoculated with 200 ng *TMV* RNA (20 plants) or mock treated (20 plants). 24 hours after inoculation, upper non-treated leaves (virus-free) of these plants were grafted onto 10 week-old healthy plants (20 plants with leaves of virus-treated plants and 20 plants with leaves of mock-treated plants) from which the tops were previously removed. Darker green shows that these plants generate a systemic signal that can be transmitted via grafting. Seeds derived from newly emerged tissues were collected and named "the progeny of infected" (PI) or "the progeny of control" (PC). To show that this is the first progeny, these plants were named PI_1 and PC_1. To analyze changes in the next generation, these plants were propagated with and without *TMV*. The

second generation of plants was obtained and named PI_2 for the progeny of infected PI_1 plants, PI_1_C1 for the progeny of non-infected PI_1 plants, and PC_2 for the progeny of non-infected PC_1 plants. These seeds were used to analyze HRF and stress tolerance.

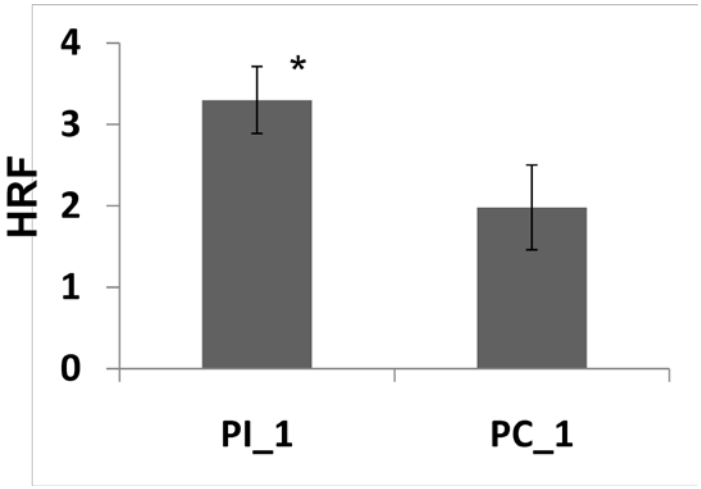


Figure 3.2: Increased frequency of recombination in the first progeny of infected plants

Spontaneous HRF was analyzed in the population of 50-100 plants taken from each of 10 independent progenies of both infected (PI_1) and mock-treated (PC_1) lines, and the analysis repeated three times. The data are shown as an average number of spots per single plants with s.d. The asterisk shows a statistically significant difference between PC_1 and PI_1 plants ($P<0.05$).

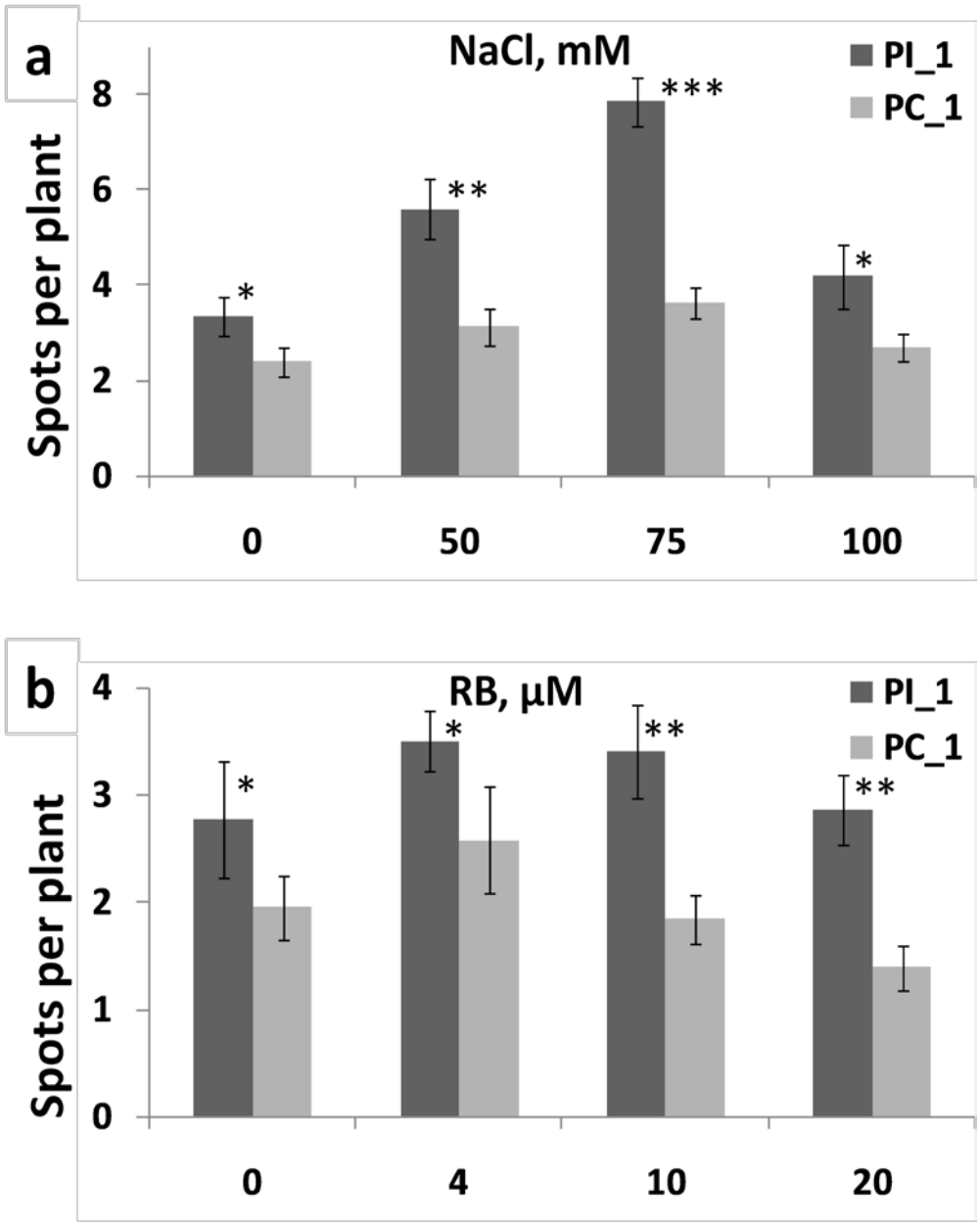


Figure 3.3: PI_1 plants show a higher increase in HRF upon exposure to abiotic stress

PI_1 and PC_1 plants were germinated and grown for one week on sterile MS media. At the age of one week, the plants were moved to MS plates containing various concentrations of either NaCl (a) or rose Bengal (RB) (b). Recombination frequency was assayed at three weeks post germination. The data are shown as an average HRF (with s.d.) per single plant

scored in the population of ~ 100 plants taken from each of 10 independent PI_1 and PC_1 lines. Asterisks show significant differences in exposed plants as compared to control non-exposed plants; a single asterisk indicates $P<0.05$, a double asterisk – $P<0.01$, and a triple asterisk – $P<0.001$.

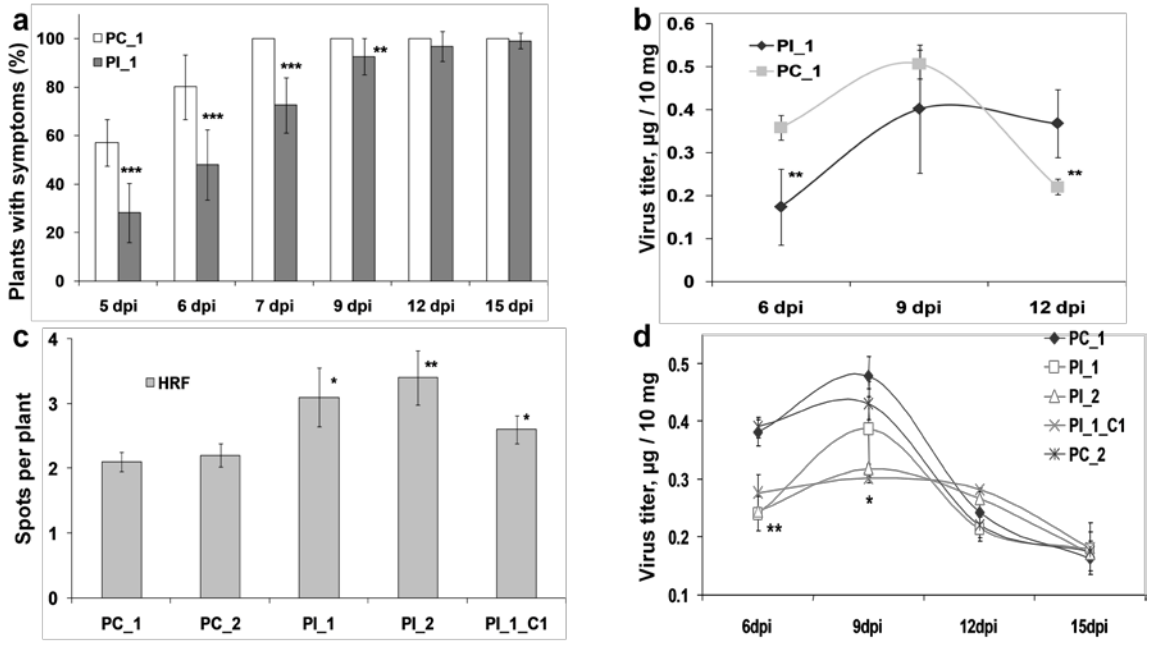


Figure 3.4: PI_1 and PI_2 plants exhibit delayed symptoms and a lower viral titer upon *TMV* infection

Single leaves of five-week-old PI_1 and PC_1 plants were infected with 200 ng *TMV* (see Materials and methods for details). For analysis, 100 PC_1 plants (10 plants from each of ten lines) and 360 PI_1 plants (20 plants from each of 18 different PI lines) were infected. The experiment was performed in three independent sets. Asterisks show a statistically significant difference, where one asterisk stands for $P<0.05$, two asterisks - $P<0.01$ and three asterisks - $P<0.001$.

a. Symptoms were monitored daily, and the data for 5, 6, 7, 9, 12 and 15 dpi are presented. The graphic shows the average percentage (from three independent experiments, with s.d.) of plants with symptoms (out of the total number of infected plants).

b. Virus concentrations were measured in infected plants at 6, 9 and 12 dpi. The graphic shows the average viral concentrations (from 18 different PI_1 lines and 10 different PC lines) as calculated from three independent repeats (in µg/mL of *TMV* with s.d.).

c. Spontaneous non-induced HRF analyzed in three-week-old plants is shown as the average number of recombination events (from three independent experiments with s.d.) in the population of 100-200 (10-20 per individual line) plants in each experimental group. Statistical analysis was performed to identify differences between PI_1 and PC_1, PI_2 and PC_2, and PI_1_C1 and PC_2 plants. Asterisks show statistically significant differences: one asterisk represents $P<0.05$ and two asterisks represent $P<0.01$.

d. Single leaves of five-week-old plants (twenty plants per treatment) were infected with 200 ng *TMV*. The viral titer was analyzed at 6, 9, 12 and 15 dpi. The graphic shows the average viral concentrations as calculated from 3 independent experiments (in µg/mL of *TMV* with s.d.). Statistical analysis was performed to identify differences between PI_1 and PC_1, PI_2 and PC_2, and PI_1_C1 and PC_2 plants. Asterisks show statistically significant differences: one asterisk represents $P<0.05$ and two asterisks represent $P<0.01$.

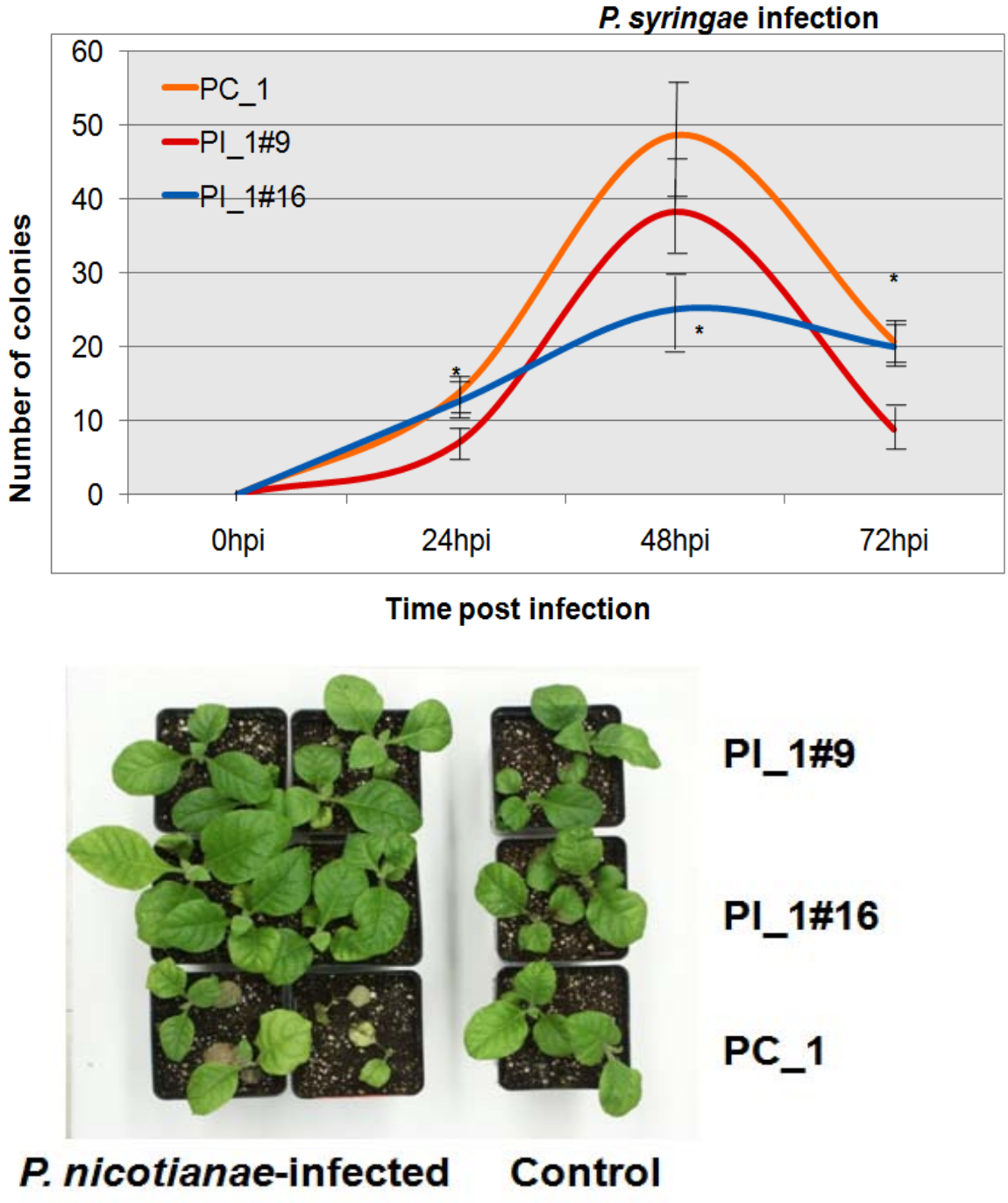


Figure 3.5: PI_1 plants exhibit higher resistance to P. syringae and P. nicotianae

Plants were germinated on soil and infected with the pathogen either at 4 (*P. syringae*) or 6 (*P. nicotianae*) weeks post germination.

a. Figure shows the number of bacterial colonies recovered from infected PC_1 or PI_1 plants (see Methods for details). Bars show standard deviation. Asterisks show significant differences between PC_1 and either of PI_1 lines ($P<0.05$).
b. Disease response of tobacco inoculated with an aggressive isolate of
c. *Phytophthora nicotianae*. PC_1 plants (bottom row) showed significantly more black shank disease symptoms than PI_1#9 (middle row) and PI_1#16 (top row). Uninoculated plants of each line are shown in the pots to the far right.

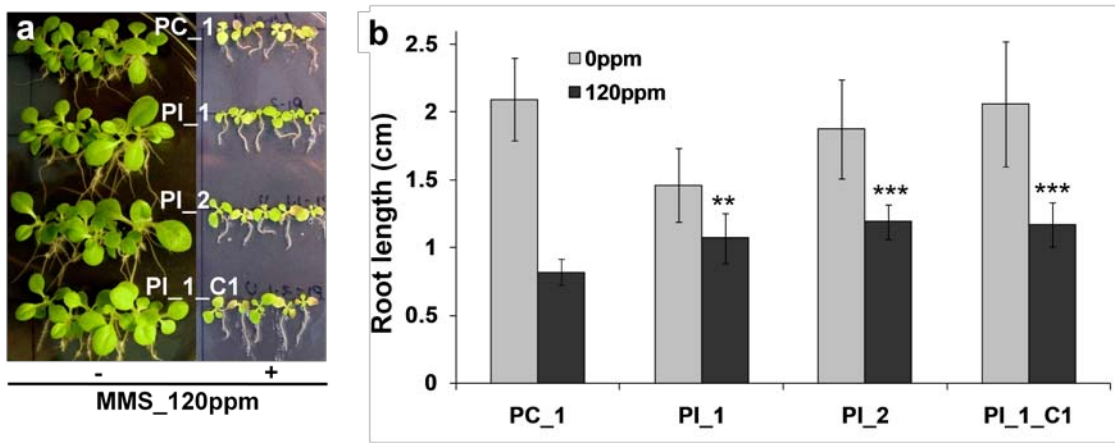


Figure 3.6: The Progeny of virus-infected plants exhibit higher tolerance to MMS

Plants were germinated on either normal media or media supplemented with 120 ppm methyl methane sulfonate.

a. The picture shows representative phenotypes of plants grown with and without methyl methane sulfonate.

b. Root length was measured at two weeks after exposure. The average root length (in cm with s.d.) of 20 plants (4 plates with 5 plants in each) was calculated. Asterisks show statistically significant differences as analyzed between PI_1 and PC_1, PI_2 and PC_1, PI_1_C1 and PC_1 plants exposed to 120 ppm MMS; two asterisks represent $P<0.01$, and three asterisks represent $P<0.001$.

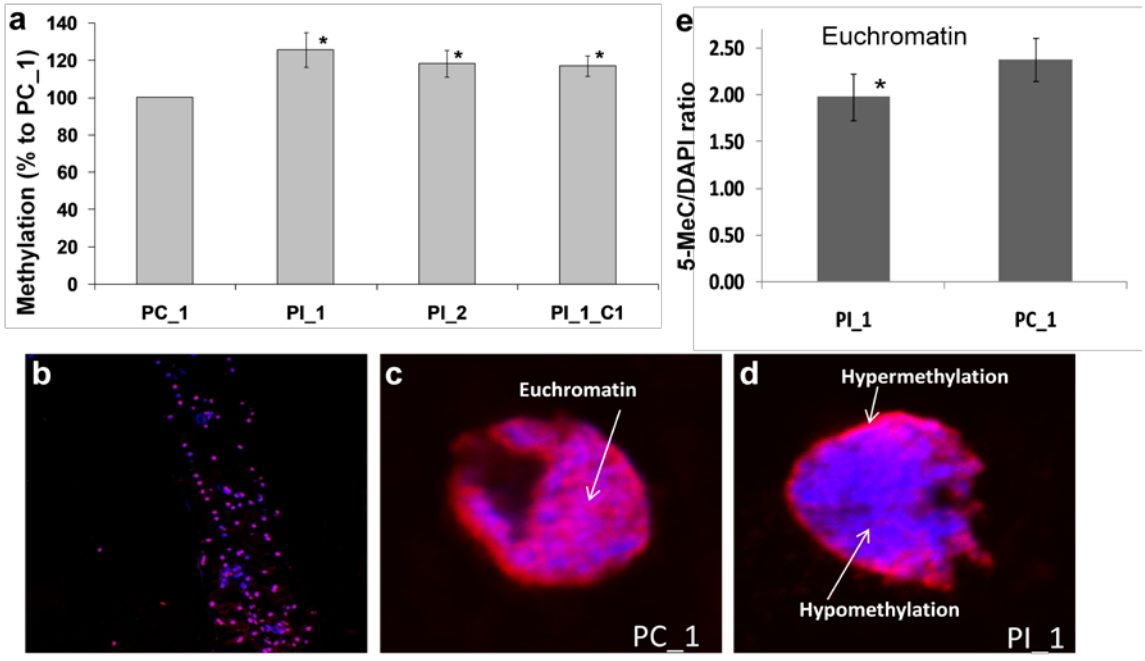


Figure 3.7: PI plants exhibit methylation changes

a. Analysis of methylation using the cytosine extension assay. The data are shown as a percentage of methylation relative to methylation in PC_1 plants (with s.d.). Asterisks show a significant difference ($P<0.05$).

b. Analysis of methylation via immunohistochemistry using anti-5-MeC antibodies. The picture shows a crossectioned leaf at low magnification. Methylated cells are in red.

c. A picture of the cell nucleus of PC_1 plants showing different degrees of methylation in the center of the nucleus (red staining for methylated areas and blue staining for DNA) and on the periphery (the margin of the nucleus, mostly red).

d. A picture of the cell nucleus of PI_1 plants

e. The intensity of staining in the center of the nucleus was analyzed using ImageJ. The data are shown as an average ratio (with s.d.) of 5-MeC to DAPI

staining in the cells of PI_1 and PC_1 plants. Asterisks shows a significant difference between PI_1 and PC_1 plants ($P<0.05$).

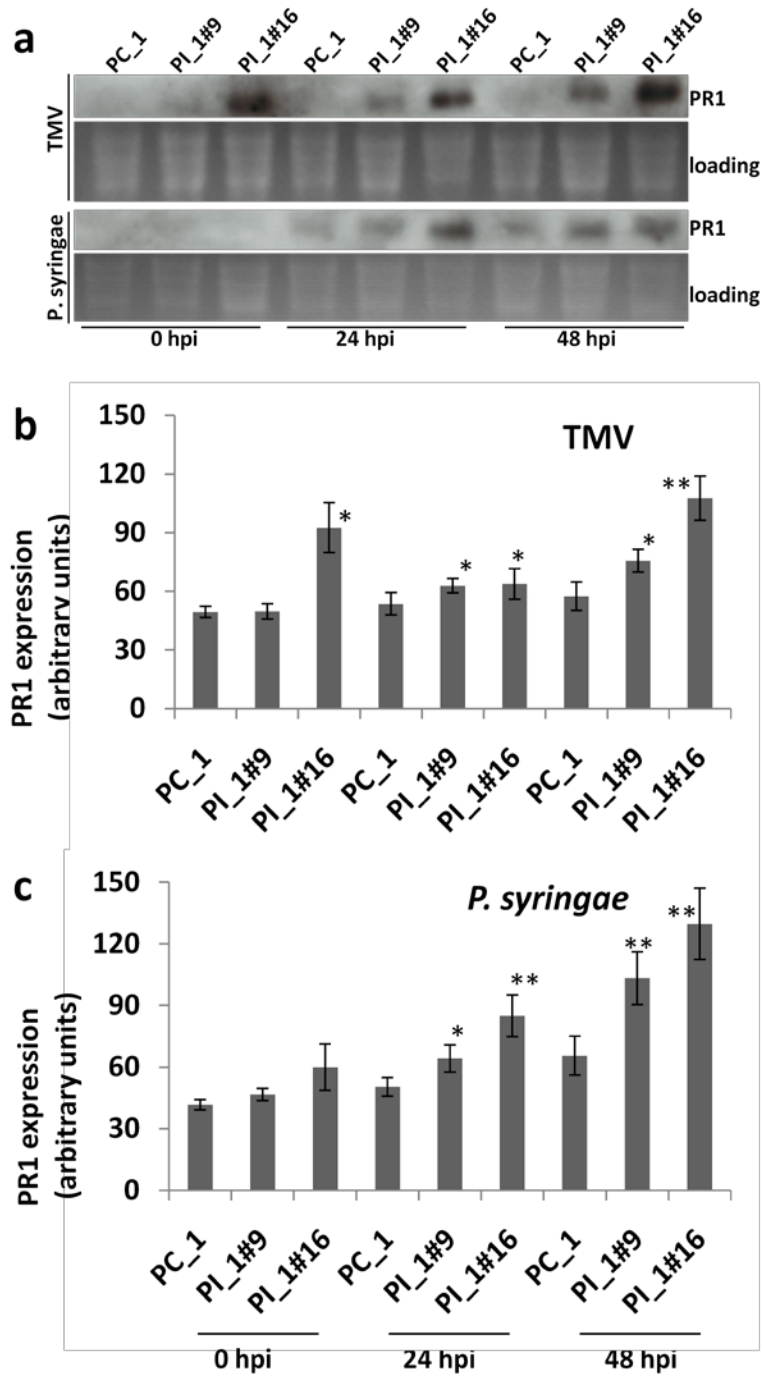


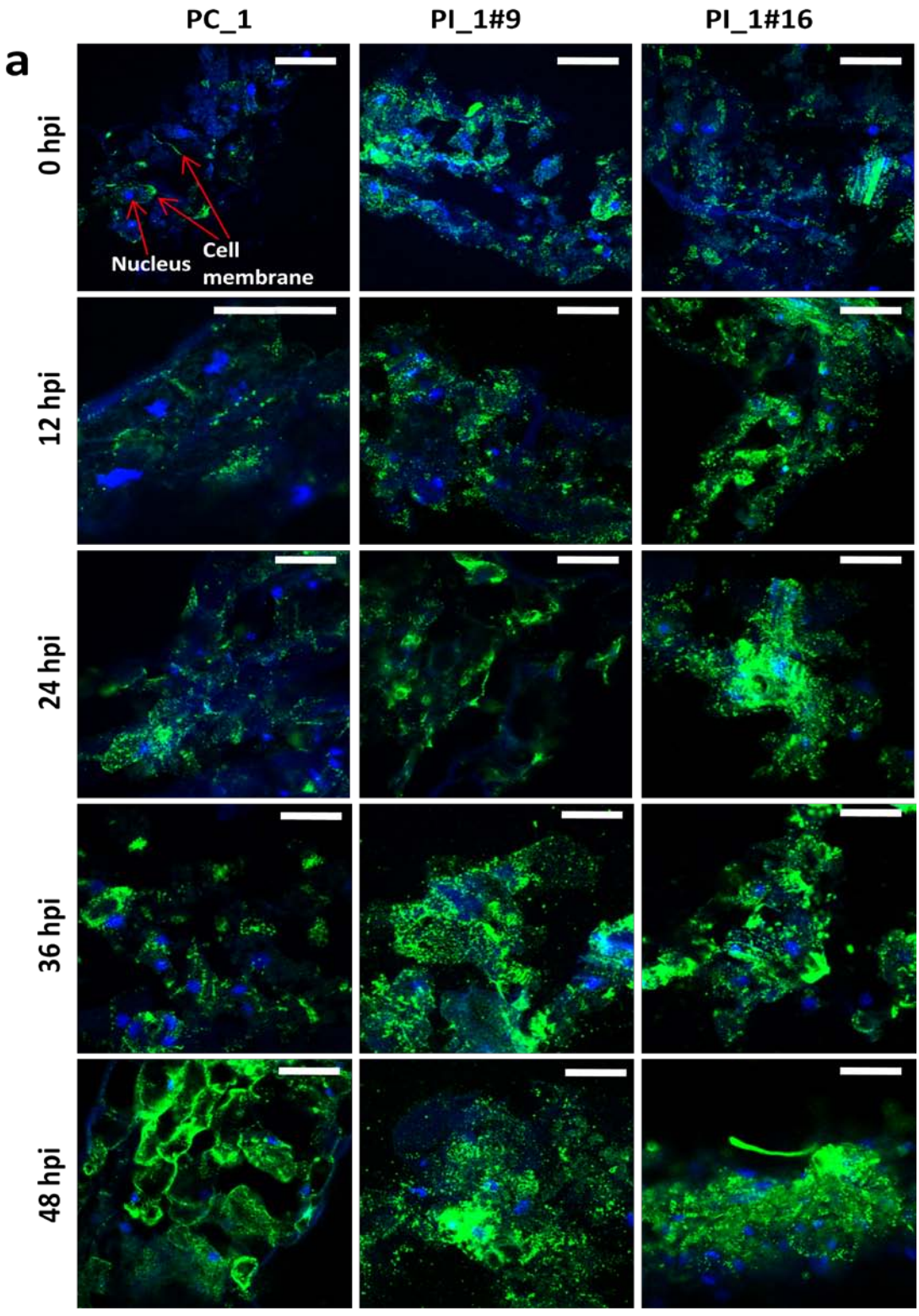
Figure 3.8: PI_1 plants show a higher level of PR1 expression and higher callose deposition

a. The figure shows Northern blot analysis of PR1 expression in PI_1#9, PI#16 and PC_1 plants either infected with *P. syringae* or *TMV* or mock treated.

Samples were taken at 24 and 48 hours post inoculation (hpi). Electrophoresis of total RNA from each sample is shown as a loading control.

b. The figure shows quantification analysis of PR1 expression in *TMV*-infected samples (see Figure 8a). The data are shown as an average of four independent repetitions (in arbitrary units with s.d.). Asterisks show significant differences between infected PI_1 plants and infected PC_1 plants; one - $P<0.05$ and two - $P<0.01$.

c. The figure shows quantification analysis of PR1 expression in *P. syringae*-infected samples (see Figure 8a). The data are shown as an average of four independent repetitions (in arbitrary units with s.d.). Asterisks show significant differences between infected PI_1 plants and infected PC_1 plants; one - $P<0.05$ and two - $P<0.01$.


a
PC_1
PI_1#9
PI_1#16
0 hpi
12 hpi
24 hpi
36 hpi
48 hpi
Nucleus
Cell membrane

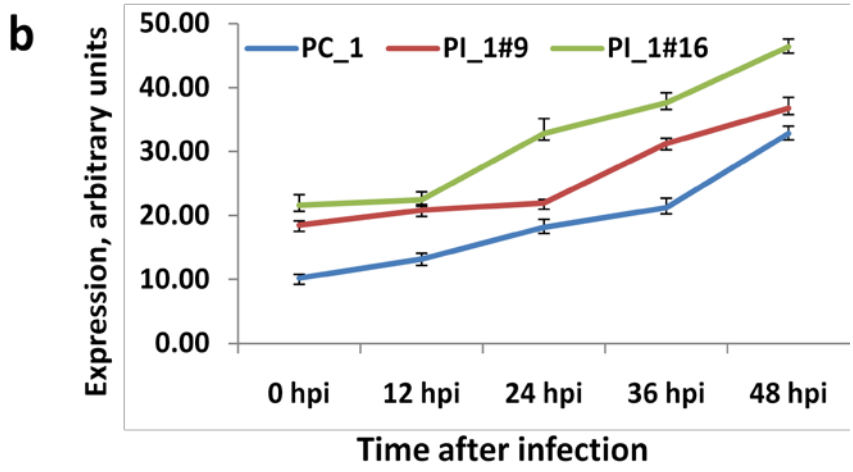
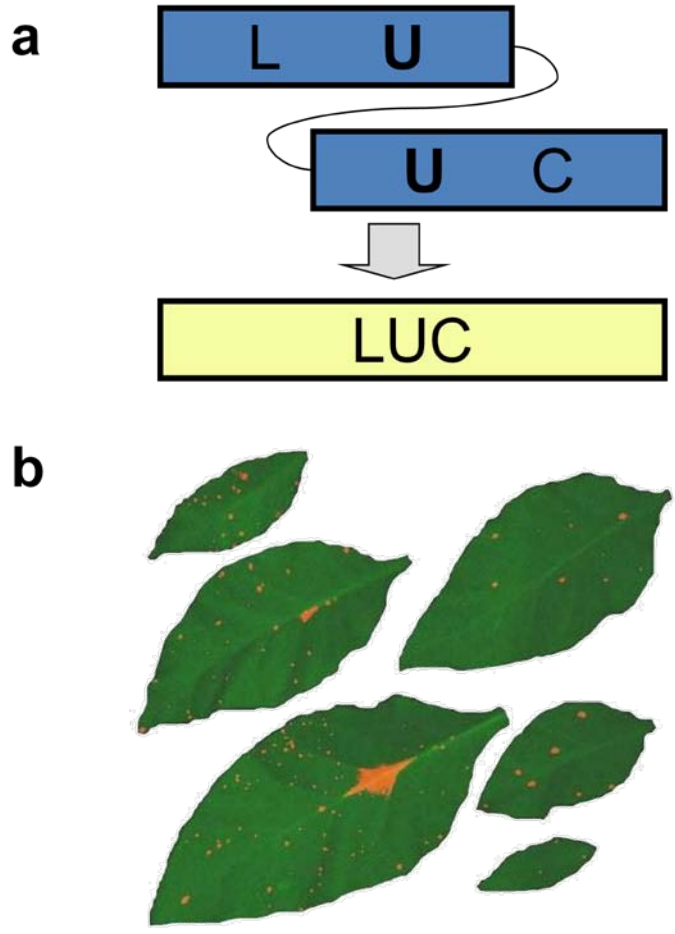


Figure 3.9: PI_1 plants show a higher callose deposition

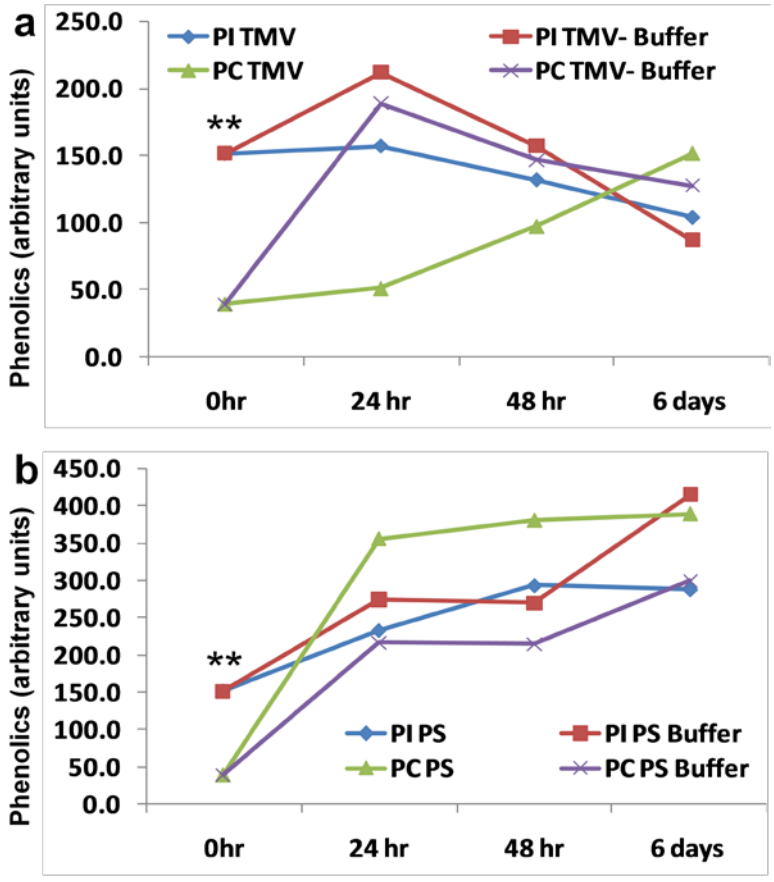
a. Immunohistochemical staining for callose deposition (in green). The callose is labeled by green fluorescence and DNA stain DAPI is observed in blue. Scale bar indicates 40 µm length.

b. Quantification of the intensity of callose-positive staining (with s.e.) (as in Figure 3.9 a). The data were analyzed in three independent experiments (3 leaves from each of 5 plants).



Supplementary Figure 3.1: Reporter gene based recombination construct for analysis of sequence independent recombination rate.

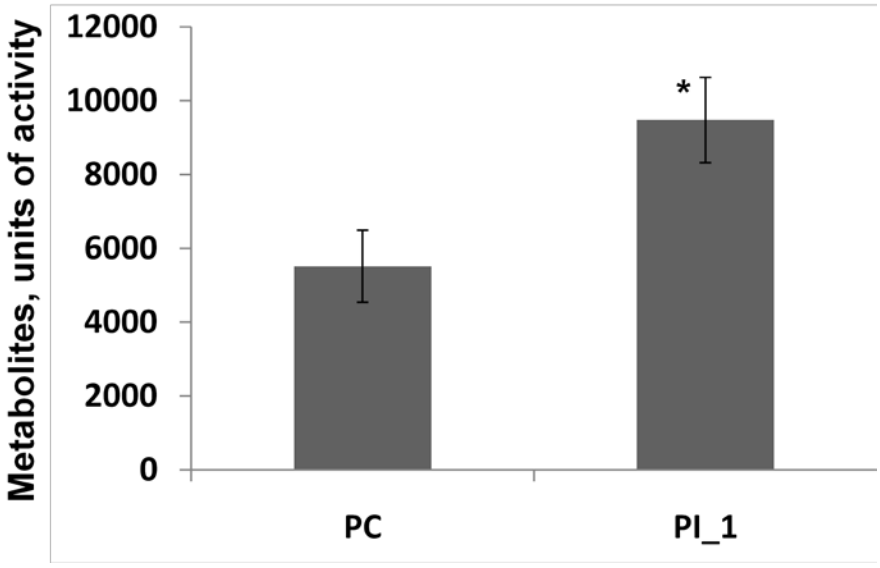
a. Two truncated copies of reporter genes are separated by a filler sequence. The copies are substrate for double strand break repair via homologous recombination. A functional copy of reporter gene results from such a repair.
b. Tobacco plants carrying stable integration of recombination construct. These plants show reporter gene expression in certain tissue sectors. The rate of recombination is determined by calculating number of average spots of reporter gene expression in these plants using a CCD camera.



Supplementary Figure 3.2: Analysis of polyphenols in progeny of infected lines.

a. Amount of polyphenols in PI and PC plants following *TMV* infection at 0, 24, 48h and 6days post infection. The amount of polyphenols was significantly higher in PI plants at 0h time point.

b. Amount of polyphenols in PI and PC plant after infection with *Pseudomonas syringae* at 0, 24, 48h and 6days post infection. The PI plants show significantly higher content of polyphenolics at 0h time point.



Supplementary Figure 3.3: Total secondary metabolite content in PI and PC plants.

The total amount of secondary metabolite present in PI and PC plants was determined using NMR analysis. PI plants had significantly higher content of secondary metabolites in absence of any stress.

4.0 Progeny of tobacco mosaic virus-infected *Nicotiana tabacum* plants exhibit trans-generational changes in metabolic profiles

4.1 Abstract

Our previous work showed that infection of *Nicotiana tabacum* (tobacco) plants with tobacco mosaic virus (*TMV*) results in transgenerational changes in DNA methylation patterns and rearrangements in resistance gene-like loci. We hypothesized that these transgenerational pathogen-induced genetic changes could potentially be associated with transgenerational phenotypic or chemical responses. To test whether a transgenerational phenotypic response can be detected in the progeny of infected (PI) plants, we analyzed their metabolic profiles using Nuclear magnetic resonance (NMR) spectroscopy and gas chromatography-mass spectrometry (GC-MS). In particular, we tested the progeny of two different infected plants and compared their metabolic profiles to the progeny of non-treated control (PC) plants. In total more than 25 water-soluble metabolites were identified and quantified from these three plant classes. Our results showed that PI plants had significantly higher concentrations (50-100% higher) of almost all metabolites in their leaves. Specifically, PI plants had higher levels of various amino acids such as alanine, valine, serine, threonine, leucine and glutamine, as well as various sugars, including sucrose, glucose and fructose. The high abundance of free amino acids and free carbohydrates suggests that the PI plants are in a heightened state of metabolic readiness. This metabolic readiness potentially allows the plants to more easily deploy their small molecule reserves to the production of radicals and salicylates in an effort to combat and potentially localize the

TMV infection. This work shows that viral infections can lead to trans-generational or heritable metabolic changes to plants.

4.2. Introduction

Viral infections in plants are perhaps among the most well studied plant/pathogen interactions. Understanding how plant viruses spread through plants and kill their hosts or become localized and lead to a chronic state of infection has important implications in the prevention, treatment and containment of many crop viruses. In particular, certain types of plant viral infections can result in viral localization at the site of infection; this type of infection is called "incompatible" and occurs when the virus carries an avirulence gene (*Avr*) and the plant carries a resistance (*R*) gene. The most common and well-described case of incompatible infection is the infection of certain cultivars of *Nicotiana tabacum* (tobacco) plants with tobacco mosaic virus (*TMV*). The tobacco cultivar known as "Big Havana" carries in its genome a specific N-gene that makes it resistant to *TMV*. As a result, Big Havana cultivars, when infected with *TMV*, exhibit classical features of an incompatible infection. On the other hand, "compatible" infection is characterized by viral spread from the site of infection and it occurs when either the virus or the plant lack the components for interaction (an *Avr* or *R* gene, respectively). The tobacco cultivar SR1 does not have an active *N*-gene and thus, when infected with *TMV*, cannot restrict viral movement.

Incompatible infection results in what is termed as an immediate local hypersensitive response (HR) consisting of activation of the salicylic acid dependent pathway leading to massive radical production, cell death and localization of viral progression. This is followed by a systemic increase in pathogen tolerance, named systemic acquired resistance (SAR). On

the other hand, a compatible infection does not lead to HR or SAR, but it does result in an immediate plant response in the form of increased radical production. Apparently such a response occurs regardless of whether the plant is resistant (incompatible) or sensitive (compatible) to the pathogen. A second, larger burst of radicals occurs only if the plant has the required *R* gene.

Local infection of SR1 tobacco plants with *TMV* results in virus spread to non-infected tissues, typically occurring within 48 hours post infection, with the first symptoms appearing at 7-8 days post infection. Despite not having a *TMV* resistance gene, these SR1 cultivars are generally able to mount a response to *TMV* infection. The fact that such plants are still able to respond with a transient but noticeable increase in radical production could suggest the generation of some sort of system-wide pathogen warning signal. It is possible that in plants that have an *R*-gene, this second radical burst inactivates the initial warning signal. As a result, plants that do not contain the appropriate *R* gene have the warning signal systematically spread throughout all plant tissues. The findings of a dramatic change in the transcriptome of *Arabidopsis* plants exposed to a compatible pathogen support this hypothesis (De Vos et al., 2005; Love et al., 2005). Tao et al. (2003) actually suggested that a large part of the differences in transcriptional changes between compatible and incompatible viral infections is quantitative rather than qualitative.

As plants do not have a predetermined germline, and plant cells are generally totipotent, it is likely that the changes triggered by this signal can be transmitted to the next generation, thereby allowing later progeny to be better prepared for similar pathogen infections. Our previous experiments showed that local compatible infection of tobacco plants with *TMV* results in a systemic increase in genome instability (Kovalchuk et al., 2003). A more recent report showed that this infection resulted in an increase in the instability of *N*-gene-like *R*-

gene loci in the progeny. This response was paralleled by a global genome hypermethylation and hypomethylation at the *R*-gene loci (Boyko et al., 2007). This is remarkable, as the infected plants belonged to the SR1 cultivar, a cultivar that does not contain an active *N*-gene. Changes observed in *R*-gene loci that exhibit similarity to the *N*-gene could be an indication of a transgenerational response to a compatible infection. Our most recent work indicated that the progeny of infected tobacco plants had an increase in spontaneous non-induced recombination frequency, increased tolerance to *TMV* and a higher level of certain phenolic compounds.

Given the significant transcriptomic changes seen in the progeny of *TMV* infected plants we hypothesized that these progeny would also have other downstream (i.e. metabolic) changes that allow these plants to tolerate pathogen infection. To test the hypothesis, we have checked the levels of metabolites in control and infected plants and found the progeny of infected plants to have a dramatically different metabolome as compared to the progeny of non-infected plants.

4.3 Experimental procedures

4.3.1. Infection of SR1 plants with *TMV* and generation of the progeny of infected plants

The generation of transgenic tobacco cultivar SR1 plants line #A carrying in the genome a single copy of the luciferase (LUC) recombination substrate was described previously (Kovalchuk et al., 2003). For these experiments, single leaves of twenty *Nicotiana tabacum* cultivar SR1 plants were infected with 200 ng (100 ng/μl in phosphate buffer) of the *TMV*

U1 strain. Twenty control plants were mock-inoculated with phosphate buffer only. Carborundum was used as an abrasive to facilitate the viral infection. The experiments were carried out in two independent rounds, with 10 plants from each group in each round. Infected and mock-inoculated plants were kept at 32°C to allow faster progression of the viral infection. At 24 h after infection, the infected leaves were cut and the plants were allowed to produce seeds. Cutting leaves at 24 h post infection prevented the spread of virus to non-infected tissue (as checked by PCR) and symptom appearance at two weeks after infection (Kovalchuk et al., 2003). The progeny of infected and non-infected plants was used for the analysis of metabolites. The progeny plants were also tested for the presence of *TMV* and no virus was found (data not shown).

4.3.2. Infection of the progeny of infected and progeny of control plants with *TMV*

To prepare material for metabolomic analysis, a total of twenty three-week-old plants, corresponding to the progeny of infected (10 plants) and the progeny of control plants (10 plants), were used. For the progeny of infected plants we used two independent lines, corresponding to the progeny of infected #9 (PI #9) and the progeny of infected #19 (PI #19); for the progeny of control (PC) we used five plants each from two independent lines, PC #1 and PC #3. Three-week-old plants were either infected with *TMV* (100 ng per each infected leaf; 200 ng total), or mock-treated. To reduce the number of samples for metabolic profiling, we combined plants from PC #1 and PC #3. Samples (treated leaves) were collected 48 h after treatment; control samples were collected before infection. Three independent repeats were made. Thus, we obtained 27 samples, 3 samples of PI #9, PI #19

and PC plants at time “0”, 3 samples of mock-treated leaves of PI #9, PI #19 and PC plants at 48 h, and 3 samples of virus-infected PI #9, PI #19 and PC plants at 48 h.

4.3.3 Metabolite extraction

Up to 3 frozen leaves from each sample (wt. 0.95 g) were taken into a 50 mL glass tube and were quickly immersed in 3 mL isopropanol at 75ºC (preheated) with 0.01% BHT (butyrated hydroxytoluene) (Sigma-Aldrich) for 15 min. 1.5 mL chloroform and 0.6 mL water were added and vortexed; then agitated at room temperature for 1h. The liquid extracts were then transferred to a centrifuge tube. Four milliliters of 2:1 chloroform: methanol solution with 0.01% BHT were again added to the leaves and shaken for 30 min. The liquid extracts were then transferred back to the original centrifuge tube.

This extraction procedure was repeated until the leaves of every sample became white (usually about 5 extractions were needed, including the one with isopropanol). One milliliter of 1M KCl was added to the combined extract, vortexed, centrifuged and the upper phase (containing the water soluble metabolites) was collected. A second extraction was performed with 2 mL of distilled water which was added to the lower phase, vortexed, centrifuged, and the upper phase was collected. All the water soluble extracts were combined and stored in the freezer at -20^{o}C until further analysis.

4.3.4. NMR sample preparation and spectroscopy

The extracted tobacco leaf solution (1 mL) was evaporated to dryness using a Speedvac concentrator. The dried samples (0.035 g) were reconstituted in 250 μL H_2O. Fifty

microliters of 50 mM NaH_2PO_4 buffer (pH 7), 35 µL of D_2O and 15 µL of a standard buffer solution (3.73 mM DSS (disodium-2, 2-dimethyl-2-silapentane-5-sulphonate and 0.47% NaN_3 in H_2O) were added to the sample. The sample was vortexed for 1 minute and sonicated for 30 minutes, and then transferred to a standard Shigemi microcell NMR tube for subsequent spectral analysis. All ^{1}H-NMR spectra were collected on a 500 MHz Inova (Varian Inc., Palo Alto, CA) spectrometer equipped with a 5 mm HCN Z-gradient pulsed-field gradient (PFG) room-temperature probe. ^{1}H-NMR spectra were acquired at 25 °C using the first transient of the tnnoesy-presaturation pulse sequence, which was chosen for its high degree of quantitative accuracy. Spectra were collected with 4096 transients using a 4 s acquisition time and a 1 s recycle delay.

4.3.5 NMR compound identification and quantification

All FIDs were zero-filled to 64k data points and subjected to line broadening of 0.5 Hz. The singlet produced by the DSS methyl groups was used as an internal standard for chemical shift referencing (set to 0 ppm) and for compound quantification. All ^{1}H-NMR spectra were processed and analyzed using the Chenomx NMR Suite Professional software package version 5.0 (Chenomx Inc., Edmonton, AB). The Chenomx NMR Suite software allows for qualitative and quantitative analysis of an NMR spectrum by manually fitting spectral signatures of metabolites from an internal database of metabolite spectra to the spectrum of interest (Weljie et al., 2006). Specifically, the spectral fitting for each metabolite was done using the standard Chenomx 500 MHz, pH 4-9 metabolite library. It has been previously shown that this fitting procedure provides absolute concentration accuracies of 90 % or better (Saude et al., 2004). Moreover, concentration data was corrected for bandpass filter

attenuation as previously described (Saude et al., 2009). Each spectrum was processed and analyzed by at least two NMR spectroscopists to minimize compound misidentification and misquantification. We also used sample spiking to confirm the identities of a number of compounds. Sample spiking involves adding 20-200 µM of the suspected compound to selected tobacco leave samples and examining whether the relative NMR signal intensity changed as expected.

4.3.6 GC-MS compound identification and quantification

Identification of metabolites and quantification of metabolite concentrations is often much more tedious via GC-MS than it is with NMR spectroscopy. To derivatize the metabolites for GC-MS analysis, 40 µL of 20 mg/mL methoxyamine hydrochloride (Sigma-Aldrich) in ACS grade pyridine was added to the water-soluble tobacco leaf extracts and incubated at room temperature for 16 hours. Then 50 µL (N-Methyl-N-trifluoroacetamide) MSTFA (Thermo Scientific) with 1% TMCS (Trimethylchlorosilane) derivatization agent were added and incubated at 37 °C for 60 minutes on a hotplate. The samples were vortexed twice throughout the incubation period to ensure complete dissolution. Samples were refrigerated at 4°C for no longer than 48 h prior to analysis in order to avoid any degradation of the derivatized compounds.

Derivatized extracts were analyzed using an Agilent 7890-5975C GC-MS instrument operating in an Electron Impact (EI) ionization mode. For GC-MS analysis, 2 µL of the derivatized tobacco leaf solution were injected using a split/splitless injector with a split ratio of 5:1 onto a HP-5MS capillary column (Agilent J&W Scientific, 30m × 250 µm × 0.25 µm). The injector port temperature was held at 250 °C and the helium carrier gas flow rate

was set to 1 mL/min at the initial oven temperature of 70 °C. The oven temperature was increased at 1 °C /min to 76 °C, and then at 6.1 °C/min to 310 °C. The total run time was 45 minutes. The full scan mode of the quadrupole MS was used at a mass range of 50-500 m/z, with a solvent delay of 6 minutes. The MS ion source temperature was 230 °C and the Quadrupole temperature was 150 °C. In GC-MS, a faster scan speed generally provides more data points across a chromatographic peak, but it tends to lower the ion statistics. In contrast, a slower scan rate produces few scans over the peak and results in better spectra. The scan speed of our quadrupole MS was optimized over a number of samples and it was found that a relatively slow scan rate of 1.7 scans/s gave the best results was used throughout all experiments.

The AMDIS spectral deconvolution software (Version 2.62) from NIST (National Institute of Standards and Technology) was used to process the total ion chromatogram and the EI-MS spectra of each GC peak. After deconvolution, the purified mass spectrum of each of the trimethylsilated metabolites was identified using the NIST MS Search program (version 2.0d) which was linked to the NIST mass spectral library (2005). Retention Indices (RIs) were calculated using a C8-C20 alkane mixture standard (Fluka, Sigma-Aldrich). Metabolites were identified by matching the EI-MS spectra with those of reference compounds from NIST library. In addition safety checks were performed by using additional published GC-MS libraries (Schauer et al., 2005) for the rapid identification of metabolites in complex biological samples. Peaks having no match to published retention indices and/or no match to the AMDIS GC/MS spectral library were identified by matching the experimental RI of each metabolite with an in-house RI library (containing 312 TMS-derivatized metabolites) developed in our laboratory (Wishart et al., 2009). Thirty-five pure standards obtained from the Human Metabolome Library (Wishart et al., 2007) were run

through the GC–MS (using the same protocol described above) to confirm their identity, retention indices and EI spectra and were subsequently used for producing external five-point calibration curves (for quantification). Where peak baseline resolution was not observed, peak deconvolution software was used to separate peaks based on Gaussian shape recognition. Using the mass spectral information obtained in this manner, peaks were successfully identified and peak areas successfully quantified.

4.4 Results

4.4.1 Experimental set-up and generation of samples

For all the experiments described here, we used tobacco cultivar SR1 plants (Boyko et al., 2007). As described above, single leaves of three week-old plants were infected with 200 ng (100 ng/leaf) of *TMV*, whereas control plants were mock-treated. The seeds from the collected plants were used for future analysis. In this particular pilot experiment, the plantlets from the progeny of infected plants (PI) were tested for viral presence by Real-Time PCR or by infection of another group of five-week old SR1 plants. Both methods confirmed that the seeds of PI plants were virus-free (data not shown). The progeny of the control plants (PC) were also confirmed to be virus-free.

Based on previous genomic and transcriptomic studies, we hypothesized that PI plants would have different metabolic expression profiles than PC plants. To test whether significant metabolic changes occur upon infection with *TMV*, we inoculated three-week-old PI and PC plants with 200 ng of *TMV*.

4.4.2 NMR Identification and Quantification

As seen in figure 4.1, the NMR spectra from the tobacco leaf samples are relatively simple. Typically 90% of all visible peaks in each NMR spectrum could be definitively assigned to a compound and more than 90% of the spectral area could be routinely fit using the Chenomx spectral analysis software. As shown in Figure 4.1, most of the visible peaks are annotated with a compound name. Among the 27 tobacco leave samples, a total of 20 compounds were routinely identified and quantified, with the most abundant compounds being malic acid (0.9 mM), fructose (0.6 mM), glucose (0.5 mM), nicotine (0.5 mM) and choline (0.3 mM). In general, the pattern was similar for all the samples, however, for a few samples, some metabolites such as sucrose, fructose, glucose and serine could not be identified/quantified, but are present in large quantities in other samples.

4.4.3 GC-MS Identification and Quantification

Most of the visible peaks in our GC-MS spectra could be annotated with a compound name. Unlike in the situation with NMR, where peak coverage approached 90%, the fraction of all GC-MS peaks that could be positively identified was approximately 70%. This relatively low level of coverage is a common problem in global or untargeted GC-MS metabolomics studies (Kopka, 2006; Schauer et al., 2005). It is possible that some of the unidentified peaks may be non-metabolic by-products from the chemical derivatization process, or they may be degraded by-products of derivatized metabolites. However, the lack of a comprehensive GC-MS library specifically for plant metabolites (the NIST mass spectra library contains

only a relatively small portion of metabolically relevant compounds), is likely an equally important reason for this incomplete coverage.

All the peaks corresponding to identified metabolites were verified by spiking or subsequent GC-MS runs with pure standards. In total we were able to identify and quantify 15 metabolites via GC-MS. For full identification, the mass spectra of the identified compounds not only had to match the EI-MS spectra in the NIST database, but also the RI of the compounds in our RI library which consists of 312 TMS-derivatized metabolites. All these metabolites were quantified by GC-MS using external calibration. The lower limit of detection for most GC-quadrupole-MS instruments is about 1 uM.

Comparison of the metabolites pattern detected by NMR and GC-MS methods determined a common set of 10 compounds, while NMR detected 10 compounds that GC-MS methods could not detect (Table 4.1). Additionally our GC-MS analysis led to the detection of 5 compounds that NMR could not detect (Table 4.1). The reasons for these differences could be manifold. For those compounds found in NMR but not in GC-MS, it may be that the metabolites of interest were either too volatile for GC-MS detection, lost in sample preparation or eluted during the solvent delay. For those compounds found by GC-MS but not by NMR, the concentrations could have been below the NMR detection limits. In general, GC-MS appears to have better sensitivity (sub-µM) compared to NMR spectroscopy (a few µM) (Wishart, 2008). In all cases, the existence of NMR detectable metabolites was explicitly checked in our GC-MS analyses and vice versa.

4.4.4 The progeny of infected plants exhibited substantial differences in metabolite levels

The analysis of differences between PC and PI plants showed that both PI lines, PI #9 and PI #19 had the higher total concentration of detected metabolites prior to infection (Figure 4.2). The total metabolite concentration was ~ 100% higher in PI #9 and ~50% higher in PI #19 lines as compared in PC lines (Figure 4.2; "0 h" samples). A t-test indicated that the concentration data for line PI #9 was significantly different, whereas the data for line #19 was not ($P<0.01$ and $P>0.1$, respectively). Comparison of the concentrations measured by NMR and GC-MS methods showed a very high degree of similarity; with six different metabolites showing identical concentrations, as analyzed for all three groups in the "0 h" samples (Figure 4.3, Figure 4.4).

Analysis of individual metabolites in the "0 h" samples showed most of them being higher in PI plants as compared to PC plants, although significant differences (for at least one of the PI lines) were observed for 15 out of the 25 unique metabolites (Figures 4.5-4.8). Metabolites were grouped into four different groups: group 1 - amino acids (alanine, glutamate, leucine, serine, threonine and valine); group 2 – sugars (fructose, glucose, sucrose) and those involved in sugar metabolism (malate, fumarate); group 3 and group 4 that included all the other detected metabolites. A general trend was observed for group 1; there was a higher level of most of the amino acids in PI plants at 0h and there was an increase in the amount of amino acids in samples of 48 h and 48 h_V in PC plants, especially noticeable for 48 h_V samples (Figure 4.5). For the group 2, that included sugars, there also was a trend; we found higher level of sugars in PI plants and the amount of all sugars drastically dropped after mock treatment (48 h) or viral infection (48 h_V) (Figure 4.6).

Group 3 (Figure 4.7) and group 4 (Figure 4.8) included all other detected metabolites. Group 3 involved the metabolites that had different levels in PC and PI plants in non-infected tissue but had changes in the metabolites level in mock-treated and infected tissue (Figure 4.7). Group 4 included metabolites that had higher level in PI plants than in PC plants in non-infected tissue but did not change in mock-treated or infected tissue (figure 8).

It is interesting to note that the concentration of most metabolites was found to be higher in the PI #9 line as compared to PI #19 line; with the difference being most obvious for 4-aminobutyrate, 3-hydroxyisovalerate, 3-hydroxybutyrate, choline, threonine, alanine and fumarate (Figures 4.5-4.7). Analysis of changes in the infected samples (48 h_V) and in the mock-inoculated samples (48 h) showed that most metabolite concentrations were similar. After infection, however, higher levels were still observed in 9 out of 26 unique metabolites in at least one of PI lines. At the same time, the level of sugars such as fructose, glucose and sucrose dropped substantially in PC, PI #9 and PI #19 (Figure 4.6). In fact, the concentrations of fructose and sucrose fell below detectable levels (Figure 4.6).

4.5 Discussion

Metabolites in plants play a major role in many resistance and stress responses. The identification and quantification of metabolites is of paramount importance to study the dynamics of the metabolome, to analyze fluxes in metabolic pathways and to translate the role of each metabolite following various stimuli. By a simple extraction, a broad variety of primary and secondary compounds can be observed, and conclusions can be drawn about the major changes occurring in the metabolic pathways leading to the various metabolites. Here,

we tested and compared the metabolic profiles of the progeny of two different infected plants to the progeny of non-treated control plants.

The progeny of stressed plants had higher concentrations of sugars such as sucrose, glucose, and fructose and amino acids such as proline, alanine, phenylalanine, leucine and threonine (Figure 4.5 and Figure 4.6). Sucrose, glucose and fructose are the major carbohydrate store in majority of the plants when compared to most other plant metabolites. Thus, carbohydrate accumulation has been, historically, one of the most commonly used indicators for virus-induced alterations in plant metabolism during both compatible and incompatible interactions. Abnormal accumulation or disappearance of carbohydrates, especially sucrose is diagnostic for net alterations in the balance between those processes responsible for creation and utilization of carbohydrate namely, photosynthesis and respiration. Stress often results in substantial changes in metabolite profiles. Elevated levels of sugars metabolites and amino acids were found in plants exposed to temperature (Cook et al., 2004; Kaplan et al., 2007), water and salinity (Brosche et al., 2005; Cramer et al., 2007; Kim et al., 2007), sulfur (Nikiforova et al., 2005), phosphorus (Hernandez et al., 2007), oxidative (Baxter et al., 2007) and heavy metal (Le Lay et al., 2006) stresses as well as a combination of multiple stresses (Rizhsky et al., 2004). The present results of indicating an elevated level of sugar metabolites in progeny of infected plants at 0h compared to non infected control plants at 0h also agree with the previous findings.

Our metabolic profiling experiments also showed that *TMV* infection of PI or PC plants reduced the levels of glucose and fructose. Curiously, Sanchez et al. (2008) also observed a dramatic dose-dependent depletion of the level of fructose and glucose but not sucrose upon exposure of *Arabidopsis* to salt stress (Sanchez et al., 2008). A comparison of

steady-state metabolite levels in *Arabidopsis* and a related extremophile shrub known as *T. halophila* showed higher amounts of sugars (e.g., sucrose, fructose and glucose), along with various organic and amino acids (e.g., proline and citric, malic and succinic acids) in the shrub even prior to stress exposure (Gong et al., 2005). It has been known for many decades that changes in the accumulation of carbohydrates frequently precede the appearance of virus symptoms (Handford and Carr, 2007). Earlier studies (Handford and Carr, 2007) have also demonstrated that early in the infection process, even before symptoms are discernible, virus infection alters both the starch production during the day, and its degradation and/or mobilization at night. Many studies carried out on a variety of host-virus systems, have shown that partitioning of newly fixed carbon between soluble sugars (sucrose, fructose and glucose) and organic and amino acids is perturbed. A decrease in soluble sugar content in infected tissue has been observed in Chinese cabbage infected with Turnip yellow mosaic virus (TYMV) and Squash mosaic virus (SqMV)-infected squash. In contrast, sucrose levels in Zucchini yellow mosaic virus (ZYMV)-infected marrow plants increased relative to levels in healthy controls (Handford and Carr, 2007). Metabolomics approach to study the carbohydrate partitioning on viral infection in *A. thaliana* by feeding *Col 0* plants with $^{13}CO_2$ 17 days after inoculation with two strains of *TMV* known to cause mild or severe symptoms in tobacco was carried out (Handford et al., 2004). The results demonstrated that the incorporation of ^{13}C-label into carbohydrates, amino acids and organic acids was 2-4 times higher in virus-inoculated plants compared to the mock-inoculated control (Handford et al., 2004). Thus, in a wide variety of plant-virus interactions there is a shift in carbon partitioning away from soluble sugars towards organic and amino acids, although this is not necessarily the case in every system analyzed.

The local infection is mainly characterized by sucrose, glucose, alanine, glutamine, proline metabolites. At 0h high levels of primary metabolites such as sugars and tricarboxylic acid (TCA) cycle metabolites particularly malate were observed (Figure 4.6). In 48h post-infected leaves a definitely different metabolome was observed (Figure 4.6). The NMR profiles corresponding to alanine, glutamine and valine clearly increased when compared with leaves at 0h. On the contrary, sucrose, glucose, nicotine and malate, showed a lower level in the infected leaves when compared with control leaves (Figures 4.6 and Figure 4.7). The dicarboxylic acid malate has a multitude of functions in plant metabolism and homeostasis. These range from its most prominent roles in the mitochondrial TCA cycle and in crassulacean acid metabolism (CAM) and C4 metabolism, to its roles as an osmoticum, as a regulator of pH homeostasis, as a reducing equivalent that is shuttled between subcellular compartments (Martinoia et al., 2000). As an intermediate of the TCA cycle, malate is intimately associated with mitochondrial energy metabolism and is also the origin of carbon skeletons exported from the mitochondrion in support of amino acid biosynthesis (Schneidereit et al., 2006).

In addition, we observed a significant increase in 4-aminobutyrate concentration in the progeny of infected plant at 0 h compared to the control plants at 0h (Figure 4.7). A similar trend in the level of 4-aminobutyrate was also observed in mock treated (48 h) and virus-treated plants (48 h post infection). However, this level was significantly less compared to that observed in control and progeny of infected plants at 0 h. 4-aminobutyrate, a non-protein amino acid represents an important component of the free amino acid pool in plants (Shelp et al., 1999). It has been demonstrated that GABA is involved in many physiological incidents in plants such as stress protection, contribution to C: N balance, regulation of cytosolic pH, defense against pathogens, and a role of osmoregulation in the cell (Bouche

and Fromm, 2004). In addition a new role has been attributed to GABA, the role of a signaling compound involved in plant glutamate regulation (Bouche et al., 2003).

Figure 4.7 shows variations in the levels of methanol and formate in control and progeny of infected plants at 0 h compared to plants at 48 h post infection. Leaves contain small pools of methanol and the major source of this methanol is probably pectin demethylation (Fall and Fall, 1998). Plants also contain small, metabolically active pools of formate. Like methanol, formate can be emitted from leaves to the atmosphere, although emission rates appear to be lower than for methanol. In leaves of C3 plants in the light, the formate is generally considered to come from the non-enzymatic, H_2O_2-dependent decarboxylation of glyoxylate, perhaps facilitated by the peroxidatic activity of catalase. Further, environmental stresses, phytopathogenic infections strongly reduce the levels of formate by strongly inducing formate dehydrogenase (Li et al., 2001; Olson et al., 2000). In addition, the patterns of development in plants and environmental regulation imply that there may be major metabolic fluxes via formate in certain tissues and conditions (Hanson and Roje, 2001). However, the sources and control of the amounts of formate in plants remains unclear.

In summary, the results may indicate that *TMV* infection results in a redirection of metabolic activity from a growth oriented phase to a stress response oriented phase. The elevated levels of metabolites in the progeny of infection at 0 h compared to control plants at 0 h may represent a role in stress tolerance which warrants further investigation. The present results extends our previous findings that the progeny of infected plants exhibit many changes in plant genome stability, methylation pattern, stress tolerance, amount of phenolic compounds as well as the concentration of various metabolites. Thus the present approach of

metabolic profiling using both NMR and GC-MS could provide a detailed understanding of metabolic fluxes involved in complex mechanisms regulating metabolic pathways in plants.

4.6 References:

Baxter, C.J., Redestig, H., Schauer, N., Repsilber, D., Patil, K.R., Nielsen, J., Selbig, J., Liu, J., Fernie, A.R., and Sweetlove, L.J. (2007). The metabolic response of heterotrophic Arabidopsis cells to oxidative stress. Plant Physiol *143*, 312-325.

Bouche, N., and Fromm, H. (2004). GABA in plants: just a metabolite? Trends Plant Sci *9*, 110-115.

Bouche, N., Lacombe, B., and Fromm, H. (2003). GABA signaling: a conserved and ubiquitous mechanism. Trends Cell Biol *13*, 607-610.

Boyko, A., Kathiria, P., Zemp, F.J., Yao, Y., Pogribny, I., and Kovalchuk, I. (2007). Transgenerational changes in the genome stability and methylation in pathogen-infected plants: (virus-induced plant genome instability). Nucleic Acids Res *35*, 1714-1725.

Brosche, M., Vinocur, B., Alatalo, E.R., Lamminmaki, A., Teichmann, T., Ottow, E.A., Djilianov, D., Afif, D., Bogeat-Triboulot, M.B., Altman, A., *et al.* (2005). Gene expression and metabolite profiling of Populus euphratica growing in the Negev desert. Genome Biol *6*, R101.

Cook, D., Fowler, S., Fiehn, O., and Thomashow, M.F. (2004). A prominent role for the CBF cold response pathway in configuring the low-temperature metabolome of Arabidopsis. Proc Natl Acad Sci U S A *101*, 15243-15248.

Cramer, G.R., Ergul, A., Grimplet, J., Tillett, R.L., Tattersall, E.A., Bohlman, M.C., Vincent, D., Sonderegger, J., Evans, J., Osborne, C., *et al.* (2007). Water and salinity stress in grapevines: early and late changes in transcript and metabolite profiles. Funct Integr Genomics *7*, 111-134.

De Vos, M., Van Oosten, V.R., Van Poecke, R.M., Van Pelt, J.A., Pozo, M.J., Mueller, M.J., Buchala, A.J., Metraux, J.P., Van Loon, L.C., Dicke, M., *et al.* (2005). Signal signature and transcriptome changes of Arabidopsis during pathogen and insect attack. Mol Plant Microbe Interact *18*, 923-937.

Fall, A.L., and Fall, R. (1998). High-level expression of ice nuclei in Erwinia herbicola is induced by phosphate starvation and low temperature. Curr Microbiol *36*, 370-376.

Gong, Q., Li, P., Ma, S., Indu Rupassara, S., and Bohnert, H.J. (2005). Salinity stress adaptation competence in the extremophile Thellungiella halophila in comparison with its relative Arabidopsis thaliana. Plant J *44*, 826-839.

Handford, M.G., and Carr, J.P. (2007). A defect in carbohydrate metabolism ameliorates symptom severity in virus-infected Arabidopsis thaliana. J Gen Virol *88*, 337-341.

Handford, M.G., Sicilia, F., Brandizzi, F., Chung, J.H., and Dupree, P. (2004). Arabidopsis thaliana expresses multiple Golgi-localised nucleotide-sugar transporters related to GONST1. Mol Genet Genomics *272*, 397-410.

Hanson, A.D., and Roje, S. (2001). One-Carbon Metabolism in Higher Plants. Annu Rev Plant Physiol Plant Mol Biol *52*, 119-137.

Hernandez, G., Ramirez, M., Valdes-Lopez, O., Tesfaye, M., Graham, M.A., Czechowski, T., Schlereth, A., Wandrey, M., Erban, A., Cheung, F., *et al.* (2007). Phosphorus stress in common bean: root transcript and metabolic responses. Plant Physiol *144*, 752-767.

Kaplan, F., Kopka, J., Sung, D.Y., Zhao, W., Popp, M., Porat, R., and Guy, C.L. (2007). Transcript and metabolite profiling during cold acclimation of Arabidopsis reveals an intricate relationship of cold-regulated gene expression with modifications in metabolite content. Plant J *50*, 967-981.

Kim, J.K., Bamba, T., Harada, K., Fukusaki, E., and Kobayashi, A. (2007). Time-course metabolic profiling in Arabidopsis thaliana cell cultures after salt stress treatment. J Exp Bot *58*, 415-424.

Kopka, J. (2006). Current challenges and developments in GC-MS based metabolite profiling technology. J Biotechnol *124*, 312-322.

Kovalchuk, I., Kovalchuk, O., Kalck, V., Boyko, V., Filkowski, J., Heinlein, M., and Hohn, B. (2003). Pathogen-induced systemic plant signal triggers DNA rearrangements. Nature *423*, 760-762.

Le Lay, P., Isaure, M.P., Sarry, J.E., Kuhn, L., Fayard, B., Le Bail, J.L., Bastien, O., Garin, J., Roby, C., and Bourguignon, J. (2006). Metabolomic, proteomic and biophysical analyses of Arabidopsis thaliana cells exposed to a caesium stress. Influence of potassium supply. Biochimie *88*, 1533-1547.

Li, R., Wong, N., Jabali, M.D., and Johnson, P. (2001). CD44-initiated cell spreading induces Pyk2 phosphorylation, is mediated by Src family kinases, and is negatively regulated by CD45. J Biol Chem *276*, 28767-28773.

Love, A.J., Yun, B.W., Laval, V., Loake, G.J., and Milner, J.J. (2005). Cauliflower mosaic virus, a compatible pathogen of Arabidopsis, engages three distinct defense-signaling pathways and activates rapid systemic generation of reactive oxygen species. Plant Physiol *139*, 935-948.

Martinoia, E., Massonneau, A., and Frangne, N. (2000). Transport processes of solutes across the vacuolar membrane of higher plants. Plant Cell Physiol *41*, 1175-1186.

Nikiforova, V.J., Kopka, J., Tolstikov, V., Fiehn, O., Hopkins, L., Hawkesford, M.J., Hesse, H., and Hoefgen, R. (2005). Systems rebalancing of metabolism in response to sulfur

deprivation, as revealed by metabolome analysis of Arabidopsis plants. Plant Physiol *138*, 304-318.

Olson, B.J., Skavdahl, M., Ramberg, H., Osterman, J.C., and Markwell, J. (2000). Formate dehydrogenase in Arabidopsis thaliana: characterization and possible targeting to the chloroplast. Plant Sci *159*, 205-212.

Rizhsky, L., Liang, H., Shuman, J., Shulaev, V., Davletova, S., and Mittler, R. (2004). When defense pathways collide. The response of Arabidopsis to a combination of drought and heat stress. Plant Physiol *134*, 1683-1696.

Sanchez, D.H., Siahpoosh, M.R., Roessner, U., Udvardi, M., and Kopka, J. (2008). Plant metabolomics reveals conserved and divergent metabolic responses to salinity. Physiol Plant *132*, 209-219.

Saude, E.J., Lacy, P., Musat-Marcu, S., Mayes, D.C., Bagu, J., Man, S.F., Sykes, B.D., and Moqbel, R. (2004). NMR analysis of neutrophil activation in sputum samples from patients with cystic fibrosis. Magn Reson Med *52*, 807-814.

Saude, E.J., Obiefuna, I.P., Somorjai, R.L., Ajamian, F., Skappak, C., Ahmad, T., Dolenko, B.K., Sykes, B.D., Moqbel, R., and Adamko, D.J. (2009). Metabolomic biomarkers in a model of asthma exacerbation: urine nuclear magnetic resonance. Am J Respir Crit Care Med *179*, 25-34.

Schauer, N., Steinhauser, D., Strelkov, S., Schomburg, D., Allison, G., Moritz, T., Lundgren, K., Roessner-Tunali, U., Forbes, M.G., Willmitzer, L., *et al.* (2005). GC-MS libraries for the rapid identification of metabolites in complex biological samples. FEBS Lett *579*, 1332-1337.

Schneidereit, J., Hausler, R.E., Fiene, G., Kaiser, W.M., and Weber, A.P. (2006). Antisense repression reveals a crucial role of the plastidic 2-oxoglutarate/malate translocator DiT1 at the interface between carbon and nitrogen metabolism. Plant J *45*, 206-224.

Shelp, B.J., Bown, A.W., and McLean, M.D. (1999). Metabolism and functions of gamma-aminobutyric acid. Trends Plant Sci *4*, 446-452.

Weljie, A.M., Newton, J., Mercier, P., Carlson, E., and Slupsky, C.M. (2006). Targeted profiling: quantitative analysis of 1H NMR metabolomics data. Anal Chem *78*, 4430-4442.

Wishart, D.S. (2008). Applications of metabolomics in drug discovery and development. Drugs R D *9*, 307-322.

Wishart, D.S., Knox, C., Guo, A.C., Eisner, R., Young, N., Gautam, B., Hau, D.D., Psychogios, N., Dong, E., Bouatra, S., *et al.* (2009). HMDB: a knowledgebase for the human metabolome. Nucleic Acids Res *37*, D603-610.

Wishart, D.S., Tzur, D., Knox, C., Eisner, R., Guo, A.C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S., *et al.* (2007). HMDB: the Human Metabolome Database. Nucleic Acids Res *35*, D521-526.

Table 4.1: Unique metabolites detected by NMR and GC-MS

NMR-specific	GC-MS-specific	Common: NMR/GC-MS
2-Methylglutarate	Butanoic acid	3-Hydroxybutyrate
3-Hydroxyisovalerate	Propanoic acid	4-Aminobutyrate
Choline	Acetamide	Alanine
Formate	Valine	Fructose
Fumarate	Butylamine	Glucose
Glutamate		Leucine
Glutamine		Malate
Methanol		Serine
Nicotine		Threonine
Phenylalanine		Sucrose

Table shows the list of detectible metabolites as measured by NMR or GC-MS. 'NMR-specific' shows the list of metabolites detected only by NMR. 'GC-MS-specific' shows the list of metabolites detected only by GS-MS. Whereas 'NMR/GC-MS' shows the list of metabolites detected by both.

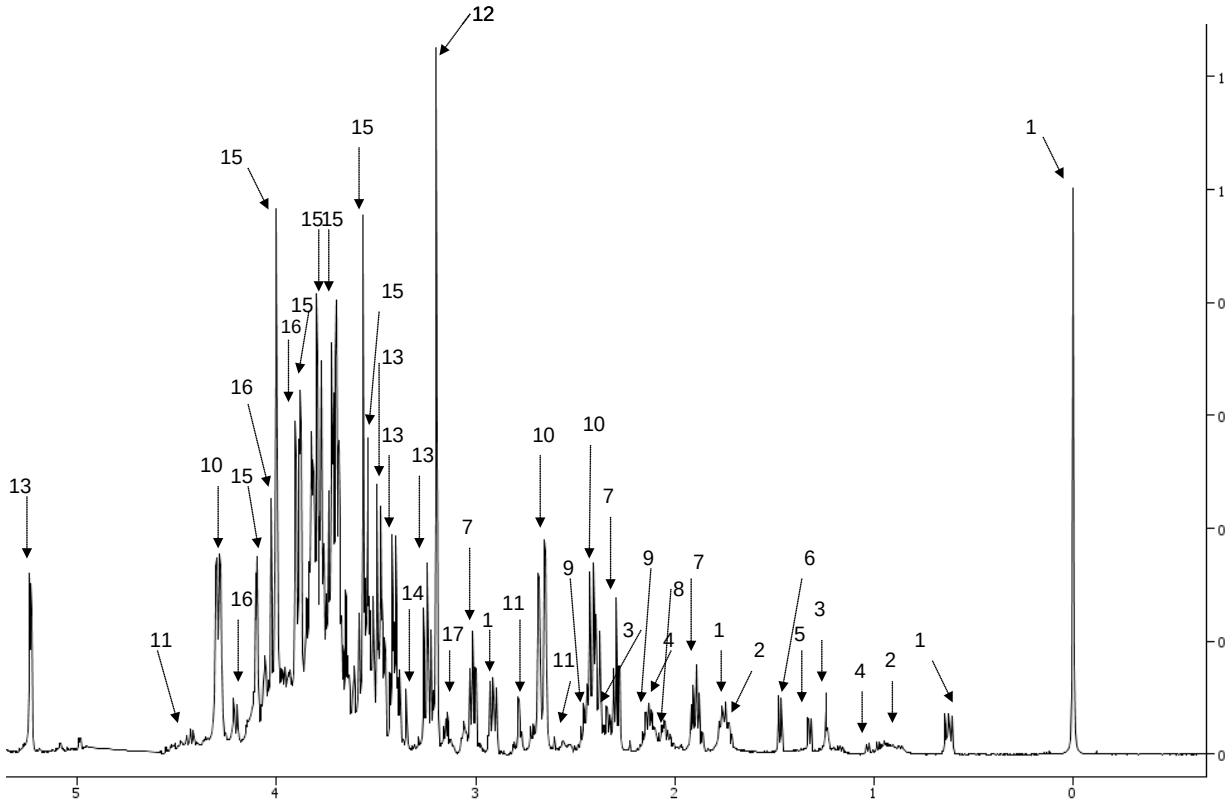


Figure 4.1: Typical 500 MHz 1H-NMR spectrum of a Tobacco leaf sample.

Numbers indicate the following metabolites: 1. DSS; 2. Leucine; 3. 3-hydroxyisovalerate; 4. 2-methylglutarate; 5. Threonine; 6. Alanine; 7. 4-aminobutyrate; 8. Glutamate; 10. Glutamine; 10. Malate; 11. Nicotine; 12. Choline; 14. Methanol; 15. Fructose; 16. Sucrose; 17. Phenylalanine

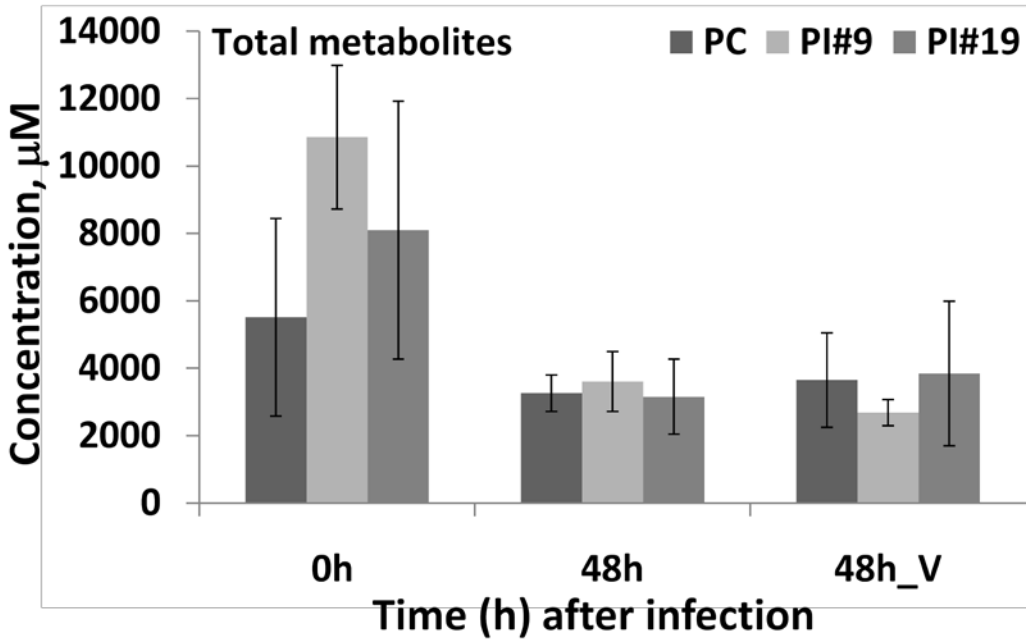


Figure 4.2: Total metabolite levels, as measured by NMR

PC, PI#9 and PI#19 plants were infected with *TMV* and samples were taken at 48 hpi (hour post infection) – "48h_V". The samples from non-infected tissues were taken at 0h – "0h" and at 48h – "48h". The metabolites were detected by NMR. Total concentrations of detected metabolites for each sample were obtained by adding up the concentrations of each individually detected metabolite. The data for PC, PI#9 and PI#19 plants are shown as averages (with SD) from three independent experiments.

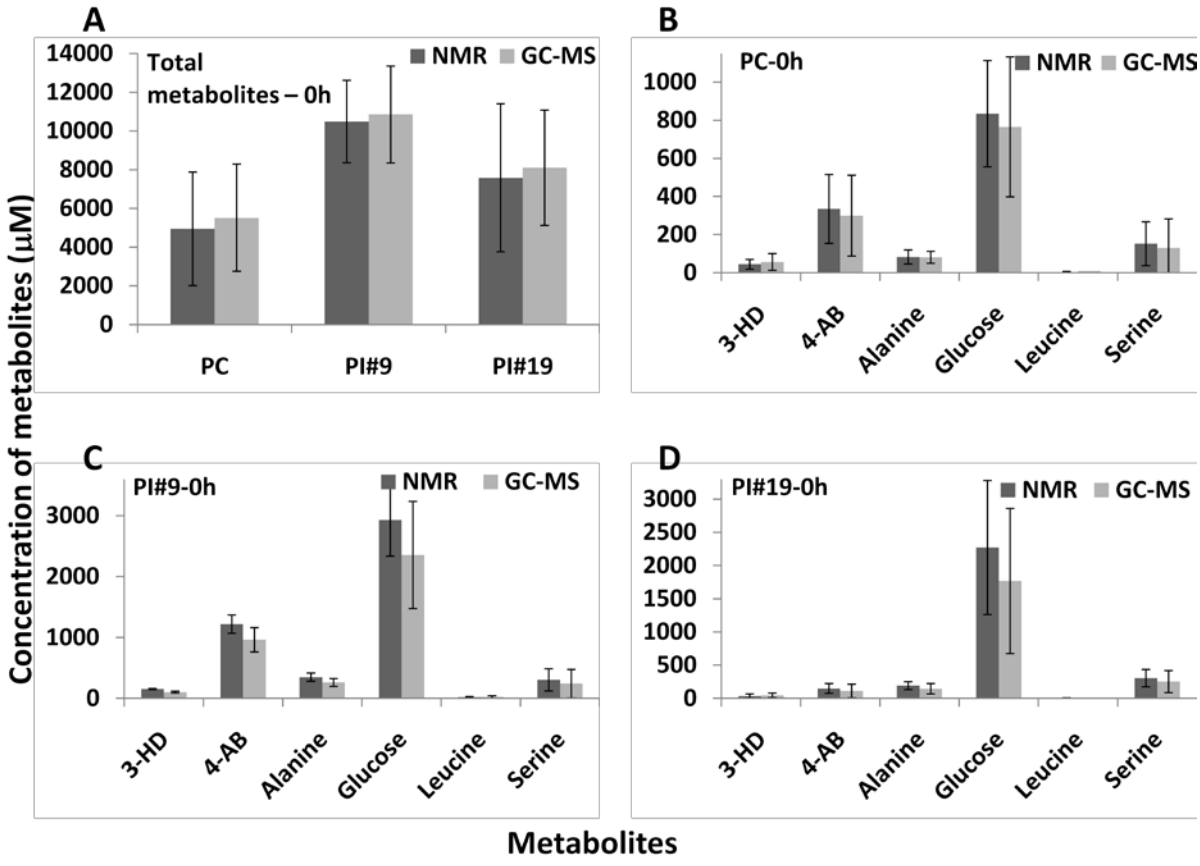


Figure 4.3: Comparison of the average metabolite concentrations in non-infected PC, PI #9 and PI #19 plants, as measured by NMR and GC-MS

The concentrations of metabolites in non-infected tissue of PC, PI #9 and PI #19 plants at 0 h time point was compared for NMR or GC-MS methods. "3-HB" is 3-Hydroxybutyrate, whereas "4-AB" is 4-Aminobutyrate.

A. Shows the comparison for total amount of all metabolites.

B. Shows the comparison for six common metabolites as measured in PC plants.

C. Shows the comparison for six common metabolites as measured in PI#9 plants.

D. Shows the comparison for six common metabolites as measured in PI#19 plants.

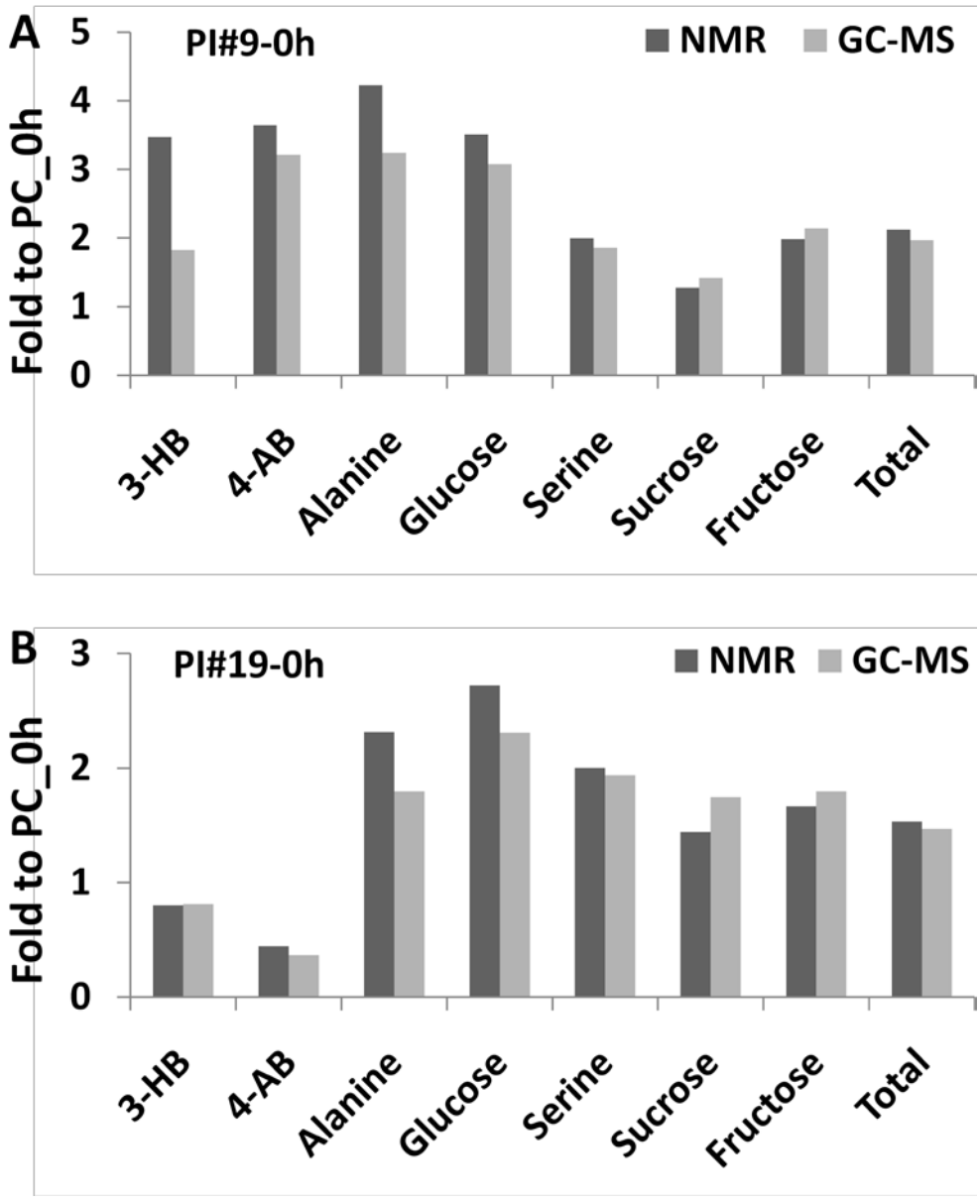


Figure 4.4: Comparison of fold difference in the amount of metabolites as measured by NMR and GC-MS

Fold difference in the amount of metabolites between PI #9 (A) and PC or PI #19 (B) and PC samples collected from non-infected plants was analyzed. The comparison of fold differences is shown for NMR and GC-MS analyses. "3-HB" is 3-Hydroxybutyrate, whereas "4-AB" is 4-Aminobutyrate.

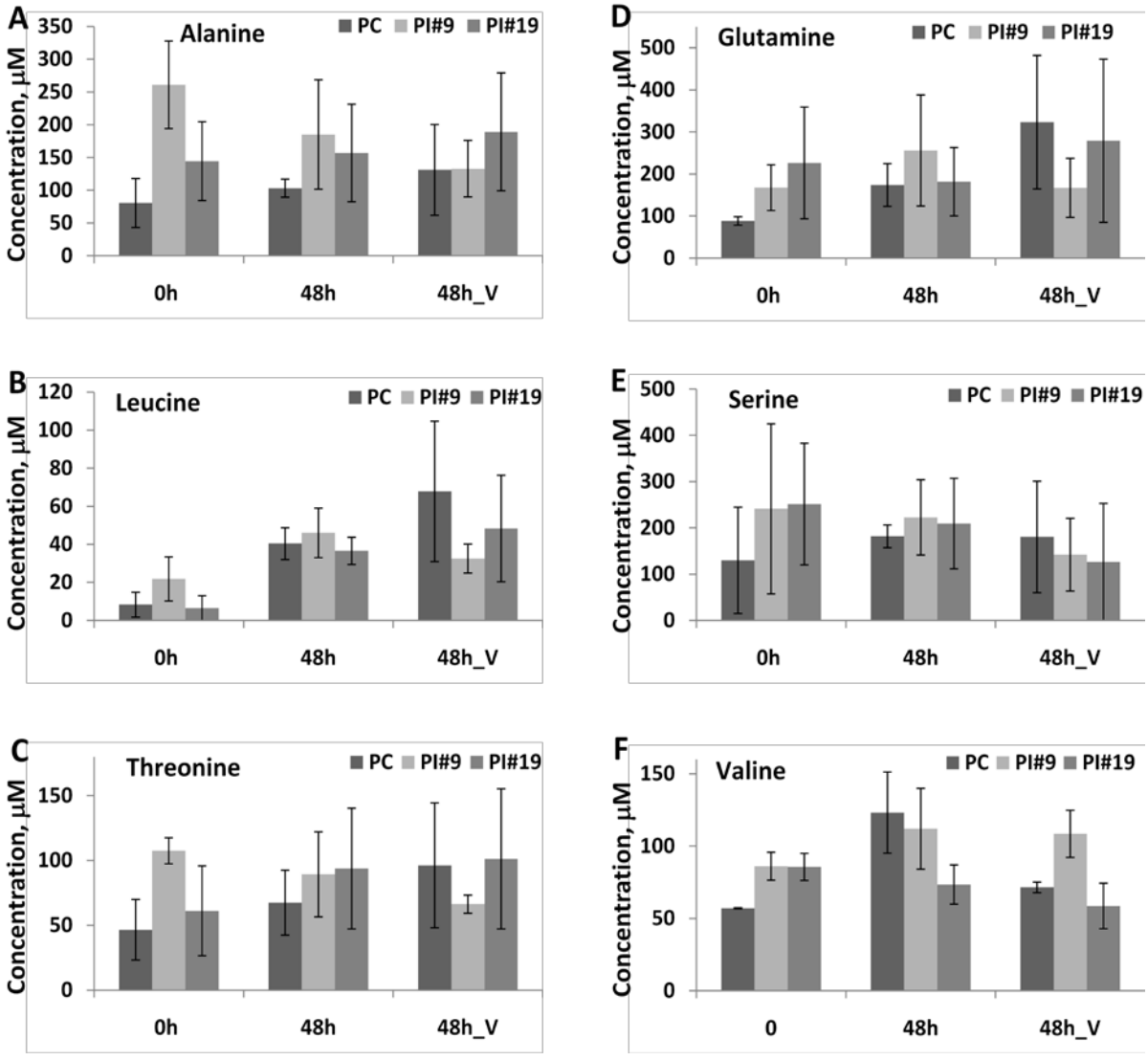


Figure 4.5: the concentrations of amino acids, as measured by NMR

Figures A-F shows the average (with SD) concentration of alanine, leucine, threonine, glutamine, serine and valine, respectively in PC, PI #9 and PI #19 plants.

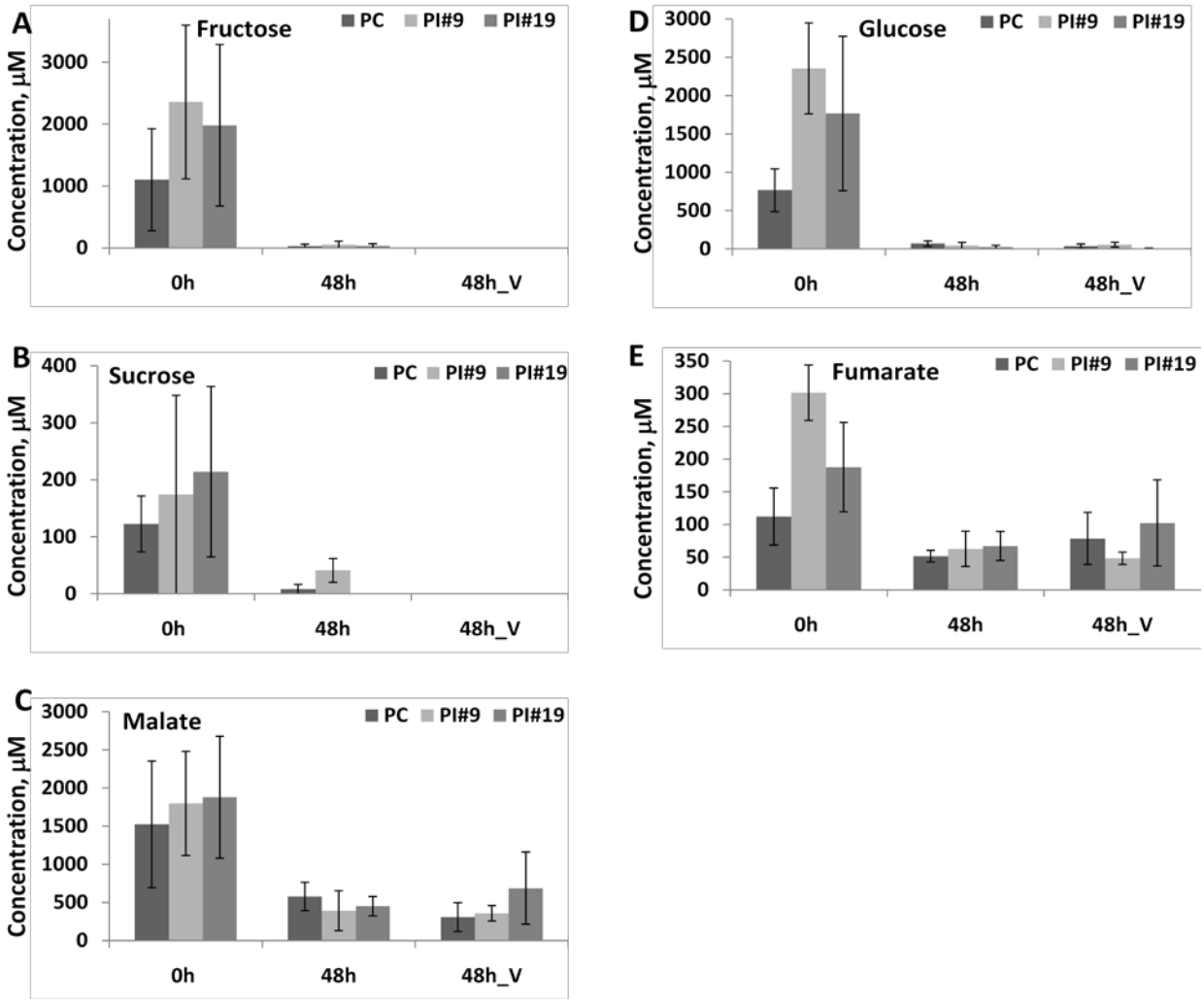


Figure 4.6: the concentrations of sugars and metabolites involved in sugar metabolism in PC and PI plants, as measured by NMR or GC-MS

Figures A-E show the average (with SD) concentrations of fructose, sucrose, malate, glucose and fumarate, respectively in PC, PI #9 and PI #19 plants.

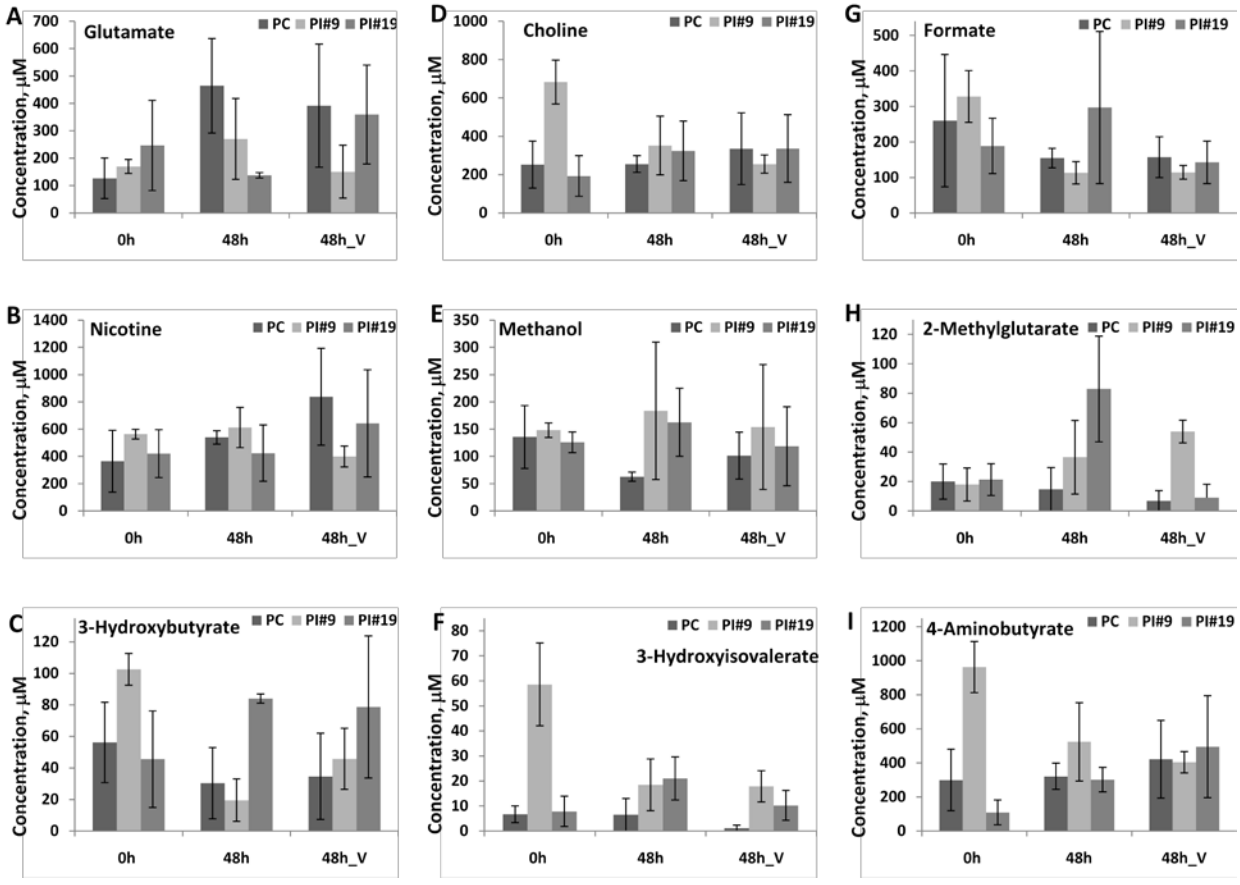


Figure 4.7: The concentrations of other metabolites in PC and PI plants, as measured by NMR or GC_MS

Figures A-I show the average (with SD) level of glutamate, nicotine, 3-hydroxybutyrate, choline, methanol, 3-hydroxyisovalerate, formate, 2-methylglutarate, 4-aminobutyrate, respectively in PC, PI #9 and PI #19 plants.

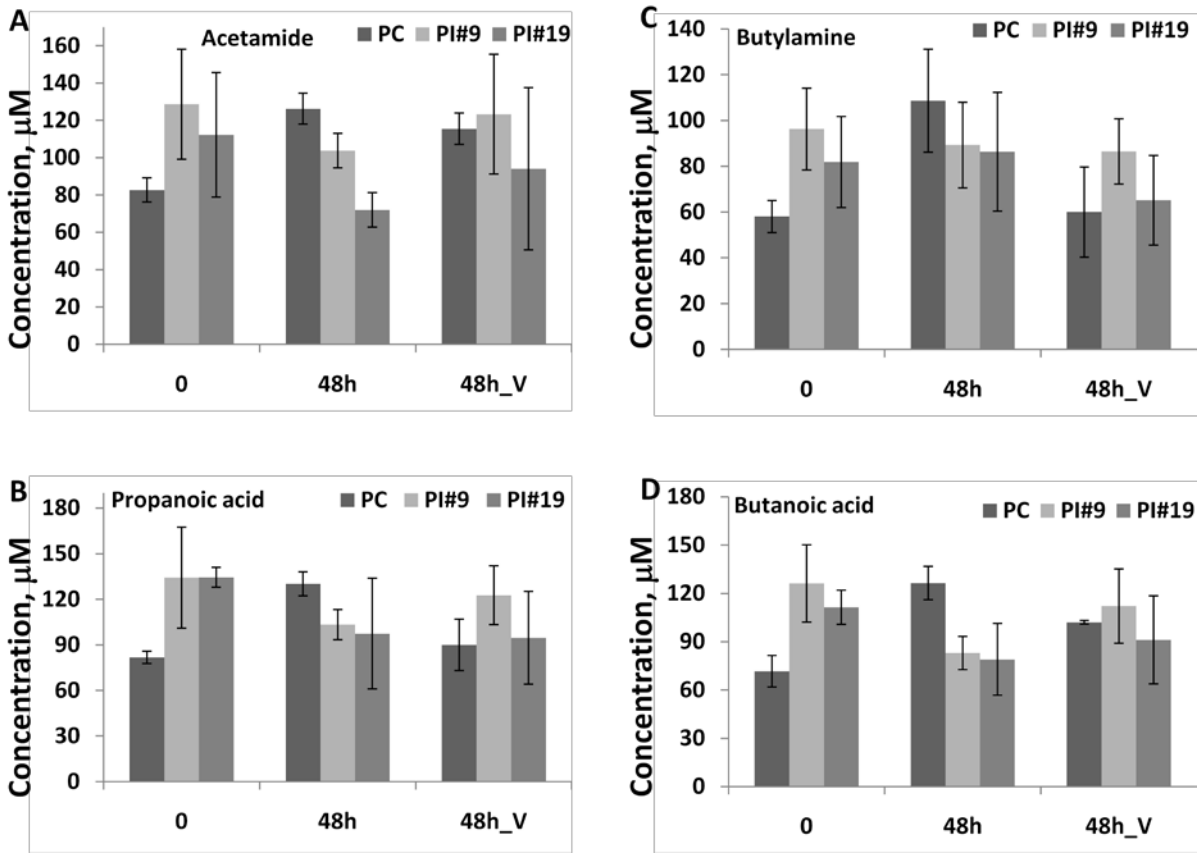


Figure 4.8: The concentrations of other metabolites in PC and PI plants, as measured by NMR or GC-MS

Figures A-D show the average (with SD) level of acetamide, propanoic acid, butylamine and butanoic acid, respectively in PC, PI #9 and PI #19 plants.

5.0 Development of novel reporter gene-based constructs for the study of sequence-specific recombination rate

5.1 Abstract

Homologous recombination is crucial for repair of DNA double-strand breaks in plants. The rate of recombination in plants can be monitored using reporter gene-based recombination constructs. These constructs are efficient indicators of recombination and have been used for a wide variety of scientific studies. However, a construct which can provide information on the recombination rate in specific plant DNA sequences was still not developed. Recently, homologous recombination has been associated with response of plants to different stresses. In one such study, a higher rate of somatic recombination was observed upon infection with a compatible pathogen. The progeny of infected plants exhibit lower levels of DNA methylation in the LRR regions of plant resistance (*R*) genes paralleled by an increased rate of rearrangements in these regions. It can be hypothesized that pathogen infection leads to a higher rate of recombination in the LRR regions as compared to that in the NBS regions or housekeeping genes. Further analysis of the hypothesis would require analysis of the recombination rate in specific sequences which was not feasible using available techniques. In this study, we have developed a novel concept of reporter gene-based sequence-specific recombination constructs for studying the recombination rate in the LRR, NBS regions of *R* genes and in the random regions of housekeeping genes. Transgenic *Arabidopsis thaliana* plants carrying any of these constructs indeed allow the analysis of recombination events in plant tissue.

5.2 Introduction

Living organisms face challenges during their growth period in the form of various biotic and abiotic stresses. Exposure to these stresses results in various metabolic changes, direct or indirect damage to DNA, activation of various molecular stress response pathways, and inheritance of genetic and epigenetic changes in progeny. Stresses like radiation, genotoxic chemicals and other abiotic factors are major causes of DNA double-strand breaks (DSBs). DNA strand breaks represent one of the most lethal DNA lesions, and thus they need to be dealt with promptly (Molinier et al., 2006; Ries et al., 2000).

Two major pathways have been evolved which are responsible for the repair of DNA DSBs. The non-homologous end joining (NHEJ) pathway is responsible for the repair of the majority of DNA DSBs in plants. Such approach is based on the joining of broken DNA strands independent of the presence of homologous sequences. As a result, NHEJ is less energy consuming but more error-prone than its counterpart – homologous recombination (HR). The proteins mediating HR-based repair entail identification and the use of homologous sequences serving as repair templates. The formation of Holliday junction is an intermediate step in the repair process during which the information from the homologous sequence is copied onto damaged DNA strands. The end result is an error-free, efficient repair with exclusion of mutations (Puchta, 2005).

In addition to maintaining genomic stability, HR has been suggested to be a significant benefactor for creating higher variability in the genome. The main mechanism responsible

for HR-dependent variability is crossing-over during meiosis. Hence, HR functions dually by balancing between the maintenance of genome integrity and the creation of variability (Klein, 1995; Schuermann et al., 2005). Recent studies have suggested yet another aspect of HR. It has been observed that the rate of HR is up-regulated by many biotic and abiotic stresses such as UV radiation, salinity, draught, heavy metals and pathogens (Boyko et al., 2010; Boyko et al., 2007; Kovalchuk et al., 2003; Molinier et al., 2006; Ries et al., 2000). The exact biological significance of induced HR rates upon exposure to different stresses is still under investigation.

The majority of previously reported studies have relied on transgenic marker-based homologous recombination substrates such as β-glucuronidase or luciferase. A transgene consists of a recombination substrate sequence cloned in the form of two truncated copies of a reporter gene. The truncated copies share a significant overlap in the homologous sequences which are utilized by HR to repair DSBs in the overlapping region. The result of DSB repair in the overlapping region is the creation of a functional reporter gene. The rate of HR is then compared between plants exposed to different stresses by calculating the number of sectors of the reporter gene (Kathiria and Kovalchuk, 2010).

Using the aforementioned concept, many variations of HR constructs have been developed. One of the initial constructs created consists of direct repeats of truncated copies of the GUS reporter gene. This version of the reporter gene could be used for analysis of intra and inter-molecular recombination events as the recombination events are possible between regions of homology within the transgenic locus as well as between sister chromatids or homologous chromosomes (Swoboda et al., 1994). In contrast, constructs that carry inverted repeats of

the transgene do not allow intra-molecular recombination events to occur, and thus, they can only be used for analysis of inter-molecular recombination events such as inter-allelic and meiotic crossing-over events (Baur et al., 1990; de Groot et al., 1992; Jelesko et al., 1999; Swoboda et al., 1994).

Recombination transgenes developed till now depict the rate of recombination in reporter genes; hence they provide no information about the recombination rate in specific sequences from the plant genome. Studies requiring such sequence-specific information are hampered by the lack of sequence-specific recombination constructs in the field (Kathiria and Kovalchuk, 2010). These studies showed that tobacco plants exposed to a compatible pathogen, tobacco mosaic virus, respond to infection by generating a systemic signal. The signal leads to increased somatic and meiotic recombination frequencies in infected plants (Kovalchuk et al., 2003). Subsequent analysis of the progeny of pathogen-infected plants showed epigenetic changes in these plants (Boyko et al., 2007). Notably, levels of global genome DNA methylation in the progeny of infected plants were significantly higher compared to control plants. At the same time, the level of methylation in the leucine rich repeats (LRR) regions of genomic areas carrying homology to plant resistance (*R*) genes has been found to be decreased. These alterations were paralleled by an increased rate of rearrangements in these LRR regions (Boyko et al., 2007).

Plant *R* genes confer efficient resistance against pathogens. In plants, if there is an R gene present against a particular pathogen, localized cell death in infected tissues is induced which helps limiting a systemic spread of a pathogen in a host plant. Such a plant-pathogen interaction is termed incompatible interaction. In contrast, the absence of an *R* gene results in

the spread of a pathogen and symptom development known as compatible plant pathogen interactions (Shirasu and Schulze-Lefert, 2000). The majority of *R* genes consist of several well-characterized functional regions, including NBS and LRR regions (Hammond-Kosack and Jones, 1997). The NBS region is involved in signal transduction during various cellular processes, while the LRR region provides the specific ability of binding to cellular molecules. In R genes, the LRR region has been theorized to be a determinant of specific pathogen recognition. LRR regions are highly variable in DNA sequence composition in contrast to NBS regions which are highly conserved. Even single nucleotide changes are able to switch pathogen specificity of an *R* gene from one pathogen to a completely different pathogen (Farnham and Baulcombe, 2006; Parniske et al., 1997).

Based on the abovementioned findings by Boyko et al. (2007), a hypothesis can be developed that an increased rate of rearrangements can lead to higher rates of recombination in LRR regions. As LRR regions are responsible for pathogen recognition, the development of novel recombinant LRR regions may result in recognizing a previously compatible pathogen. Hence, these epigenetic alterations may be the main cause of rapid evolution observed in LRR regions. The fact that the level of DNA methylation which is inversely correlated with recombination in a DNA sequence also decreases in the LRR regions of R genes in the progeny of infected plants further strengthens the hypothesis.

Unfortunately, direct analysis of recombination in LRR regions of pathogen-infected plants is not feasible. The frequency of rare event occurrence in somatic tissue prohibits the comprehensive analysis. The analysis of changes in the progeny using Restricted fragment length polymorphism (RFLP) is more feasible, although it still requires a large number of

Southern blots (Boyko et al., 2007) and the information about the nature of rearrangements is limited.

To test the hypothesis whether pathogen infection leads to more frequent recombination in the LRR regions of resistant genes, we designed several reporter gene-based recombination constructs. One of these constructs carried sequences of the LRR regions of RPS5, RPP5 and N gene fused to luciferase. RPS5, RPP5 and N gene are well-known *R* genes conferring resistance to *Pseudomonas syringae* (bacterium), *Hyaloperonospora parasitica* (fungus) and Tobacco mosaic virus (virus), respectively. Another reporter construct carried sequences of the NBS regions of these genes, yet another carried sequences of housekeeping genes - actin, ubiquitin and RENT. Thus, these reporter genes should allow the analysis of recombination events in plants in a sequence-specific manner. We hypothesized that infection with pathogen will preferably increase the recombination frequency in constructs carrying the LRR sequences. We have regenerated several transgenic Arabidopsis lines carrying each of the aforementioned constructs and found that recombination events were detected by using each construct. It still remains to be established whether exposure to pathogens increases the recombination rate more efficiently in the LRR-based constructs.

5.3 Materials and methods

5.3.1 Plant growth conditions

Plants were grown in growth chambers under long-day conditions (16:8 light: dark). The intensity of light was approximately 13,000 lux. Growth chambers were set at 22 °C temperature unless stated otherwise. The humidity in the chambers was maintained at 65% constantly. Plants were grown in 2.5 X 2.5 inches pots, one plant per pot. Generic black soil mixed with vermiculite (3:1) was used for the purpose. Plants were fertilized with 50 mg of a generic fertilizer mix at 2 to 6 weeks of age.

5.3.2 Plant screening for antibiotic resistance

Screening for the stable transformation and segregation ratio was carried out with the help of hygromycin resistance gene. Hence, MS medium containing 30 mg/L hygromycin was used as plant growth medium. Either 8.5 or 14 cm plastic Petri dishes were used. Seeds were sterilized with the help of 1% sodium hypochlorite and plated on media. The plates were incubated at 4°C for 48 hours. Following stratification, the plates were moved to a growth chamber. The chambers were maintained at 22°C temperature and light intensity of 10,000 lux under long day conditions (16:8 light: dark). Seedlings were allowed to grow for 2 to 3 weeks until the difference between hygromycin resistant and sensitive seedlings can be determined. Resistant plants were transferred to soil.

5.3.3 Construct design

A conceptual design of constructs was an utmost important factor for the study. After a substantial review on the subject, a reporter gene-based construct was designed. The construct consisted of three different LRR regions from three different *R* genes in tandem repeats. This will provide an increased probability of identifying recombination events in somatic tissue. A second copy of the same LRR regions will be present at 3' region (Figure 5.1). The 5' LRR repeats will be fused with a reporter gene, while the second copy of the repeats will be preceded by the Cauliflower mosaic virus (CaMV) 35S promoter region. Hence, the two overlapping LRR regions would be in reverse orientation. Only an interchromosomal recombination within the LRR regions would result in expression of a reporter gene.

5.3.4 Selection of sequences

For isolating LRR regions, three different R genes were chosen. *RPS5*, *RPP5* and *N* genes are known *R* genes, and their molecular functioning have been studied. *RPS5* provides resistance to *Pseudomonas syringae*, while *RPP5* can efficiently resist *Hyaloperonospora parasitica,* and N gene makes plants resistant against *TMV*. These *R* genes impart resistance against three classes of pathogens and are from two plant species, Arabidopsis (*RPS5* and *RPP5*) and tobacco (*N* gene).

For selection of LRR sequences, the three R genes were searched for conserved regions. The Pfam software (Sanger institute, UK) was used for the purpose. DNA sequences concomitant with the conserved LRR regions were identified and isolated. For uniformity, equal sizes of LRR regions were selected. Each LRR region was 276 bps in length. Similarly, NBS regions from the three R genes were also isolated.

For analysis of recombination in LRR and NBS regions, a control sequence was absolute necessity. Three housekeeping genes were selected for the purpose – actin, ubiquitin and RENT.

A firefly luciferase gene was selected as a reporter gene for the study. The reporter gene was designed to be a fusion gene. The fusion gene consisted of a 5' located Hemagglutinin (HA) tag, a central Kanamycin resistance (*hpt II*) gene, and a 3' positioned luciferase gene (figure 5.1). The HA tag provides a high specific binding ability, while the *hpt II* gene provides resistance to kanamycin and can be used as a selective agent in tissue culture. Luciferase expression can be visualized *in vivo* or can be assayed *in vitro*. For efficient functioning of kanamycin and luciferase genes, a sequence of 8-weak amino acids was introduced between them. We expected the luciferase protein to be cleaved from the rest of the protein upon recombination in the substrate.

5.3.5 Artificial synthesis

For cloning, sequences were assembled in required configuration using Vector NTI software® (Invitrogen, USA). Three different types of assemblies were organized from each of LRR, NBS and housekeeping sequences. The first assembly consisted of 5' repeat

regions. The second assembly included a reporter gene cassette, and the last assembly had 3' repeats. The downstream cloning following synthesis was carried by the Gateway®-based cloning system. Appropriate attB1, attB2, attB3 and attB4 sequences were included in the flanking region of the assemblies. The sequences were sent to CODON DEVICES Inc. (USA) for artificial synthesis. The assemblies were received in the form of pUC19 inserts. The sequences were confirmed by sequencing of pUC19 plasmid inserts.

5.3.6 The Gateway-based cloning system

The Gateway®-based cloning system relies on sequence specific recombinases. Four sequences are recognized by the recombinases – attB, attP, attL and attR. The recombination between attB and attP sequences, when mediated by BP Clonase results in an attL sequence. Similarly, the recombination between attL and attR sequences creates an attB sequence. This process is utilized in cloning projects.

The assemblies containing LRR, NBS and housekeeping genes were received in pUC19 vectors. The assemblies were transferred to Gateyway® pDONR vectors using BP reaction (Figure 5.2). Briefly, 50 fmoles of pUC19 and pDONR plasmids were incubated in the presence of BP Clonase for overnight at room temperature. The amount of plasmid was calculated using the following equation:

$ng = (x\ fmoles)(N)(660fg/fmoles)(1ng/10^6\ fg)$

Where x is the number of fmoles required and N is the size of DNA in bp.

The BP reaction was introduced into *E. coli* (TOP10, Invitrogen) cells, and colonies were obtained by plating the cells on LB medium under 50 mg/L kanamycin antibiotic. Insertion of appropriate fragments in pDONR vectors was confirmed using restriction analysis.

In the second phase of cloning, the assemblies inserted into pDONR vectors were transferred to a binary vector pHM43GW. LR reactions were carried out for the purpose (Figure 5.2). 10 fmoles of pDONR vectors and 20 fmoles of pHM43GW plasmids were incubated with LR Clonase at room temperature overnight. The reactions were inserted into TOP10 cells and colonies were obtained under 50 mg/L spectinomycin selection. The LR reaction puts together the three fragments and results in the creation of the final recombination construct. Clones from the LR reactions were verfied using restriction analysis. Bands of expected size were obtained from each of the LRR, NBS and housekeeping sequences containing binary vectors (supplementary figure 5.1).

5.3.7 Plant transformation by the Agrobacterium-mediated gene delivery system

The binary vectors, pHM43GW-LRR, pHM43GW-NBS and pHM43GW-HK, were transferred to *Agrobacterium tumefaciens* (strain GV3101) via electroporation. Single colonies were obtained from yeast extract-peptone (YEP) medium plates in the presence of 25 mg/L rifampicin, 25 mg/L gentamycin and 50 mg/L spectinomycin. These single colonies were used to prepare a liquid overnight culture in YEP medium containing antibiotics (rifampicin 25 mg/L, gentamycin 25 mg/L and spectinomycin 50 mg/L).

For transient expression analysis, overnight cultures were obtained from the LRR, NBS and housekeeping binary vectors as described above. The cultures were centrifuged, and pellets of bacteria were resuspended in IM. Leaves of 6-week-old *Nicotiana benthamiana* plants were infiltrated with bacterial cultures. The plants were incubated at 22°C for 2 days. The leaves were analyzed using either CCD analysis or immunohistochemistry (Figure 5.3 and 5.4). Agrobacterium without any binary vector was used as a negative control.

For Arabidopsis transformation, the floral dip method was performed (Clough and Bent, 1998). Arabidopsis (ecotype Columbia) plants were grown in a pot at 22°C, and then initial shoots were removed to induce an increase in the number of secondary shoots. These plants were grown until the flowering stage. A fresh overnight culture of *A. tumefaciens* was harvested and resuspended in inoculation medium (IM). Floral buds were immersed in the solution for 5 min and then transferred to high-humidity conditions. Seeds were harvested from the plants and screened for transformation events in the presence of 25 mg/L hygromycin. Positive transformants were raised to obtain a homozygous transgenic plant. An expression pattern of luciferase was analyzed in stable transformants, and lines with appropriate expression patterns were selected for further multiplication (Figure 5.5 and 5.6).

5.3.8 CCD analysis

Luciferase expression was analyzed using a CCD camera (Gloor Instruments, Basel, Switzerland). Such camera consists of a charged coupled device (CCD) and is able to record

very faint light signals. For the analysis, the 0.5 mM luciferin solution was sprayed uniformly over the plant surface. The plants were incubated in dark for 20 minutes at room temperature for removal of any autoluminescence. The plants were placed in the CCD camera chamber, and the camera was fine-focused on plants. Images were acquired using Pixcel software (PerkinElmer, Cambridge, UK). First, a control image was generated in the presence of external light by exposing the plants for 100 µseconds. Subsequently, an external light source was turned off, and plants were exposed for 5 minutes in complete darkness. Both control and dark images are in black and white format. The images were exported to the ANAlysis software (Soft Imaging Systems, Munster, Germany) in TIFF format for post-processing. Using the ANAlysis software, the dark image was given a red pseudo color and superimposed on the control image. The final images obtained were saved in JPEG format for further evaluation (Figures 5.3, 5.5 and 5.6).

5.3.9 Direct whole-mount immunohistochemistry

Direct whole-mount immunohistochemical analysis was performed to check the expression of a Hemagglutinin (HA) tag sequence during transient expression analysis. For the purpose, *Nicotiana benthamiana* mesophyll tissues infected with *Agrobacterium* were excised from the plant and fixed using a 4% paraformaldehyde solution. The tissues were washed briefly in Phosphate buffer saline (PBS), and blocking was carried out using 10% goat serum in PBS for 4 hours. Anti-HA fluorescein isothiocynate (FITC)-labelled antibodies were diluted to 1:400 in blocking solution. The tissues were incubated overnight in the antibody solution at room temperature. The tissues were further washed using PBS buffer with 0.05% Tween

20 (PBST) solution three times, one hour each. A DNA-specific fluorescent dye, 4, 6-diamidino-2-phenylindole (DAPI), was used as counter stain. The tissues were incubated in 50% glycerol overnight in dark at room temperature. The tissues were mounted in antifade medium, and slides were analyzed using confocal microscopy (Nikon Inc, Japan). Images were obtained in 488 and 546 nm wavelengths and combined together for further analysis (Figure 5.4)

5.3.10 Western blot analysis

Western blot analysis was carried out to confirm the expression of HA-tag and luciferase and to confirm the process of dissociation between luciferase and rest of the fusion protein. The total size of the transgene is approximately 120 kDa. Dissociation would result in generation of the 60 kDa luciferase protein and the 60 kDa protein consisting of the rest of the transgenic protein. Hence, using anti-HA-tag antibodies, two proteins are expected: the whole 120 kDa transgene and the 60 kDa truncated transgene. Antibodies against the luciferase protein can also recognize two proteins: the 120 kDa whole transgene and the 60 kDa luciferase protein.

Three individual transgenic lines, L60 (LRR), L8 (LRR), H399 (housekeeping) and wildtype Arabidopsis were used for the experiment. Tissues were collected by flash-freezing in liquid nitrogen from 4-week-old plants. A fine powder was obtained from the tissues by crushing the tissues in liquid nitrogen. The powder was homogenized in PBS for 15 minutes. The homogenate was centrifuged at 13000 RPM for 15 minutes. The amount of protein was

quantified by the Bradford protein quantification assay using bovine serum albumin as standard. 20 µg of protein from each transgenic line were loaded on 10% polyacrylamide gel. After electrophoresis, the proteins were transferred to polyvinylidene fluoride (PVDF) membranes (Roche applied biosciences, USA). The membranes were blocked in 5% milk in the PBST solution. After blocking, the membranes were either incubated with anti-HA (Abcam Inc, USA) or anti-luciferase (Abcam Inc, USA) antibodies at 1:500 dilution overnight at 4C. Primary antibodies were washed using 0.5% milk in PBST. Secondary antibodies containing Horse radish peroxidase (HRP) conjugates (Abcam Inc, USA) were applied at 1: 10000 times dilution for 2 hours at room temperature. The enhanced chemiluminescence (ECL) method was used for detection of HRP-conjugated secondary antibodies. Briefly, 2.5 mL of ECL reagent (PerkinElmer, USA) was applied to each membrane. After removing excess ECL reagent, the membranes were covered in a transparent plastic wrap and exposed to BioMax scientific films (Kodak Inc, USA). The films were developed with an in-house film developer and used for further analysis (Figure 5.7).

5.3.11 PCR confirmation of recombination events in somatic tissues

It was utmost important to confirm recombination events in the transgenic DNA sequence after stable transformation in Arabidopsis plants. However, a limited amount of the recombined DNA present in plant tissues rule out many of the possible assays. A PCR-based approach was designed to identify the actual recombination in transgenes. PCRs were carried out by using a forward primer from the 35S promoter region (35S f: 5’

GTGGAGCACGACACACTTGT 3') and one of the reverse primers from the reporter gene region (KanLuc r 1: 5' TTGTCAATCAAGGCGTTGGT 3' or KanLuc r 2: 5' TTTCCACCATGATATTCGGC 3'). PCR amplification is possible only if the 35S promoter and the reporter gene are brought together by a recombination event (Figure 5.8). In addition to the primers, Takara exTaq (TAKARA BioInc, Japan) and 1µg of genomic DNA from transgenic lines (L60, L8, H399 and wildtype) were used in the PCR. Initial denaturation was carried at 95˚C for 10 minutes. The actual PCR cycle consisted of a denaturation step at 95˚C for 30 seconds, an annealing step at 55˚C for 35 seconds and an extension step at 72˚C for 2 minutes. The 35 PCR cycles were carried out followed by the final extension step at 72˚C for 5 minutes. Electrophoresis of PCR reactions was performed using ethidium bromide (0/5 µg/ml) containing 1% agarose gel. The bands were visualized on a transilluminator, and the image was acquired for further analysis (Figure 5.9).

5.4 Results

5.4.1 Synthesis of binary vectors

Due to complexity of designed recombination constructs, artificial synthesis of sequences was opted for. The sequences were obtained in the pUC19 vector after artificial synthesis. Sequence analysis of fragments showed that the fragments were exact match to the ordered sequences. Subsequent BP reactions resulted in the inclusion of the fragments in the pDONR vectors. The cloning was efficient and precise as no alterations in the fragments were observed, as checked by restriction analysis. The LR reaction was used to put together all the fragments in the pHM43GW binary vector. This phase of cloning was also precise; hence,

all the planned steps of cloning (as shown in Figure 5.2) were successful. Restriction analysis revealed a proper integration of the fragments in the binary vector in all positive clones checked (Supplementary Figure 5.1).

5.4.2 Transient expression analysis

To analyze the functionality of recombination reporters, a transient assay was performed. The Agrobacterium-mediated gene delivery system was used to deliver plasmids into *Nicotiana benthamiana* plants. Expression of the reporter gene was analyzed either by a CCD camera or by immunohistochemistry.

CCD analysis of infected leaves was carried out to confirm a proper expression of the fusion gene and functioning of the luciferase reporter gene. This analysis revealed that luciferase is expressed from all the three binary vectors during a transient expression experiment. No expression from the negative control was recorded (Figure 5.3).

Immunohistochemical analysis was done to check the expression and detection efficiency of HA-tag. In the negative control, there was some expression recorded in plant vasculature when the FITC range was analyzed. However, no expression was observed in plant mesophyll tissues. In the case of binary vectors, there was a significant expression present in mesophyll tissues (Figure 5.4). The images indicated that the HA-tag sequence was present in these tissues.

5.4.3 Stable expression analysis

Arabidopsis plants were transformed with the binary vectors using the floral dip method of the Agrobacterium-mediated gene transfer system and screened for stable integration of transgenes. Approximately 200 positive transformants from each of the constructs were isolated. These positive transformants were analyzed for the expression pattern of reporter genes originating from transgenes. CCD camera analysis revealed variations in expression patterns (Figure 5.5). Some of the plants showed the expected spot-like appearance of luciferase expression in somatic tissues (Figure 5.5a). However, many plants either showed no expression of luciferase or had an aberrant expression pattern (Figure 5.5b). Plants with undesired expression patterns were removed from the selection process. Plants with a moderate amount of luciferase spots were selected so that any increase or decrease in recombination can be noted (Figure 5.5c). Plants homozygous for the transgene were screened in further generations.

5.4.4 Reporter gene analysis

The homozygous plants were used to study the structure of the protein generated after recombination events. The design of the reporter gene included 8 weak amino acids between kanamycin and luciferase genes. As a result, the luciferase gene should dissociate from the rest of the protein post-translationally. The western blot analysis using anti-HA-tag antibodies revealed a large protein in a size range of 120 kDa present only in line L60 and absent in lines L8, H399 and in the negative control line (figure 5.6a). The western blot

analysis with anti-luciferase antibodies identified a much smaller protein in a size range of 60 kDa (Figure 5.6b). The band was absent in L8, H399 and wildtype samples and present only in L60 samples. This analysis confirmed the process of dissociation of luciferase protein from the rest of the fusion protein.

5.4.5 Analysis of recombination event in the transgene

To check the occurrence of recombination events in transgenes at the level of plant genomic DNA, the PCR-based approach was devised. PCR analysis using 35S forward and KanLuc1 reverse primers showed the presence of a 2.4 kb fragment that was expected to form only after a recombination event. However, the fragments were identified in lines L8 and H339. No fragments were present in line L60 or in negative control samples. Similarly, PCR analysis using the 35S forward primer and KanLuc2 reverse primers resulted in a 2.6 kb fragment. This fragment was present in L8 and H399 samples but absent in L60 and negative control samples, similar to results obtained by 35S-KanLuc1 PCR analysis. Hence, it is possible to identify recombination events in genomic regions of transgenic plants. Also, it confirms the proper functioning of recombination constructs.

5.5 Discussion

Plants are faced with stresses that often lead to DNA double-strand breaks. DNA double-strand breaks are repaired via NHEJ or HR pathways. The HR pathway, although less prominent, is more reliable for DSB repair (Puchta, 2005). The most convenient method for

analysis of HR is the use of reporter gene-based recombination constructs. The rate of recombination can be estimated by counting the number of events of reporter gene expression (Kathiria and Kovalchuk, 2010). Recent studies which used such recombination constructs revealed changes in the rate of HR upon exposure to stresses such as radiation, salinity, drought and pathogens (Boyko et al., 2010; Boyko et al., 2006; Kovalchuk et al., 2003; Molinier et al., 2006; Ries et al., 2000). In one such study, a systemic signal was identified during a compatible plant-pathogen interaction (the absence of R-gene-based recognition of a pathogen) which increases the somatic recombination rate in infected plants (Kovalchuk et al., 2003). Further experiments revealed a decrease in DNA methylation and an increase in the rate of rearrangements in the LRR region of *R* genes in plants (Boyko et al., 2007).

In R genes, different regions are responsible for different functions. The NBS region includes an ATP-binding site. This site is involved in ATP hydrolysis and subsequent signal transduction (Traut, 1994). The LRR region is a key player in avirulence factor recognition (Dangl and Jones, 2001). Various alleles of one specific resistance gene *L* possess specificity against different strains of flax rust pathogen *Melampsora lini* (Ellis et al., 1999). These alleles of the *L* gene have differences only in DNA sequence of LRR regions.

The evolution of *R* genes is yet another interesting phenomenon. Many *R* genes are found in clusters either at one locus or at multiple closely associated loci on the same chromosome (Ronald et al., 1992; Williams et al., 1996). It has been proposed that gene duplication events followed by diversification of these genes are a major reason behind the evolution of *R* gene families (Lawrence et al., 1995; Whitham et al., 1994). Molecular analysis of the *Cf-*

x gene family has shown that gene duplication by unequal crossing over plays an important role in its evolution. Due to the use of unequally aligned sequences for recombination, the number and cluster composition of gene family members can change. This results in an increase in the variation present in the population (Parniske et al., 1997).

The results obtained by Boyko et al. (2007) suggest that recombination in LRR regions might play an important role even in somatic tissues. The rate of rearrangements is increased upon infection with a compatible pathogen (Boyko et al., 2007). Such increase would create high variability in LRR regions. Under pathogen selection pressure, this variability may result in rapid development of novel LRR regions. The LRR regions might be able to recognize previously compatible pathogens. As the gametes in plants are produced from somatic tissues in later part of their life, such changes in somatic tissues of the LRR regions can be transmitted to the next generation.

In order to mitigate the hurdles in the study, a new reporter gene-based sequence-specific recombination construct was designed. The construct consisted of two repeats of either LRR or NBS, or housekeeping genes. The LRR and NBS sequences were obtained from three different R genes – RPS5, RPP5 and N gene, while the housekeeping sequences were derived from actin, ubiquitin and RENT genetic sequences. RPS5, RPP5 and N genes are well-studied R genes conferring resistance against *Pseudomonas syringae*, *Peronospora parasitica* and *TMV* (Ade et al., 2007; Erickson et al., 1999; Parker et al., 1997). R genes conferring resistance to bacterial, fungal and viral pathogens were selected deliberately in order to allow a broader analysis of the influence of the type of pathogen and plant species on DNA rearrangements triggered by the infection process.

The analysis of transient expression showed that *Nicotiana benthamiana* plants infected with either one of three recombination constructs, LRR, NBS and HK, showed the expression of the reporter gene (Figure 5.3). We noticed that the expression level of the HK construct was the highest, and the expression level of the LRR construct was the lowest. Difference in the expression levels between the constructs can be attributed to either differential infection rate or differences in recombination in different regions. However, either of the possibilities cannot be concluded based on only these results.

The expression and detection ability of HA-tag peptide was also checked during transient expression analysis. Immunohistochemistry was used for the purpose. The results obtained indicated that the expression of the HA tag was present in LRR, NBS and HK constructs in mesophyll cells of infected leaves (Figure 5.4). Hence, not only the presence of a reporter sequence but also the possibility of using the HA tag for assays was confirmed. Hence, the results of transient expression analysis showed that recombination constructs do result in the expression of the reporter gene. Also, it showed that luciferase and HA-tag peptide sequences formed an appropriate confirmation to be detected by specific antibodies. Since the expression of the reporter gene from extrachromosomally located DNA is only possible upon a recombination event, this transient assay showed that recombination events in constructs are indeed possible.

Once the expression of the reporter gene from the recombination construct was confirmed, the binary vectors were used to stably integrate transgenes in *Arabidopsis* plants. Positive transformants were selected by screening in the presence of hygromycin. Patterns of luciferase gene expression were checked by CCD camera analysis. There were variations in

expression patterns in the first generation of transformants. Some plants showed a spot-like expression pattern of luciferase as expected (Figure 5.5a); while some plants either had no expression or revealed an unexpected expression pattern (Figure 5.5b). The unexpected or aberrant expression can be a result of the influence of ectopic sequences on the transgene, as it was observed when a functional copy of reporter genes was inserted into the plant genome (Alvarado et al., 2004; Koncz et al., 1989). Such lines, except line L60 (the LRR construct), were removed from the further selection procedure. Homozygous plants were obtained from the transformants which possessed a moderate number of luciferase spots (Figure 5.5c)

Homozygous transgenic plants were used for further experiments which inquired into functioning of recombination constructs. Western blot analysis was used for checking the reporter gene expression. Three different lines were used for the experiment. Line L8 and H399 showed a moderate and low level of luciferase expression in the form of spots. Line 60 had a very high expression which was not limited to precise spots (Figure 5.6). Arabidopsis wild-type was used as a negative control. The reporter gene expression was checked by using antibodies against HA-tag and luciferase peptides. Since our construct included weak amino acids placed between kanamycin and luciferase genes, we hypothesized that by using anti-luciferase antibodies, we would be able to detect both type of proteins: one representing an entire fusion protein consisting of HA-tag, kanamycin and luciferase proteins having a size of approximately 120 kDa and another one consisting of only luciferase of 60 kDa in size. The western blot using anti-HA-tag antibodies showed the presence of 120 kDa protein, detected only in line L60. When an anti-luciferase antibody was used, a 60 kDa protein was detected in line L60 (Figure 5.7). The size of the whole protein is expected to be 120 kDa. A split in the reporter gene would result in a 60 kDa truncated reporter protein and a 60 kDa

luciferase protein. Using anti-HA-tag antibodies, one has a chance to detect either the presence of a 120 kDa protein if no dissociation occurred or a 60 kDa protein if luciferase protein dissociated fully or both the 60 and 120 kDa proteins if partial dissociation occurred. The fact that we observed only the 120 kDa protein could suggest that no dissociation of the luciferase protein occurred. However, since anti-luciferase antibodies detected only the 60 kDa protein, it might be suggested that anti-luciferase antibodies could not detect the presence of a larger protein; similarly, it is possible that an anti-HA-tag antibody could not detect a smaller protein. This is not really surprising; the confirmation of peptide sequences in the initial whole protein and subsequent peptide after the split is expected to be different. Hence, it is possible that the antibodies used in western blot analysis were able to detect only one confirmation, and thus only one band was visible. This explains the fact why the whole reporter protein is detected in HA-tag western blot and the dissociated luciferase protein alone is present in luciferase western blot. Apart from this fact, the western blot show that the dissociation of the luciferase from the rest of the reporter gene does occur, but it is not present in all reporter proteins produced. The absence of HA-tag and luciferase bands in lines L8 and H399 can be due to a lower frequency of recombination events and thus lower expression of the reporter gene in somatic tissue that are below detection range under current assay conditions.

The transgenic lines which possess the recombination construct did express the reporter gene. However, it was important to confirm that the expression was derived from a recombination event in repeats and not as a result of ectopic expression of the reporter gene. The cell which contains the recombinant transgene would be rare; hence, it might be hard to detect its expression. A PCR-based approach was devised to check the presence of

recombination events. The approach used a forward primer originating in the 35S promoter region and a reverse primer – in the reporter gene region. Amplification by such PCR reaction will result only if a recombination event has taken place in the repeats (Figure 5.8). Two different reverse primers, KanLuc1 and KanLuc2, were used in conjugation with the 35S forward primer. Lines L60, L8 and H399 were used for the experiment in addition to a wild-type sample. Line L60 was selected because it showed high-level expression of reporter genes, but it was suspected to have ectopic expression as it did not show a spot-like pattern. The results from PCR reactions indicate that the expected 2.4 kb fragment was observed in 35S-KanLuc1 PCR in lines L8 and H399. No fragments were observed either in L60 or in negative control samples. Similarly, 35S-KanLuc2 PCR revealed the expected 2.6 kb fragment in L6 and H399 samples but not in L60 and negative control samples. Therefore, it can be concluded that recombination indeed occurs in the repeats that are stably integrated in the plant genome. Hence, recombination constructs can be utilized to study the rate of recombination in the sequence of interest. Also, no bands were observed in line L60, although it showed high-level expression of the reporter gene using a CCD camera and western blot analysis. Hence, the reporter gene expression was a result of ectopic expression rather than recombination events. Thus, line L60 should not be used for recombination analysis.

In future experiments, effects of various stresses on the rate of recombination in transgenic lines containing recombination constructs will be analyzed. Also, changes in epigenetic status, including DNA methylation and histone modification changes in the repeat regions of recombination constructs will also be analyzed. Such sequence-specific recombination constructs can also be designed for various studies if conventional approaches cannot provide adequate analysis.

5.6 References:

Ade, J., DeYoung, B.J., Golstein, C., and Innes, R.W. (2007). Indirect activation of a plant nucleotide binding site–leucine-rich repeat protein by a bacterial protease. Proceedings of the National Academy of Sciences *104*, 2531-2536.

Alvarado, M.C., Zsigmond, L.M., Kovacs, I., Cseplo, A., Koncz, C., and Szabados, L.M. (2004). Gene Trapping with Firefly Luciferase in Arabidopsis. Tagging of Stress-Responsive Genes. Plant Physiol *134*, 18-27.

Baur, M., Potrykus, I., and Paszkowski, J. (1990). Intermolecular homologous recombination in plants. Mol Cell Biol *10*, 492-500.

Boyko, A., Blevins, T., Yao, Y., Golubov, A., Bilichak, A., Ilnytskyy, Y., Hollander, J., Meins, F., Jr., and Kovalchuk, I. (2010). Transgenerational adaptation of Arabidopsis to stress requires DNA methylation and the function of Dicer-like proteins. PLoS One *5*, e9514.

Boyko, A., Greer, M., and Kovalchuk, I. (2006). Acute exposure to UVB has a more profound effect on plant genome stability than chronic exposure. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis *602*, 100-109.

Boyko, A., Kathiria, P., Zemp, F.J., Yao, Y., Pogribny, I., and Kovalchuk, I. (2007). Transgenerational changes in the genome stability and methylation in pathogen-infected plants: (virus-induced plant genome instability). Nucleic Acids Res *35*, 1714-1725.

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

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

de Groot, M.J., Offringa, R., Does, M.P., Hooykaas, P.J., and van den Elzen, P.J. (1992). Mechanisms of intermolecular homologous recombination in plants as studied with single- and double-stranded DNA molecules. Nucleic Acids Res *20*, 2785-2794.

Ellis, J.G., Lawrence, G.J., Luck, J.E., and Dodds, P.N. (1999). Identification of regions in alleles of the flax rust resistance gene *L* that determine differences in gene-for-gene specificity. Plant Cell *11*, 495-506.

Erickson, F.L., Dinesh-Kumar, S.P., Holzberg, S., Ustach, C.V., Dutton, M., Handley, V., Corr, C., and Baker, B.J. (1999). Interactions between tobacco mosaic virus and the tobacco *N* gene. Philos Trans R Soc Lond B Biol Sci *354*, 653-658.

Farnham, G., and Baulcombe, D.C. (2006). Artificial evolution extends the spectrum of viruses that are targeted by a disease-resistance gene from potato. Proc Natl Acad Sci U S A *103*, 18828-18833.

Hammond-Kosack, K.E., and Jones, J.D. (1997). Plant Disease Resistance Genes. Annu Rev Plant Physiol Plant Mol Biol *48*, 575-607.

Jelesko, J.G., Harper, R., Furuya, M., and Gruissem, W. (1999). Rare germinal unequal crossing-over leading to recombinant gene formation and gene duplication in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A *96*, 10302-10307.

Kathiria, P., and Kovalchuk, I. (2010). Reporter gene-based recombination lines for studies of genome stability. Methods Mol Biol *631*, 243-252.

Klein, H.L. (1995). Genetic control of intrachromosomal recombination. Bioessays *17*, 147-159.

Koncz, C., Martini, N., Mayerhofer, R., Koncz-Kalman, Z., Korber, H., Redei, G.P., and Schell, J. (1989). High-frequency T-DNA-mediated gene tagging in plants. Proc Natl Acad Sci U S A *86*, 8467-8471.

Kovalchuk, I., Kovalchuk, O., Kalck, V., Boyko, V., Filkowski, J., Heinlein, M., and Hohn, B. (2003). Pathogen-induced systemic plant signal triggers DNA rearrangements. Nature *423*, 760-762.

Lawrence, G.J., Finnegan, E.J., Ayliffe, M.A., and Ellis, J.G. (1995). The *L6* gene for flax rust resistance is related to the Arabidopsis bacterial resistance gene *RPS2* and the tobacco viral resistance gene *N*. Plant Cell *7*, 1195-1206.

Molinier, J., Ries, G., Zipfel, C., and Hohn, B. (2006). Transgeneration memory of stress in plants. Nature *442*, 1046-1049.

Parker, J.E., Coleman, M.J., Szabo, V., Frost, L.N., Schmidt, R., van der Biezen, E.A., Moores, T., Dean, C., Daniels, M.J., and Jones, J.D. (1997). The Arabidopsis downy mildew resistance gene *RPP5* shares similarity to the toll and interleukin-1 receptors with *N* and *L6*. Plant Cell *9*, 879-894.

Parniske, M., Hammond-Kosack, K.E., Golstein, C., Thomas, C.M., Jones, D.A., Harrison, K., Wulff, B.B., and Jones, J.D. (1997). Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the Cf-4/9 locus of tomato. Cell *91*, 821-832.

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

Ries, G., Heller, W., Puchta, H., Sandermann, H., Seidlitz, H.K., and Hohn, B. (2000). Elevated UV-B radiation reduces genome stability in plants. Nature *406*, 98-101.

Ronald, P.C., Albano, B., Tabien, R., Abenes, L., Wu, K.-s., McCouch, S., and Tanksley, S.D. (1992). Genetic and physical analysis of the rice bacterial blight disease resistance locus, Xa21. Molecular and General Genetics MGG *236*, 113-120.

Schuermann, D., Molinier, J., Fritsch, O., and Hohn, B. (2005). The dual nature of homologous recombination in plants. Trends in Genetics *21*, 172-181.

Shirasu, K., and Schulze-Lefert, P. (2000). Regulators of cell death in disease resistance. Plant Mol Biol *44*, 371-385.

Swoboda, P., Gal, S., Hohn, B., and Puchta, H. (1994). Intrachromosomal homologous recombination in whole plants. EMBO J *13*, 484-489.

Traut, T.W. (1994). The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites. European Journal of Biochemistry *222*, 9-19.

Whitham, S., Dinesh-Kumar, S.P., Choi, D., Hehl, R., Corr, C., and Baker, B. (1994). The product of the tobacco mosaic virus resistance gene *N*: similarity to toll and the interleukin-1 receptor. Cell *78*, 1101-1115.

Williams, C.E., Wang, B., Holsten, T.E., Scambray, J., Silva, F., and Ronald, P.C. (1996). Markers for selection of the rice *Xa21* disease resistance gene. TAG Theoretical and Applied Genetics *93*, 1119-1122.

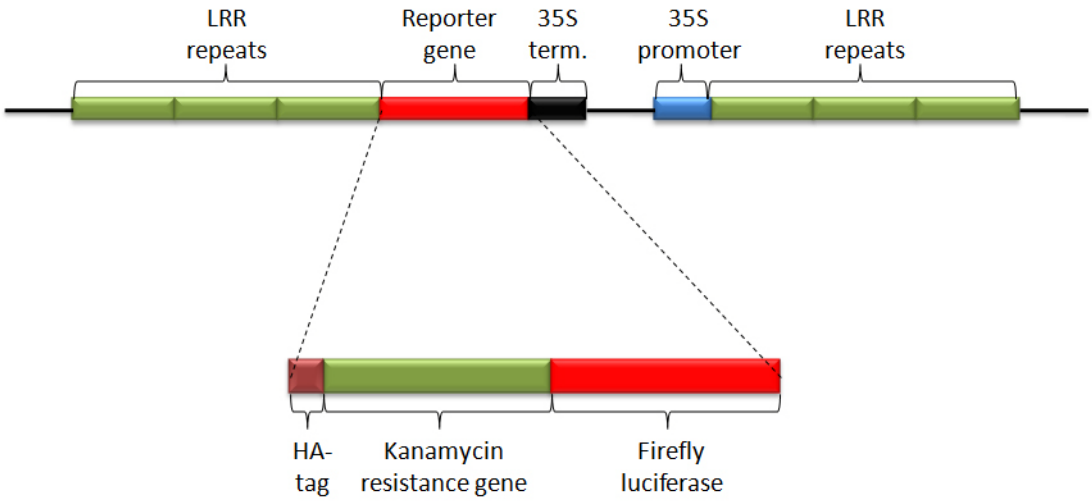


Figure 5.1: A schematic diagram representing the design of a sequence-specific recombination construct.

The reporter gene-based recombination construct consists of two LRR repeats, one – at the 5' end which is followed by a reporter gene, and the second 3' repeat which is preceded by a constitutive (35S) promoter. The reporter gene is a fusion gene consisting of the HA-tag, the kanamycin resistance gene and the firefly luciferase gene.

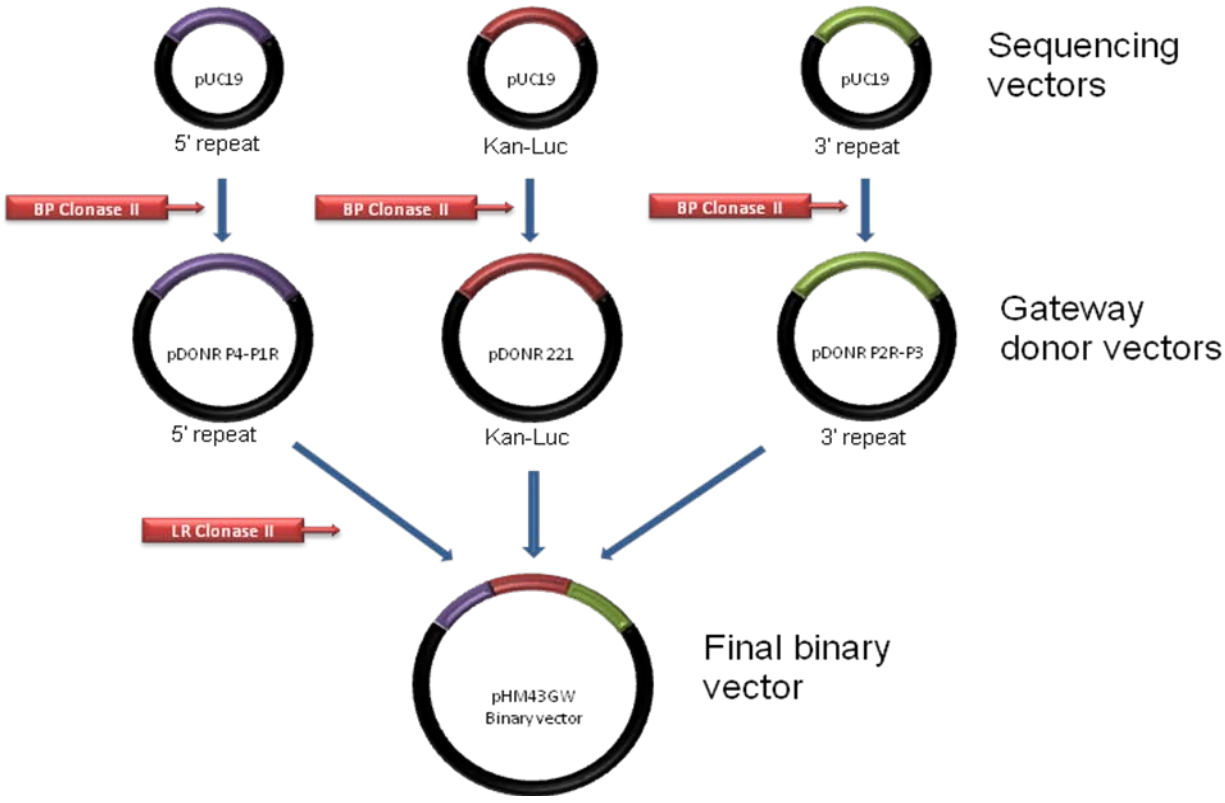


Figure 5.2: An overview of various cloning stages.

Newly synthesized sequences were obtained in pUC19 vectors. The sequences were transferred to Gateway donor vectors using a BP reaction. Three different fragments were assembled in a binary vector pHM43GW using a LR reaction. The binary vector was used for both transient and stable transformation experiments.

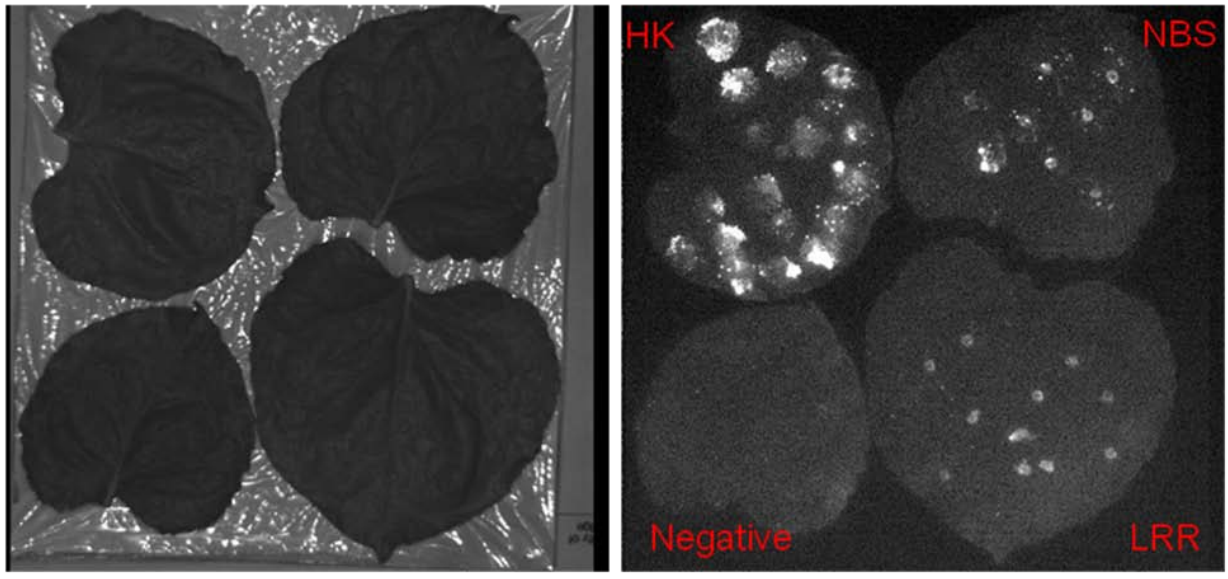
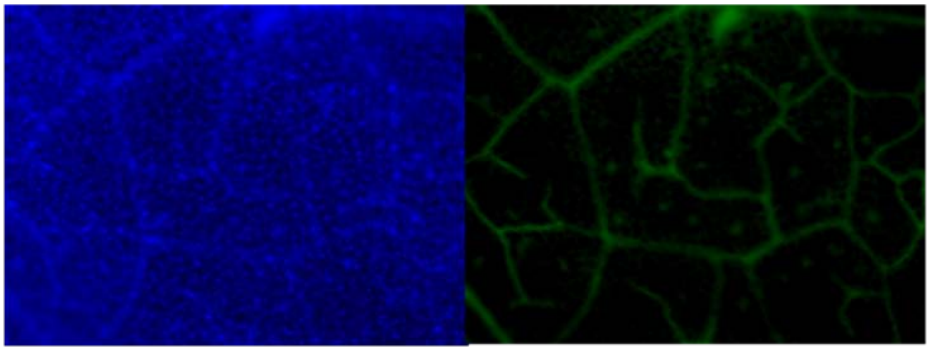
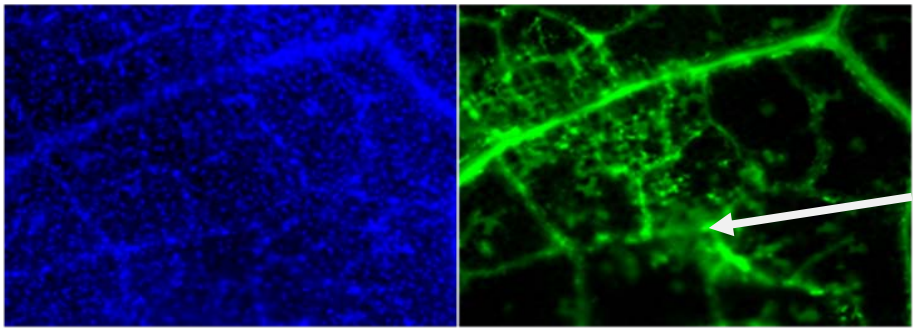


Figure 5.3: Transient expression analysis using a CCD camera.

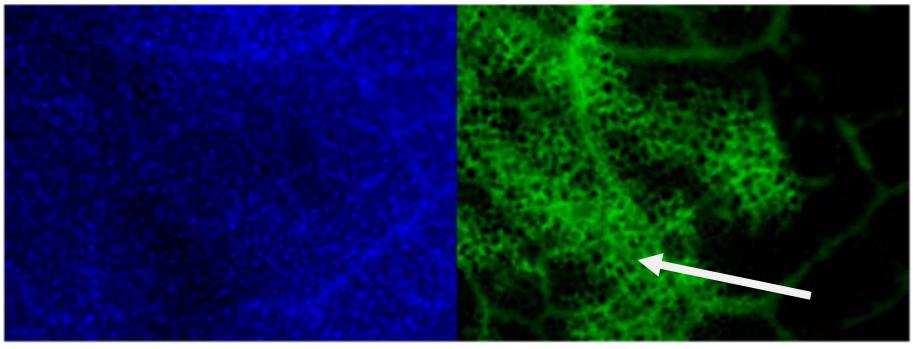
Nicotiana benthamiana plants were infiltrated with Agrobacterium containing binary vectors that consist of LRR, NBS or housekeeping (HK) constructs. Luciferase expression was checked after 2 days post infection. The image obtained in the absence of light indicated that luciferase expression was present in all three binary vectors but not in the negative control.



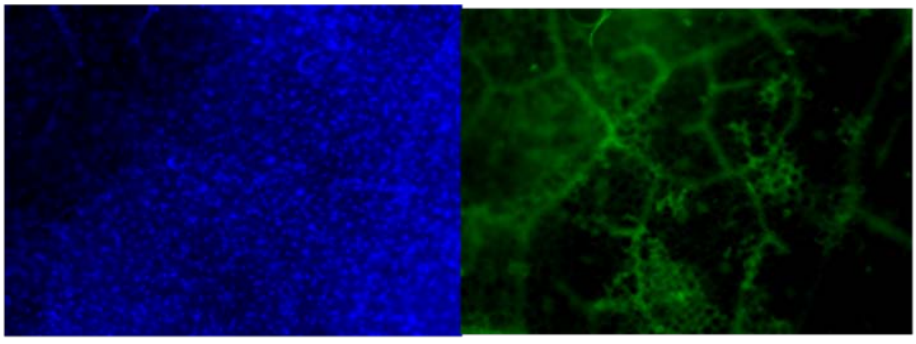
Negative treatment



House keeping gene construct



NBS region construct



LRR region construct

Figure 5.4: Transient expression analysis using immunohistochemistry

The presence of the HA-tag in infected tissue was analyzed by immunohistochemistry. DAPI (blue) indicates the DNA content in the tissue and is localized to the cell nucleus. FITC (green) expression indicates the presence of the HA-tag protein sequence in the cells. FITC expression was observed in mesophyll cells from the samples infected with binary vectors (white arrows). However, such expression was absent in the negative control.

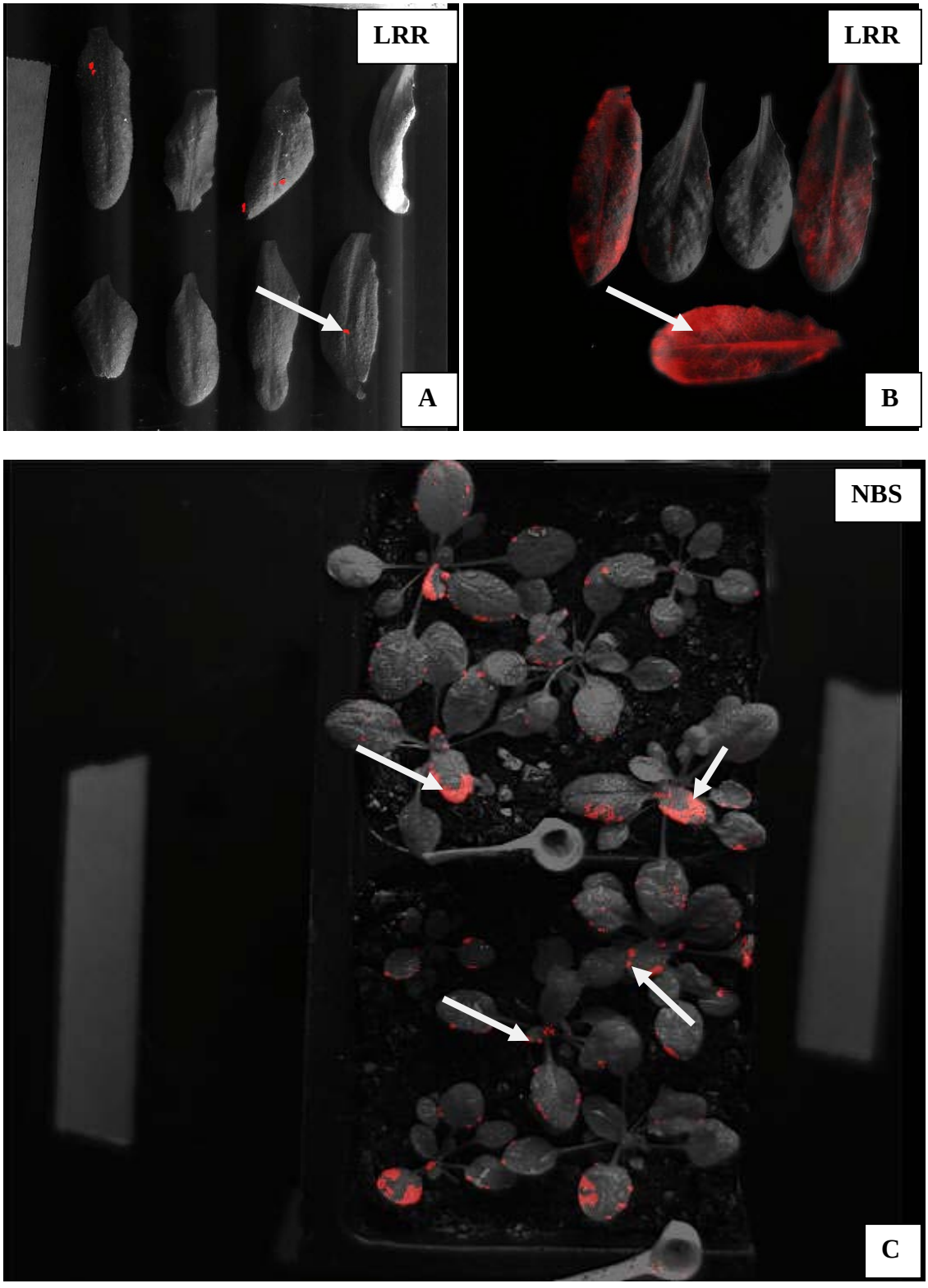

LRR
A
LRR
B
NBS
C

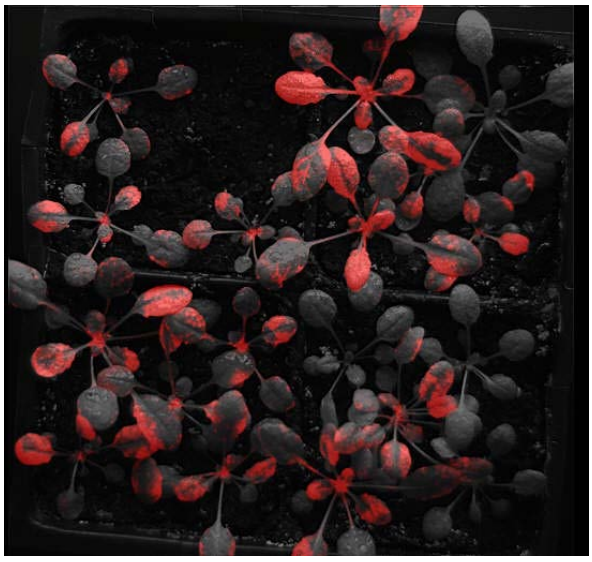
Figure 5.5 Stable expression of recombination constructs in Arabidopsis

CCD analysis of luciferase expression in transgenic plants indicates variations in expression patterns.

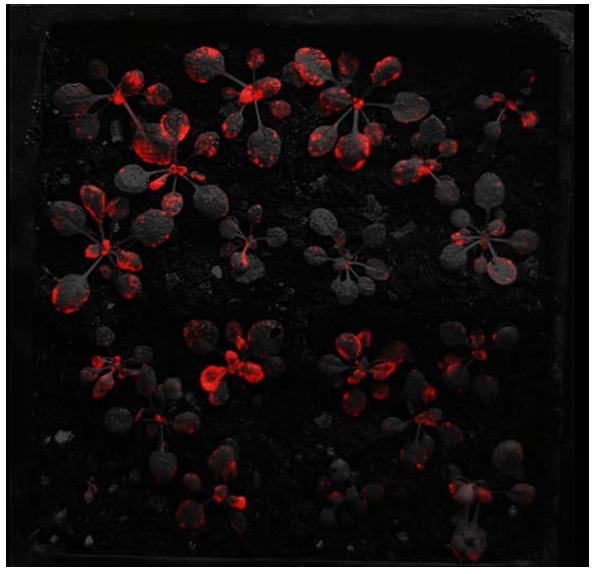
A. Some plants showed luciferase expression in the form of precise spots.

B. Plants showing either no expression or aberrant expression of luciferase.

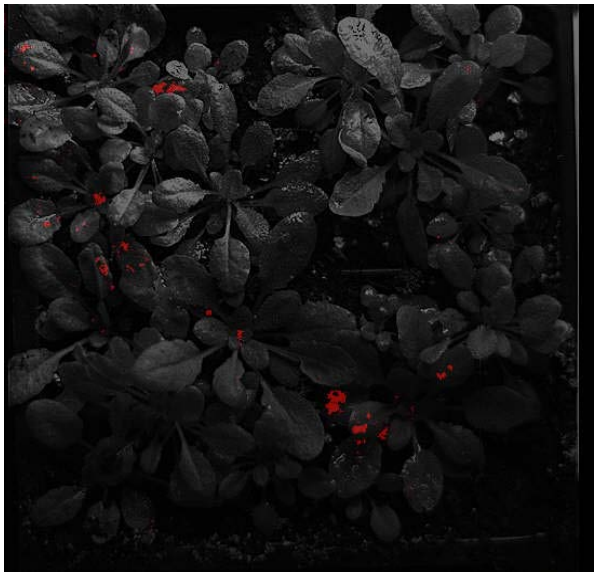
C. Plants showing an ideal number of spots were selected for further screening.



L60



L8



H399

Figure 5.6: A luciferase expression pattern in homozygous transgenic plants

Luciferase expression patterns of three individual lines used for the experiments.

A. Line L60 showed a very high expression of luciferase but no distinct spots are visible.

B. Line L8 possesses a moderate expression of luciferase in the form of spots.

C. Line H399 with low luciferase expression and precise spots.

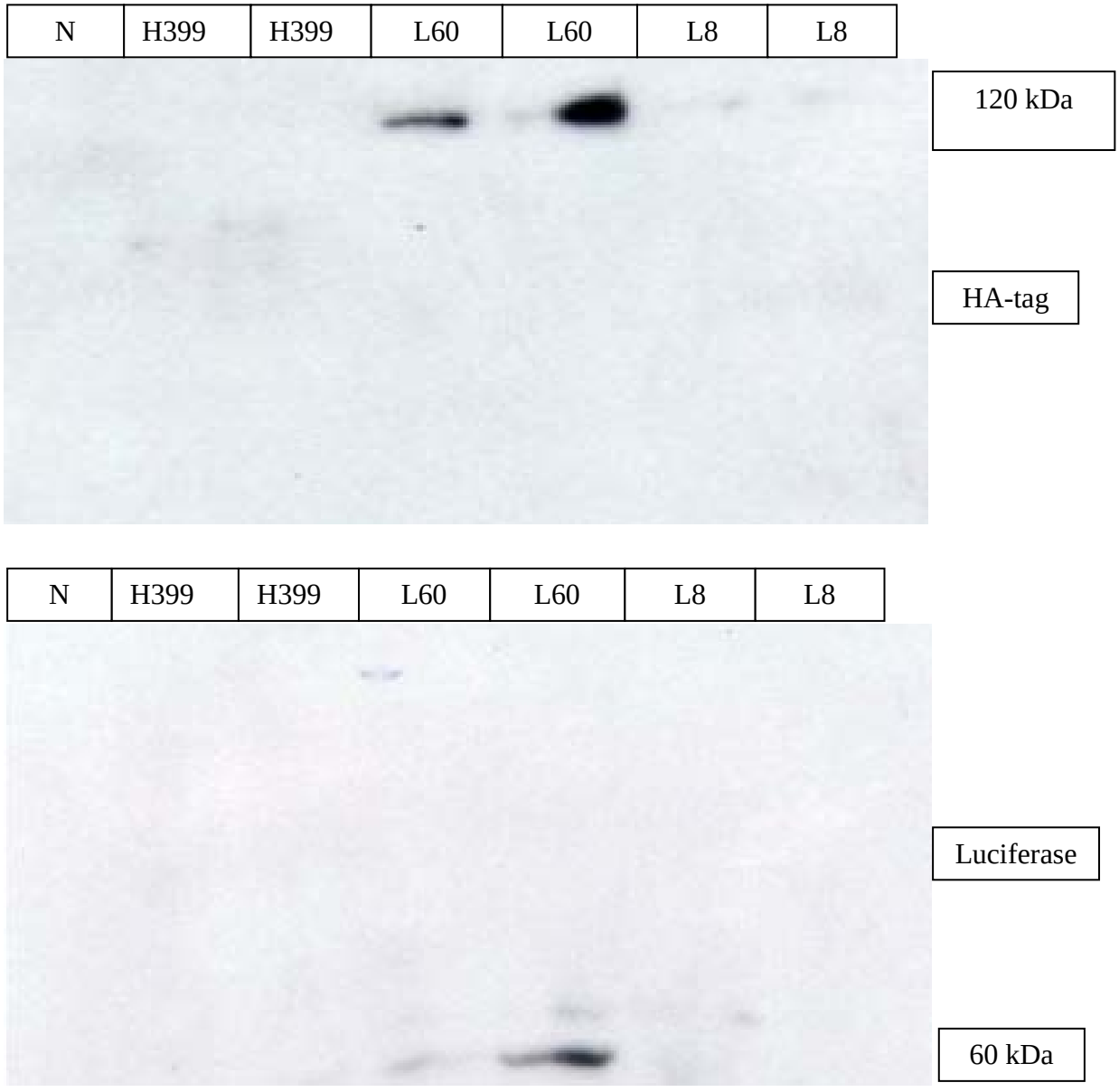


Figure 5.7: Western blot analysis of the HA-tag and luciferase proteins

A. Western blot using anti-HA-tag antibodies identified a protein in the 120 kDa size range. The band was only present in L60 samples.

B. A smaller 60 kDa protein was identified by using anti-luciferase antibodies. The band was also present only in L60 samples.

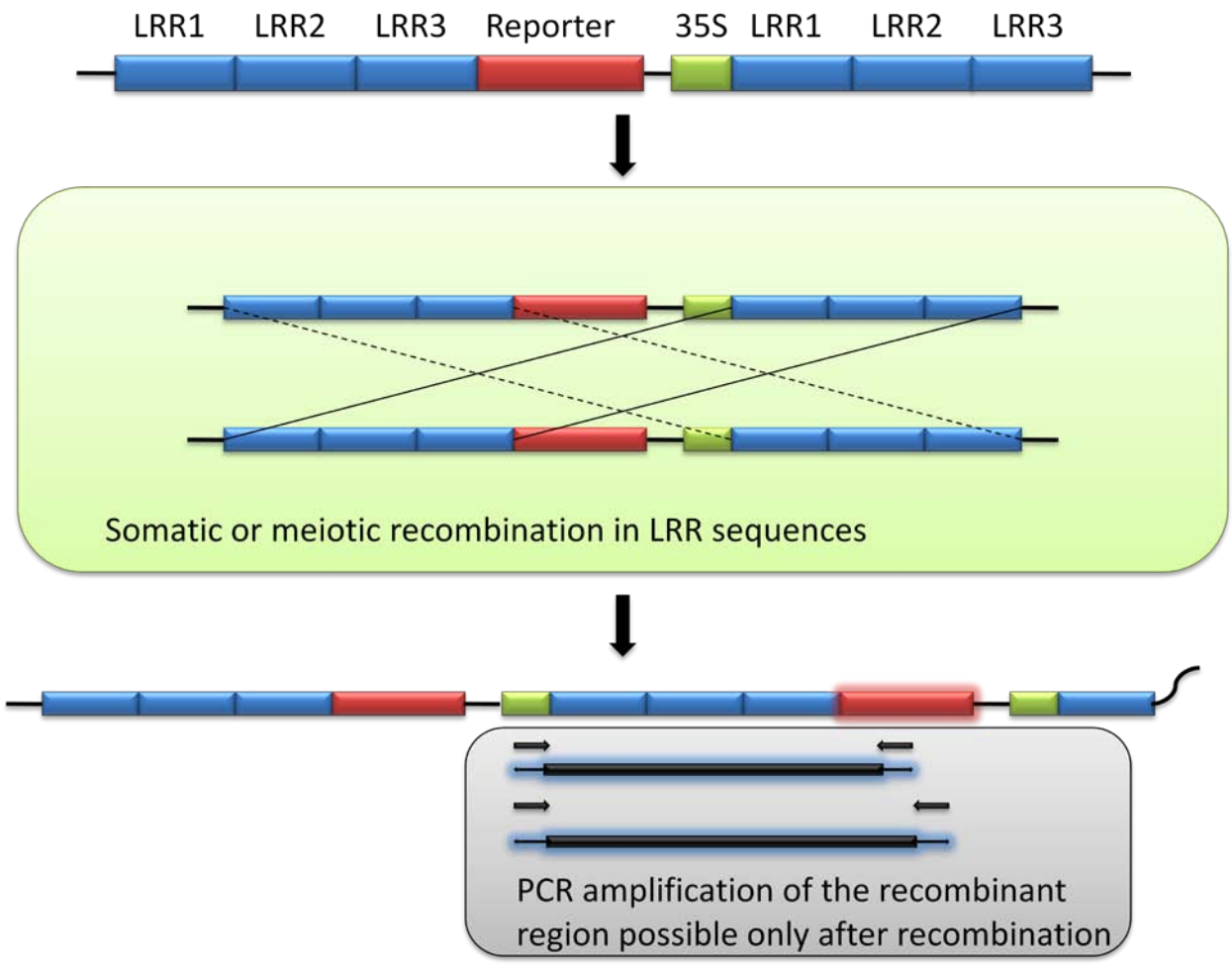


Figure 5.8: Schematic representation of a PCR approach for identification of recombination events in the transgene

Stable transformants carry transgene-based recombination constructs in the genome. The repeats present in the construct can undergo recombination with another copy of repeats present either in sister chromatids or on the homologous chromosome. This will result in luciferase gene expression as the reporter gene will be come under regulation of 35S promoter regulation. The newly recombined sequence can be identified by using a forward primer in the 35S promoter region and a reverse primer in the reporter gene region. Such amplification will not be possible in the absence of recombination in the repeats.

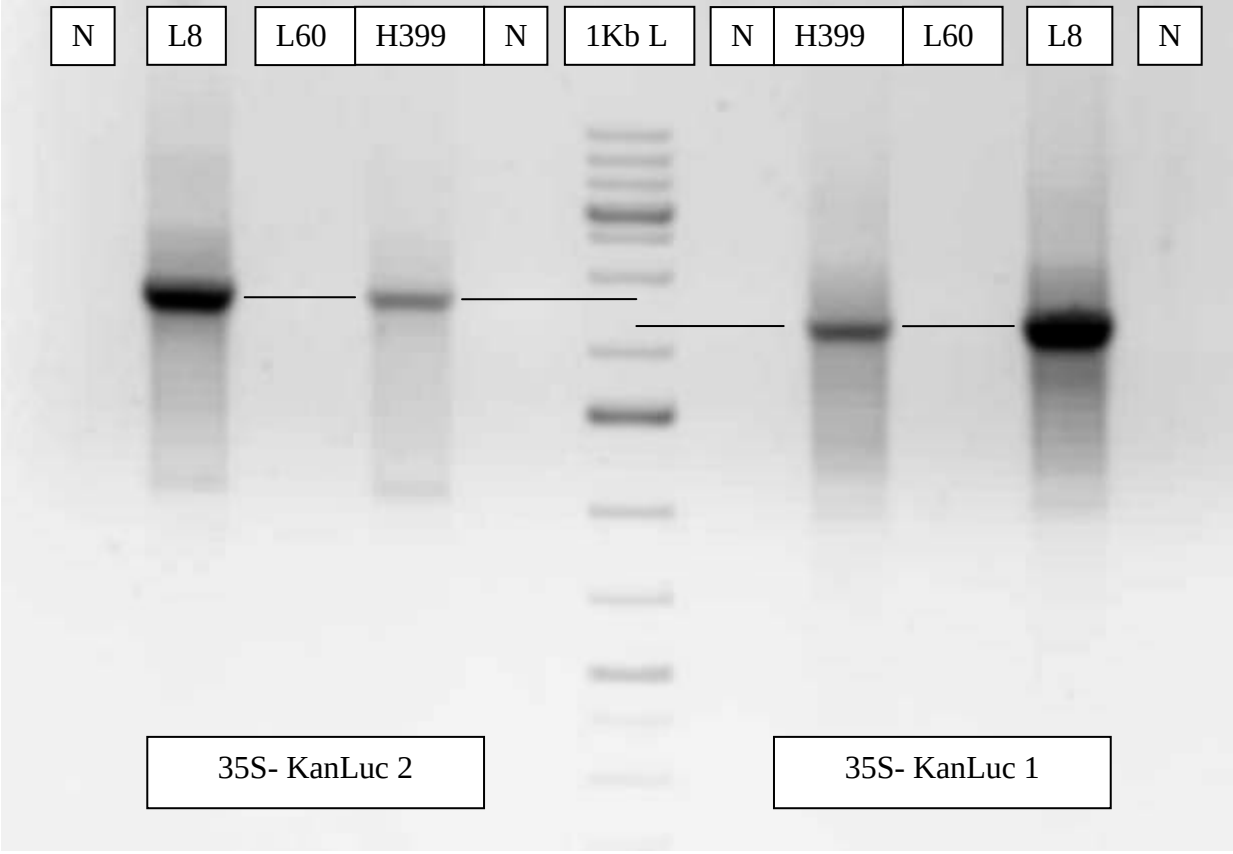
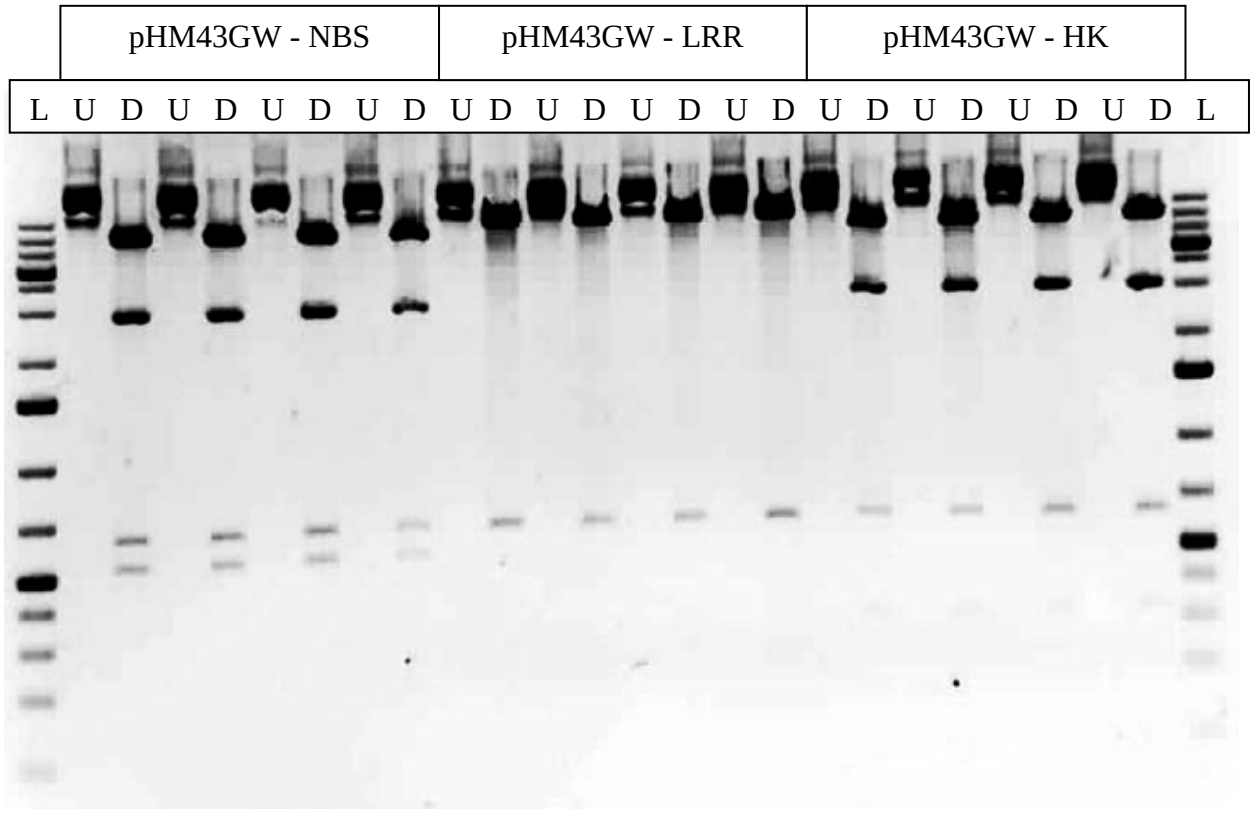


Figure 5.9: PCR confirmation of recombination events in the transgenes

PCR analysis was carried out to identify recombination events in genomic regions of transgenic plants. PCR using the 35S and KanLuc2 primers resulted in a 2.6 kb fragment in L8 and H399 samples. No fragment was present in L60 samples and in negative controls. PCRs carried out using the 35S and KanLuc1 primers revealed a similar pattern. The expected 2.4 kb fragment was observed in L8 and H399 samples but not in L60 and negative samples.



Supplementary figure 5.1: Restriction analysis of binary vectors obtained after LR reactions

Appropriate fragment insertion in the binary vectors was confirmed using restriction analysis. The binary vectors were restricted by *PstII*, and the expected fragment length was confirmed. Undigested (U) and digested (D) samples from each clone were analyzed side by side using electrophoresis.

6.0 Conclusion

Pathogens are one of the major stresses a plant encounters. Depending on the presence or absence of an resistance (*R*) gene, two types of plant pathogen interaction are possible: compatible (susceptibility) and incompatible (resistance) interactions (Hammond-Kosack and Jones, 1996). The data on the response of plants to compatible pathogens are limited. Thus, it is necessary to investigate a plant's molecular response to a compatible pathogen.

One of the observations made during compatible plant-pathogen interaction revealed the generation of a systemic signal. The signal traveled faster than the pathogen into the plant tissues and lead to transgenrational epigenetic and physiological changes (Boyko et al., 2007; Kovalchuk et al., 2003). In infected plants, the signal induced the rate of somatic recombination. The rate of recombination was measured with the aid of reporter gene-based recombination constructs (Kovalchuk et al., 2003). The next line of experiments revealed more details of the phenomenon. It was observed that DNA methylation in plants at the global genome level increased following pathogen infection. Also, the level of methylation of the leucine-rich repeats (LRR) region of *R* genes was reduced. As depicted by RFLP analysis, the rate of rearrangements in the LRR region was induced as well (Boyko et al., 2007).

In light of the abovementioned experiments, new studies were carried out to further investigate the phenomenon. One of the important questions which needed exploration was whether changes triggered by a systemic signal were able to alter the rate of pathogen

progression in host plants. The progression of three different pathogens, *TMV*, *Pseudomonas syringae* and *Phytophthora infestans*, was analyzed. The results of the experiments revealed delayed progression of pathogens in the progeny of *TMV*-infected plants (Figure 3.4 and 3.5). These results are of great importance as they represent the first observations in which the progeny of infected plants gained higher levels of resistance to pathogens. It would be interesting to carry out further experiments to find out whether the progeny of plants exposed to other pathogens also acquire a similar resistant phenotype.

Changes in plant responses to pathogens should be derived from alterations in the expression of genes involved in pathogen response. Pathogenesis-related (*PR*) genes are important candidate genes for such alterations. *PR* genes are induced by both compatible and incompatible pathogen infections, and they are able to impart resistance against pathogens (Niderman et al., 1995; van Loon, 1985). When the expression of *PR1* genes was analyzed under the influence of *TMV* and *Pseudomonas syringae* infection, a higher induction of *PR1* genes was recorded in the progeny of infected plants (Figure 3.8). The higher induction of PR genes could result from a decrease in DNA methylation or histone modifications in the promoter regions. Future studies to characterize the PR gene promoters might reveal more interesting details.

In parallel to PR gene expression, callose deposition in plants was analyzed. Callose deposition at plasmodesmata is induced following pathogen infection and limits the spread of pathogen (Bucher et al., 2001; Iglesias and Meins, 2000). Experiments monitoring callose deposition following infection with *Pseudomonas syringae* showed faster accumulation of callose in the progeny of infected plants (Figure 3.9). The results of these experiments might

explain why higher resistance to fungal and bacterial pathogens was observed. The epigenetic status of genes involved in the callose synthesis pathway might be an interesting objective for future studies.

The total polyphenolic content in the progeny of infected plants was also analyzed. Various polyphenols present in plants are induced by several pathogens (Dixon et al., 2005; Frey et al., 1997; Hain et al., 1993). The presence of higher levels of total polyphenols was observed in the progeny of *TMV* infected plants.

The second direction of research explored changes in various secondary metabolites in the progeny of *TMV*-infected plants. Interestingly, a higher amount of total secondary metabolites was observed in the progeny of these plants (Figure 4.3). When individual metabolites were profiled, a higher amount of sugars and amino acids was noticed in the progeny infected plants (Figure 4.5 and 4.6). However, most of the metabolites did not show significant differences at 24 and 48 hours post infection. The higher sugar content might be advantageous to plants for mitigating pathogen stress.

In the experiments conducted by Boyko et al. (2007), a high rate of rearrangements in the progeny of infected plants was observed. The increased rate of rearrangements may be the consequence of higher recombination rates in the LRR regions. No adequate technique was available to analyze sequence-specific recombination in somatic tissues. In response, efforts to develop a novel reporter gene-based sequence-specific recombination construct were initiated. As a result, transgenic Arabidopsis plants consisting of sequence-specific recombination constructs were produced. The presence of the reporter gene and

recombination events in the sequence under investigation was confirmed (Figure 5.7 and 5.9). Hence, a new line of recombination constructs is now available to aid a study of recombination in somatic tissues. Further experiments to analyze changes in the rate of recombination in LRR sequences need to be carried out. If such changes are observed only in the LRR sequences and not in other control sequences, it could suggest the existence of an unknown cellular mechanism which regulates genetic rearrangements allowing changes in some specific areas of the genome. This mechanism may well be of epigenetic nature. Indeed, previous studies showed that loci carrying homology to the LRR regions of resistance genes had lower levels of methylation and a higher frequency of rearrangements.

The analysis of compatible plant-pathogen interactions undertaken in the current study reveals new insights. For the first time, it has been established that the progeny of pathogen-infected plants are able to delay the progression of the pathogen. Changes in response to pathogen challenge occur in many cellular components, including changes in PR gene expression and those in the progression of symptoms in response to pathogen infection. Also the progeny of pathogen-infected plants show cross-tolerance to unrelated stresses. From the observations it can be suggested that plants undergo many drastic changes following pathogen infection. These changes include alterations in the epigenetic status of various genes. Changes in the epigenetic status may be the reason for observed alterations in PR gene induction, callose deposition and secondary metabolite accumulation. A detailed investigation into the epigenetic regulation of genetic pathways involved in pathogen response can be recommended. The use of model plant species like Arabidopsis might enhance the information retrieval from such experiments. Since many mutants in the

epigenetic pathway and in the pathway of pathogen response exist in Arabidopsis, it will also help in better understanding of the process of establishment of transgenerational changes.

6.1 References:

Boyko, A., Kathiria, P., Zemp, F.J., Yao, Y., Pogribny, I., and Kovalchuk, I. (2007). Transgenerational changes in the genome stability and methylation in pathogen-infected plants: (virus-induced plant genome instability). Nucleic Acids Res *35*, 1714-1725.

Bucher, G.L., Tarina, C., Heinlein, M., Di Serio, F., Meins, F., Jr., and Iglesias, V.A. (2001). Local expression of enzymatically active class I beta-1, 3-glucanase enhances symptoms of *TMV* infection in tobacco. Plant J *28*, 361-369.

Dixon, R.A., Xie, D.Y., and Sharma, S.B. (2005). Proanthocyanidins--a final frontier in flavonoid research? New Phytol *165*, 9-28.

Frey, M., Chomet, P., Glawischnig, E., Stettner, C., Grun, S., Winklmair, A., Eisenreich, W., Bacher, A., Meeley, R.B., Briggs, S.P., *et al.* (1997). Analysis of a chemical plant defense mechanism in grasses. Science *277*, 696-699.

Hain, R., Reif, H.J., Krause, E., Langebartels, R., Kindl, H., Vornam, B., Wiese, W., Schmelzer, E., Schreier, P.H., Stocker, R.H., *et al.* (1993). Disease resistance results from foreign phytoalexin expression in a novel plant. Nature *361*, 153-156.

Hammond-Kosack, K.E., and Jones, J.D. (1996). Resistance gene-dependent plant defense responses. Plant Cell *8*, 1773-1791.

Iglesias, V.A., and Meins, F., Jr. (2000). Movement of plant viruses is delayed in a beta-1,3-glucanase-deficient mutant showing a reduced plasmodesmatal size exclusion limit and enhanced callose deposition. Plant J *21*, 157-166.

Kovalchuk, I., Kovalchuk, O., Kalck, V., Boyko, V., Filkowski, J., Heinlein, M., and Hohn, B. (2003). Pathogen-induced systemic plant signal triggers DNA rearrangements. Nature *423*, 760-762.

Niderman, T., Genetet, I., Bruyere, T., Gees, R., Stintzi, A., Legrand, M., Fritig, B., and Mosinger, E. (1995). Pathogenesis-related PR-1 proteins are antifungal. Isolation and characterization of three 14-kilodalton proteins of tomato and of a basic PR-1 of tobacco with inhibitory activity against Phytophthora infestans. Plant Physiol *108*, 17-27.

van Loon, L.C. (1985). Pathogenesis related proteins. Plant Mol Biol *4*, 111-116.