“Gene Expression Analysis in ‘*Candidatus* Phytoplasma mali’-resistant and -susceptible *Malus* genotypes”

Dissertation

Zur Erlangung des Grades

Doktor der Naturwissenschaften

Am Fachbereich Biologie
Der Johannes Gutenberg-Universität Mainz

Mirko Moser
geb. am 29/01/1976 in Bozen (Italien)

Mainz, 2010

Dekan:

1. **Berichterstatter:**
2. **Berichterstatter:**

Tag der mündlichen Prüfung: 06.09.2010

to
my parents

Abbreviations

ADF	actin depolymerizing factor
AFLP	amplified fragment length polymorphism
AP	apple proliferation
APS	ammonium persulfate
BAP	6-benzylaminopurin
bp	base pairs
bq	becquerel
BSA	bovine serum albumin
cDNA	complementary deoxyribonucleic acid
Ci	curie
CTAB	cetyltrimethylammoniumbromid
cv	cultivar
DMSO	dimethylsulfoxid
DNA	deoxyribonucleic acid
dNTP	deoxynucleotidetriphosphate
DTT	1,4-dithiothreitol
EDTA	ethylendiaminetetraacetic Acid
EFα1	elongation Factor alpha 1 subunit
EtBr	ethidium Bromide
GA-3	gibberellic Acid-3
GD	*Malus domestica* cv Golden delicious
h	hour
IBA	indole-3-butyric acid
KAc	potassium acetate
kb	kilobase
KCl	potassium chlorid
kDa	kilo Dalton
KOAc	potassium acetate
LB	luria broth
LiCl	lithium chloride
MgAc	magnesium acetate
$MgCl_2$	magnesium chloride
min	minute
mRNA	messenger ribonucleic acid
MS	*Malus sieboldii*
NaCl	sodium chloride
Na-citrate	sodium citrate
NAD^+	nicotinamide Adenine Dinucleotide
$(NH_4)_2SO_4$	ammonium sulfate
o.p.	open pollinated
PCR	polymerase chain reaction
PVP	polyvinylpirrolidone

qPCR	quantitative polymerase chain reaction
rpm	revolutions per minute
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
RT	room temperature
RT qPCR	reverse transcription quantitative polymerase chain reaction
scDAPrimer	specific cDNA-AFLP Primer
SDS	sodium dodecyl sulphate
sec	seconds
TAE	tris-acetate-EDTA
Taq DNA polymerase	*Thermus aquaticus* DNA polymerase modified
TBE	tris-borate-EDTA
Tris	tris(hydroxymethyl)aminomethane
Tris HAc	tris acetic acid
tRNA	transfer-RNA
U	unit
V	volt
v/v, w/v	volume/volume, weight/volume
WL	wavelength
X-gal	5-bromo-4-chloro-3-indolyl-ß-galactopyranoside

1 Summary

Apple proliferation (AP) disease is the most important graft-transmissible and vector-borne disease of apple in Europe. '*Candidatus* Phytoplasma mali' (*Ca*. P. mali) is the causal agent of AP and apple (*Malus* x *domestica*) and other *Malus* species are the only known woody hosts. In European apple orchards, the cultivars are currently mainly grafted on one rootstock, *M.* x *domestica* cv. M9 because of its good agronomic traits. Unfortunately, M9 like all other *M.* x *domestica* cultivars is susceptible to '*Ca*. P. mali' remaining infected for lifetime. However, resistance to AP was found in the wild genotype *Malus sieboldii* (MS) and in MS-derived hybrids but they were characterised by poor agronomic value. The breeding of a new rootstock carrying the resistant and the agronomic traits was the major aim of a project of which this work is a part. The objective of this thesis was to shed light into the unknown resistance mechanism. For this, the interaction of the phytoplasma with the plant was studied by analysing differences between the '*Ca.* P. mali'-resistant and -susceptible genotypes related to constitutively expressed genes or to induced genes during infection. The cDNA-Amplified Fragment Length Polymorphism (cDNA-AFLP) technique was employed in both approaches. Differences related to constitutively expressed genes were investigated by comparing the transcription profiles of two '*Ca.* P. mali'-resistant hybrid genotypes (4551 and H0909) with that of the '*Ca.* P. mali'-susceptible M9. The triploid 4551 and the tetraploid H0909 share the common parent MS but H0909 has M9 as father in addition. 232 cDNA-AFLP bands present in the two resistant genotypes but absent in the susceptible one were selected and isolated. However, several different products were found associated to each band. Therefore, two different macroarray hybridisation experiments were performed with the isolated cDNA-AFLP fragments and 6 and 126 unique spots, respectively, were selected. The analysis of the corresponding cDNA-AFLP fragments yielded 40 sequences which encoded for genes of unknown function or a wide array of functions including plant defence. The analysis of the induced gene expression in the plant-pathogen interaction required a system where the plants were grown under controlled conditions. Therefore, an *in vitro* system was chosen based on healthy and '*Ca.* P. mali'-infected micropropagated plants. Infection trials using *in vitro* grafting of '*Ca.* P. mali'-strain AP showed a good transmissibility to susceptible plants (73 % GD; 33 % M9) but no (H0909, MS) or low transmission to resistant MS-derived genotypes (24% D2212). Nevertheless, H0909 and MS infected with the '*Ca*. P. mali'-strain AT2 were produced in parallel work and the resistance phenotype could be reproduced in the *in vitro* system. These healthy and infected susceptible and resistant *in vitro* plants were employed in the plant-pathogen interaction study. In addition, e*x vitro* plants were produced in order to have an independent control for the expression of genes differentially expressed in the *in vitro* plants. In the cDNA-AFLP analysis of '*Ca.* P. mali'-infected *in vitro* plants 63 bands were identified in patterns where an over-expression

in the infected state was observed in both the H0909 and MS genotypes. From the sequence analysis the major part (37 %) of the sequences showed homology with products of unknown function. The other genes were involved in plant defence, energy transport/oxidative stress response, protein metabolism and cellular growth. Real-time qPCR analysis was employed in order to validate the differential expression of the genes individuated in the cDNA-AFLP analysis. Since no internal controls were available for the study of the gene expression an analysis on housekeeping genes was performed showing that the most stable genes in our system are the elongation factor-1 alpha (EF1) and the eukaryotic translation initiation factor 4-A (eIF4A). Moreover, an algorithm for the fast and reliable analysis of the real-time PCR data was developed based on two previously published methods (DART-PCR; LinReg). Twelve out of 20 genes investigated through real-time qPCR were significantly differentially expressed in at least one genotype either in *in vitro* plants or in *ex vitro* plants. In few cases a tendency to over- or under-expression was observed though not supported by the statistical analysis due to the high variability of the real-time datasets. Overall, about 20% of the genes confirmed their cDNA-AFLP expression pattern in *M. sieboldii* or H0909 in either *in vitro* or *ex vitro* plants. On the contrary, 30 % of the genes showed down-regulation or were found not differentially expressed. For the remaining 50 % of the genes a contrasting behaviour was observed.

Genes involved in the general response against stress conditions and pathogens were observed in MS. Here, we hypothesise that in MS the '*Ca.* P. mali'-dependent repression of genes involved in processes like photosynthesis and in the electron transfer chain will unbalance the photosynthetic activity and the photorespiration. As result, and in contrast to what was observed in *M. x domestica* genotypes, an induction of genes involved in the general response against pathogens was found in MS. These genes involved the pathway of H_2O_2 and the production of secondary metabolites leading to the hypothesis that a response based on the accumulation of H_2O_2 in MS would be at the base of its resistance. This resembles a phenomenon known as "recovery" where the spontaneous remission of the symptoms is observed in susceptible plants. While the "recovery" occurs in a stochastic way in old plants the resistance in MS is an inducible but stable feature.

As additional product of this work three cDNA-AFLP-derived markers, nine confirmed and 155 putative SNP markers were developed. The three markers were independently distributed among the seedlings of two breeding progenies and were associated to a genomic region characteristic of MS. The identified markers will contribute to the development of molecular markers for the resistance as well as to map the resistance on the *Malus* genome.

2 Introduction

2.1 Apple proliferation

Apple proliferation (AP) disease (Kunze, 1989) was first described in Italy in 1950 (Rui *et al.*, 1950). Since this year the presence of the disease was reported in all countries of Central and Southern Europe (Lorenz *et al.*, 1994) classifying AP as the most important graft-transmissible and vector-borne disease of apple in Europe.

Apple (*Malus* x *domestica*) and other *Malus* species are the only known woody hosts. The specific symptoms of AP in apple are the development of witches brooms (Figure 1a and Figure 1b), due to the lack of apical dominance in affected shoots, and enlarged stipules on basal leaves (Figure 1c). Additional symptoms are associated to AP but they are not unique and can be also observed in other disorders or are peculiar to a specific cultivar. AP-infected plants show also undersized fruits (Figure 1d) with poor taste resulting in an economic loss due to reduced production (Kunze, 1989).

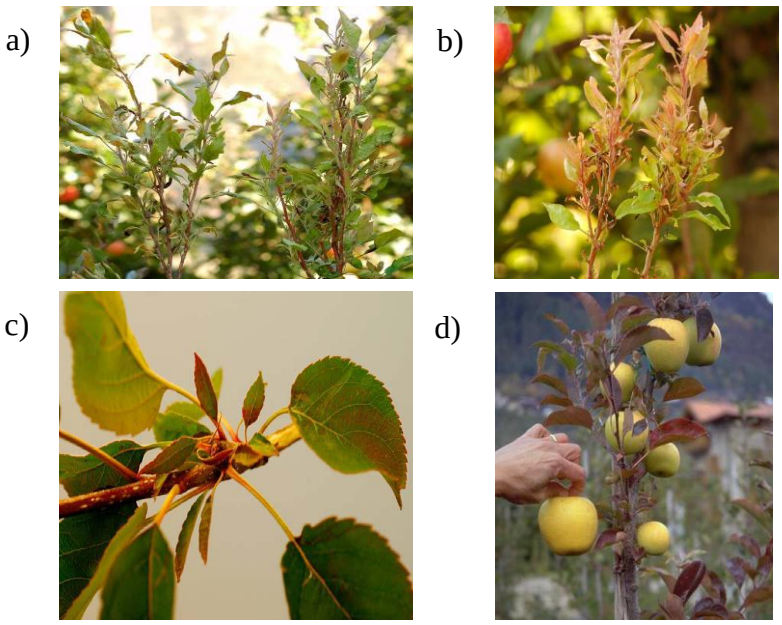


Figure 1 Symptoms of AP: a) and b) characteristic witches brooms; c) enlarged stipules; d) undersized fruits (pictures are copyright of Mauro Varner)

Apple is one of the most important fruit tree crops in Europe. In the 27 EU countries, the apple production was about 11.4 million tons in 2006 corresponding nearly to 18% of the total fruit production. Moreover, the cultivated area of about 560000 hectares represented 9% of the total surface cultivated with tree crops (data collected from the Food and Agricultural Organization - FAO stat). Fifty per cent of the European yearly apple production is concentrated in the central region formed by southeastern France, Switzerland, Austria, south Germany and northern Italy. In this area AP is of major concern (EPPO Plant Quarantine Retrieval System). The highest

economic impact of the disease in the last decade was reported from the apple growing region of Trentino (Italy) (Branz, 2003). Trentino is a particular intensively cultivated province with an annual production of more than 1 million tons corresponding to about 10% of the European apple production. In the last 15 years, AP disease has attained an epidemic character in this region causing a huge economic loss. In 2005, a survey of the local authority for the control of plant diseases estimated an up to 65 % presence of AP-symptomatic trees in those areas where the apple trees in the orchards were more than 25 years old (Ufficio Fitosanitario della Provincia di Trento; rapporto 2005). Since 2001, containment strategies to control the spread of AP in Trentino were started at different levels: the regional research center (Instituto Agrario di San Michele, IASMA) set up an interdisciplinary research project aiming to analyse the disease spread in order to find appropriate short-term control measures as well as to find a durable, long-term solution by the development of resistant plants. This thesis is part of this research activities. In parallel, insecticide treatments against the presumed insect vectors as well as eradication programs to uproot infected trees were conducted on a regional level. These control measures were able to lower the impact of the disease but represent a high economic effort for the province and the growers. In addition, replanted young apple orchards were not free of the disease but showed in severe cases up to 3% of AP-infected plants in the second year. Therefore, the demand for the development of AP-resistant plant material of good agronomic characteristics is high, not only in Trentino but in all apple growing regions of Europe where AP is an important problem.

2.2 Phytopathogenic mollicutes: phytoplasmas and spiroplasmas

In 1968, Giannotti *et al*. first reported in France the identification of a phloem-restricted phytoplasma as the proliferation agent in apple. Apple proliferation is only one of hundreds of diseases associated with phytoplasmas. However, other types of microorganisms named spiroplasmas can also cause the typical symptoms of phytoplasmoses. Phytoplasmas and spiroplasmas are wall-less, phloem-restricted, plant pathogenic gram-positive eubacteria belonging to the class *Mollicutes* (Weisburg *et al.*, 1989; Bové and Garnier, 1998). Several hundreds of plant species worldwide are hosts for both of these microorganisms (Lee *et al.*, 2000; Seemüller *et al.*, 2002) that are transmitted through leafhopper or psyllid vectors in which they also multiply (Weintraub and Beanland, 2006). Spiroplasmas have been cultured on media, are motile and are characterised by a helical shape maintained both *in planta* and *in vitro* (Saglio *et al.*, 1971). The most studied phytopathogenic spiroplasma is *Spiroplasma citri*, the causal agent of citrus stubborn disease (Saglio *et al.*, 1973; Bové *et al.*, 1989; Bové and Garnier, 2003). The possibility to cultivate *S. citri* in axenic conditions allowed the study of the morphology, the motility and the relationships with its two hosts, the plant and the insect vector. Moreover, the development of tools for the gene transfer and mutagenesis in *S. citri* has led to new knowledge about the cellular and molecular

properties of this organism making it a model to study the spiroplasma-plant interactions (Stamburski *et al.*, 1991; Gasparich *et al.*, 1993; Bove *et al.*, 2003).

In contrast, phytoplasmas (formerly called mycoplasma-like organisms, MLOs) are non-culturable, unmotile and their size and shape are variable. They are found in round shape (diameter 60-100 nm), globular shape (150-1100 nm) or filamentous particles (1-2 nm to several nm long). Because of these small dimensions and the pleomorphism, the phytoplasmas can spread in the phloem systemically cell to cell through the pores of the sieve plate colonising all the organs of the plant host. The multiplication may take place through scission or direct division (Lee *et al*., 2000; Seemüller *et al*., 2002; Christensen *et al.*, 2005). The most severely affected crops by phytoplasma-associated diseases are the deciduous fruit trees (McCoy *et al.*, 1989; Seemüller *et al.*, 1998) which are susceptible to phytoplasmoses known as “yellows diseases”, “witches’brooms disease” and “X-disease”.

2.2.1 Phytoplasma classification

Four decades ago, it was thought that phytoplasma diseases were caused by viral infections. In 1967, a Japanese research group demonstrated through electron microscope analysis the presence of wall-less microorganisms in yellows-diseased plants (Doi *et al.*, 1967).
The new class of plant-disease agents was named “mycoplasma-like organism” (MLO) because of their morphological and ultrastructural resemblance to the mycoplasmas present in animals.

In the early 1970s and based on experimental results, the yellows diseases were associated to two different types of phytopathogenic agents: spiroplasmas and virescence agents or not-helicoidal mycoplasmas.

In the following thirty years, the further characterisation of the yellows diseases-associated microorganisms demonstrated that the latter were unique and the definition MLO was substituted with the new term of “phytoplasma” (Kirkpatrick, 1992).

In 1993, the International Committee of Systematic Bacteriology (ICSB) Subcommittee Mollicutes defined specific criteria that a microorganism has to satisfy in order to be classified as phytoplasma:

- the analysis by transmission electron microscopy (TEM) must reveal cell wall less pleomorphic prokaryotes;
- the organisms have to be individuated in the sieve cells of the phloem and in general they have to be associated with diseases that cause symptoms of decline, yellowing and development anomalies, in particular of the flowers;
- the organisms are transmitted by phloem-feeding cicadellidae, cixiids or psillyds;
- the organisms are resistant to penicillin and highly sensitive to tetracyclins;
- the genome of these microorganisms is small (530-1350 kb) and the content in guanine (G) and cytosine (C) is low (23-29 mol % of the total of bases);

- the sequence analysis must indicate that the organism is phylogenetically related to other *Mollicutes.*

The impossibility to cultivate these organisms in an artificial media was and still is the major problem limiting the acquisition of knowledge about these pathogens. The detection of the phytoplasmas in the past was done only by electron microscopy of ultrathin sections of infected tissues or indirectly by light microscopy of stained-sieve cells by Diene´s stain (Deeley *et al.*, 1979) or DAPI stain (Seemüller, 1976). The identification of the phytoplasma associated with each disease was based on symptomatology, plant host range, relationships to insect vectors and sometimes on the geographic area. More specific methods employed for differentiating the phytoplasmas were based on the lack or the presence of extra-chromosomal DNA, the content in GC and the dimensions of the genome. The development of serological and nucleic acid-based molecular probes allowed a first classification of these microorganisms. However, the sensitivity of these methods was too low for the detection of several phytoplasmas associated with woody plants in which they are present at low concentration. The development of the polymerase chain reaction (PCR) technology in the 1980s improved the phytoplasma diagnostics (Chen *et al.*, 1992; Lee and Davis, 1992). This molecular tool provided relatively simple, sensitive and reliable means for detection and identification of phytoplasmas. In the late 1980s and early 1990s "universal" (generic) oligonucleotide primers were designed based on the conserved 16S rRNA gene. Restriction fragment length polymorphism (RFLP) analysis based on PCR-amplified 16S rDNA sequences allowed the creation of a comprehensive classification scheme of the phytoplasmas and further advanced the diagnostics of phytoplasmal diseases (Ahrens and Seemüller, 1992; Davis and Lee, 1993; Schneider *et al.*, 1993; Seemüller *et al.*, 1994; Gundersen and Lee, 1996; Gundersen *et al.*, 1996; Smart *et al.*, 1996; Seemüller *et al.*, 1998; Lee *et al.*, 1998b; Martini *et al.*, 2007). Furthermore, analyses based on the spacer region (SR) between the 16S and 23S rDNA, the regions that code for the ribosomal proteins (RP genes) (Schneider *et al*., 1993; Lee *et al*., 1998b) as well as analyses based on the sequence of the elongation factor Tu (Schneider *et al.*, 1997) and on the *secY* gene (encoding for a subunit of the translocation system Sec) (Lee *et al.*, 2004b) improved the classification of a broad array of phytoplasmas. Particularly effective resulted the combined analysis of the 16S rDNA sequences and the less conserved RP genes revealing the latter as useful tool for a more specific differentiation of the phytoplasmas strains that belong to the same group like the aster yellow and the elm yellow phytoplasmas (Lee *et al.*, 2004a; Lee *et al*., 2004b). Moreover, PCR-RFLP analyses on non ribosomal regions using specific primers further improved the capability to discriminate between closely related phytoplasmas (Jarausch *et al.*, 2000).

The molecular characterisation of the conserved genes has brought to a partial solution the classification problems allowing the development of a system that is independent to the characters of the disease. The genome sequencing of four phytoplasmas, aster yellows witches´broom (AY-

WB) (Bai *et al.*, 2006), onion yellows (OY) phytoplasma (Oshima *et al.*, 2004), '*Candidatus* Phytoplasma mali' (Kube *et al.*, 2008) as well as a strain of '*Ca.* P. australiense' (Tran-Nguyen *et al.*, 2008). greatly improved the number of sequences available for a better phytoplasma differentiation and classification.

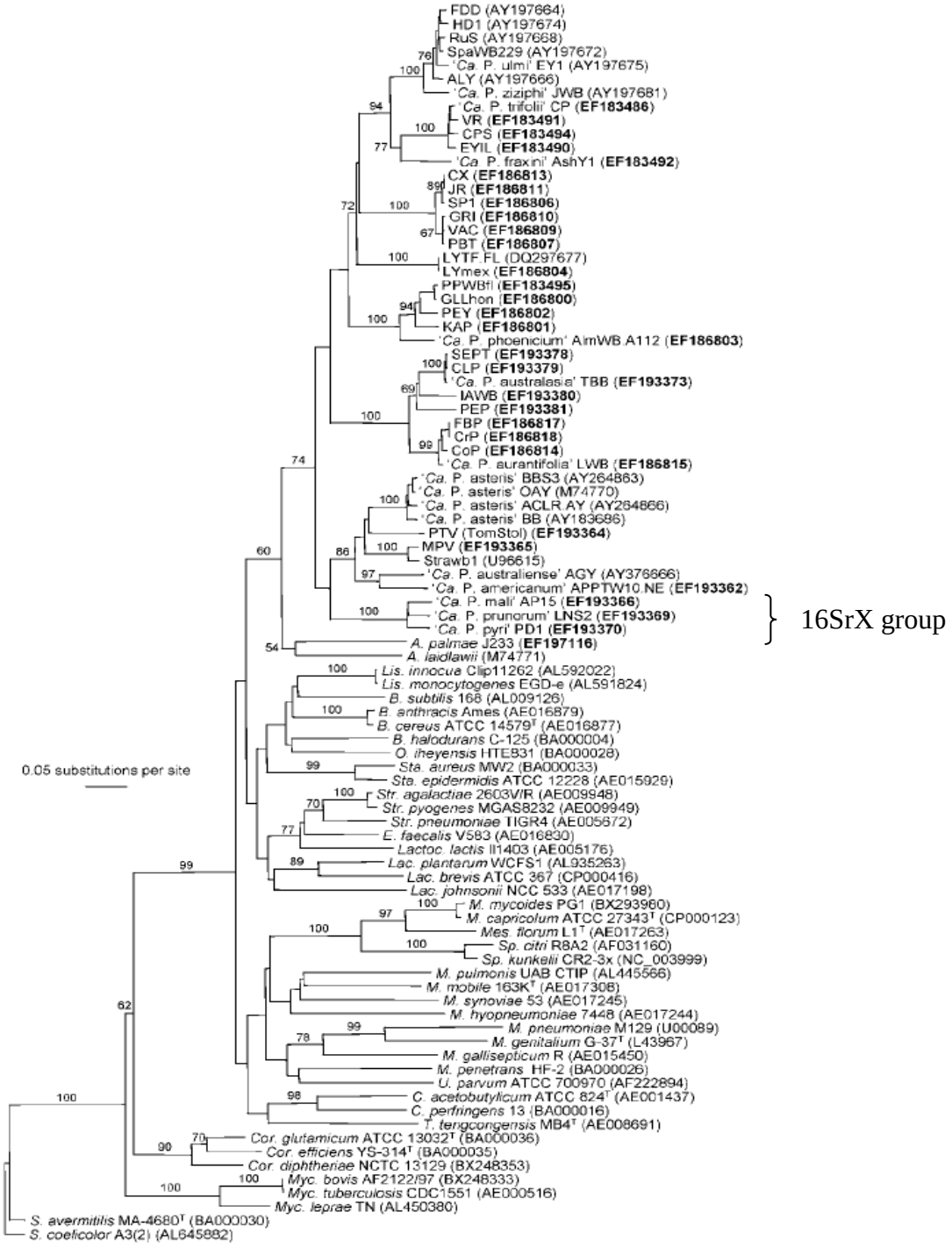


Figure 2 Phylogenetic tree constructed by parsimony analysis of full 16S rRNA sequences (Martini *et al.*, 2007)

Presently the knowledge acquired through the different molecular tools allowed for several different phytoplasmas to achieve the status of '*Candidatus*' (*Ca*) that indicates a provisional taxon status at the species level (IRPCM Phytoplasma/Spiroplasma Working Team - Phytoplasma Taxonomy Group *et al*., 2004). The European Stone Fruit Yellows (ESFY) causal agent '*Candidatus* Phytoplasma prunorum' and the Pear Decline (PD) causal agent '*Candidatus* Phytoplasma pyri' are phylogenetically grouped together with the AP causal agent '*Candidatus* Phytoplasma mali' in the AP phytoplasma group (Seemüller *et al*., 1998) or 16SrX group (Lee *et al*., 2000) within the AP subclade which is one of the major branches in the phytoplasma clade (Figure 2).

The comparison between the molecular classification of some phytoplasmas and their phenotypic characteristics like host range, vectors and induced symptoms has shown that there is no consistent correlation between the type of symptoms and the identity of the phytoplasma (Schneider and Seemuller, 1994). Closely related phytoplasmas can cause different symptoms in the same host plant (Lee *et al.*, 1992). Infection trials on *Catharanthus roseus* (*C. roseus*, syn. Vinca rosea; periwinkle), a host widely used to maintain the phytoplasmas *in vivo*, with different phytoplasmas showed different symptoms also compared to those that can be observed on their own fruit tree host (Lepka *et al.*, 1999). This was also the case of '*Candidatus* Phytoplasma mali' where the symptoms were quite different to those observed in apple showing in *C. roseus* leaves with light green colour, mild vein yellowing and smaller flower size.

2.2.2 '*Candidatus* Phytoplasma mali'

The first differentiation and specific detection of the AP causal agent '*Candidatus* Phytoplasma mali' (*Ca*. P. mali) was obtained through dot blot hybridisation using a cloned fragment of its chromosomal DNA (Bonnet *et al.*, 1990). Further, the application of the PCR facilitated the diagnosis of the AP-causing pathogen (Firrao *et al.*, 1994; Jarausch *et al.*, 1994a). The development of specific primers based on the 16S rDNA and on non ribosomal sequences of '*Ca*. P. mali' in association with the restriction fragment length polymorphism (RFLP) analysis of the PCR products showed that several genetically slightly different phytoplasmas are associated with the disease (Schneider *et al*., 1993; Kison *et al.*, 1994; Jarausch *et al.*, 1994a). Moreover, the analysis of the chromosome size of several phytoplasmas showed that two isolates of '*Ca*. P. mali' had a genome size of about 600 kb (AT from Germany) and 690 kb (AP15 from Italy). In comparison, '*Ca*. P. prunorum' had a size of 630 kb and '*Ca*. P. pyri' a size of 660 kb (Marcone *et al.*, 1999; Seemüller and Schneider, 2004; Kube *et al.*, 2008). Nevertheless, apple is non-host for '*Ca*. P. prunorum' and '*Ca*. P. pyri' (Seemüller and Schneider, 2004) while '*Ca*. P. mali' was reported in *Prunus avium*, *Prunus armeniaca* and *Prunus domestica* (Mehle *et al.*, 2007). Although the three phytoplasmas are characterised by narrow differences in chromosome size and high similarity of the 16S rDNA sequences, their inter-variability was shown to be slightly higher in the

16S-23S spacer region (Smart *et al*., 1996) and more concentrated in the sequence region that codes for the extracellular domain of the immunodominant protein (IMP) (Seemüller and Schneider, 2004). IMP may play a role in the pathogen-host interaction and this could explain the high specificity of '*Ca*. P. mali', '*Ca*. P. prunorum' and '*Ca*. P. pyri' for different plant- and insect-hosts (Berg *et al.*, 1999; Barbara *et al.*, 2001; Barbara *et al.*, 2002; Morton *et al.*, 2003).

In the last decade the introduction of more efficient molecular methods allowed a better characterisation of the phytoplasmas showing that even for '*Ca*. P. mali' several subtypes can be classified based on the PCR-RFLP pattern (Jarausch *et al*., 2000; Cainelli *et al.*, 2004). Moreover, in a recent study Seemüller and Schneider showed differences in virulence and in the capability to colonize the host and to develop the classical AP symptoms between '*Ca*. P. mali' strains (Seemüller and Schneider, 2007). Their characterisation based on RFLP patterns using a combined set of rarely and frequently cutting enzymes showed genetic diversity between the strains. Although this variability could not be associated to markers correlated with the virulence character the results showed that there is a considerable genetic heterogeneity in '*Ca*. P. mali' (Seemüller and Schneider, 2007). Mixed infections of different phytoplasma types can take place in the same host as shown in several works (Lee *et al.*, 1994; Lee *et al.*, 2000; Lee *et al.*, 2003). Similarly, mixed populations of phytoplasma strains could be present at the same time possibly determining a different host answer depending on which strain takes over (Seemüller and Schneider, 2007). Recently, the sequencing of the linear chromosome of '*Ca*. P. mali' (Kube *et al.*, 2008) allowed the selection of further genes or genomic regions to use as a molecular marker for the differentiation of '*Ca*. P. mali' strains (Schneider and Seemüller, 2009). The availability of these informations would give in the future the possibility to associate the pathogenicity character to the presence of specific genomic regions in the '*Ca*. P. mali'.

In the last years, the emerging technique quantitative PCR (qPCR) has been employed for the quantification of the number of phytoplasma cells present in both hosts, plants and insects. The method was successfully applied for the analysis of '*Ca*. P. mali' in infected plants (Baric and Dalla-Via, 2004; Jarausch *et al.*, 2004b). Because the symptom development appears to be related to the phytoplasma concentration in the host, the analysis of the qPCR data represents an important improvement. This analysis would allow the study of the disease dynamics through the relationship between the phytoplasma concentration, the disease severity and the phytoplasma diffusion in the plant-host and the insect-vector population (Bisognin *et al.*, 2008a; Bisognin *et al.*, 2008b).

2.2.3 Transmission of '*Ca*. P. mali'

The dynamics that rule the spread of AP constitute a complex system dependent on several parameters related to environmental conditions and human intervention.

'*Ca.* P. mali' can be transmitted through several ways:

- phloem-phloem connections
    - graftings
    - root bridges
- phloem sucking insects
- latently infected planting material

Two main sources are at the base of the disease spread: the introduction of a new inoculum through infected planting material that escapes the phytosanitary controls done on the young plants and the spreading of the disease from pre-existing infected plants. While the adoption of more restrictive controls on the introduction of new plant material in the orchards would eliminate the first source, the second needs the definition of a multiapproach strategy through application of defense protocols against the vector and/or the introduction of AP-resistant apple genotypes.

Marwitz *et al.* (1974) demonstrated the transmissibility of '*Ca*. P. mali' from apple to *C. roseus* by the use of the dodder *Cuscuta subinclusa* (Petzold and Marwitz, 1976). This type of transmission takes place through the formation of phloem-phloem connection in a natural way. Similarly, graftings with infected material can be one cause of spreading of the disease.

In addition, in the modern agricultural practices the short distance between the plants can favour the formation of root interconnections through which '*Ca.* P. mali' can infect healthy plants as observed in experimental trials simulating the conditions present in the orchards (Ciccotti *et al.*, 2008a). Even after eradication of the infected plants from the orchard, remaining roots in the soil can represent an inoculum source through the formation of new shoots (Baric *et al.*, 2007; Ciccotti *et al*., 2008a).

Classical monitoring methods based on the observation of the characteristic disease-related phenotype can fail to individuate infected plants for their eradication. This is due to a latent, non-symptomatic presence of the phytoplasma in the tree which is favored by several parameters like the plant age and fitness as well as the environmental conditions which can have a positive or negative influence on the symptoms development. However, the availability of new phytoplasma detection techniques has improved the sensitivity and the specificity of the screening methods allowing a more efficient monitoring of the disease spread.

Although the aspects reported above play a role in the disease diffusion, the most important spread of '*Ca*. P. mali' takes place by insect vectors. In the past, the transmission through the leafhopper *Fieberella florii* (Krczal *et al.*, 1988; Tedeschi and Alma, 2006) and the two psyllid species *Cacopsylla melanoneura* (Tedeschi *et al.*, 2002) and *Cacopsylla costalis* (syn. *C. picta*) (Frisinghelli *et al.*, 2000) was demonstrated.

Studies on the insect vector population showed that *C. picta* is the most efficient vector for the '*Ca*. P. mali' transmission in Germany and northeastern Italy (Jarausch *et al.*, 2003; Jarausch *et al.*, 2007; Mattedi *et al.*, 2008). Although in north Italy *C. melanoneura* is the most abundant psyllid (Tedeschi *et al*., 2002) its relevance as vector seems to be characteristic of the northwestern area

(Tedeschi and Alma, 2004) while in northeastern Italy and in Germany the transmission capacity of *C. melanoneura* appears very low (Mattedi *et al.*, 2008; Mayer *et al.*, 2009). Tests with pesticides conducted for the control of *C. picta* showed good efficacy in limiting the insect vector population (Mattedi *et al.*, 2007). Although the evidences of a high efficiency in pathogen transmission by *C. picta*, this indication alone does not explain the high spread of the disease in the Trentino region. All these factors together like inoculum latency in asymptomatic plants, the introduction in the orchard of infected planting material, the possibility of alternative '*Ca.* P. mali' plant hosts and the spreading through insect vectors act in a synergistic way determining the yearly high rate of new infected trees. In this context, the use of resistant plants would be the only possible solution against the diffusion of '*Ca.* P. mali' within and across orchards. This strategy would furthermore reduce the inoculums sources.

2.2.4 Phytoplasma plant-host interaction

Phytoplasmas are adapted to live in hosts belonging to different biological kingdoms: plants and insects. This involves the capability to exploit the different metabolites available in the phloem sap and in the insect hemolymph.

Studies conducted in plants after phytoplasma infection showed alteration of the normal function of the sieve cells with blocked phloem circulation or slowered phloem flux (Braun and Sinclair, 1978; Eden-Green and Waters, 1982; Kartte and Seemüller, 1991).

As consequence, the impaired translocation of photoassimilates in the phloem would result in carbohydrate accumulation in source leaves and depletion in sink organs (Batjer and Schneider, 1960; Catlin *et al.*, 1975; Kartte and Seemüller, 1991; Lepka *et al*., 1999) as observed in coconut palms infected with lethal yellows phytoplasma (Maust *et al.*, 2003). In addition, the peculiar modifications in the plant morphology could be related to an imbalance in the hormonal and growth regulators equilibrium, resulting in malformations and anomalies in the plant development (Chang and Lee, 1995).

Several symptoms observed in infected plants can be correlated to the block of the phloem sap circulation. The most common ones include alteration of the leaf pigmentation resulting in yellowing or reddening leaves as well as laminar leaves that become fragile, thick, heavy and rolled on the leaf stalk. Moreover, presence of rolled leaves, necrosis of the leaf veins, the vegetative apex and the fruits, thickness of the cortex, necrosis on the roots, reduction of the development until dwarfing and perishment of the plant are also observed (Lepka *et al*., 1999; Lee *et al*., 2000; Maust *et al.*, 2003). However, because environmental factors also have an influence on such modifications these cannot be specifically associated to the phytoplasmas presence.

Evidence supporting the possible block of the phloem flux was obtained by transmission electron microscopy (TEM) analysis. Deposit of callosium in the sieve plates, accumulation of starch in the chloroplast and their disorganisation, alteration and thickness of the cell wall and accumulation of

polyphenols were observed in phytoplasma-infected plants (Kartte and Seemüller, 1991; Lepka *et al.*, 1999; Choi *et al.*, 2004; Christensen *et al*., 2005). Other experiments evidenced abnormal deposition of callose on the sieve area that leads to the collapse of the sieve elements and the companion cells. Moreover, phloem necrosis was observed as symptom of phytoplasma diseases like pear decline (Batjer and Schneider, 1960), apple proliferation (Kartte and Seemüller, 1991) and elm yellows (Braun and Sinclair, 1978).

On the other side, interferences with the hormonal balance could be correlated to a series of morphological changes like virescence (petals are green); phyllody (transformation of the floral organs in leaves); proliferation (development of leaves from floral organs); witches brooms from lateral buds; floral malformations and aborts. In addition, flowering and vegetative cycles in the wrong season (flowering in winter and anticipation of the vegetative growth); modification of the internodes and production of small and malformed fruits of low economic value are also observed in phytoplasma-infected plants (Chang and Lee, 1995; Chang, 1998; Lee *et al*., 2000). The severity of the anatomical aberrations and the necrosis of the sieve tubes depend on the susceptibility of the host.

Specific studies on metabolism and secondary product modification showed that the presence of the pathogen caused a marked reduction in chlorophyll a, chlorophyll b and total chlorophyll content in aster yellows phytoplasma (AYP)-infected *C. roseus* plants (Chang, 1998). AYP seemed to induce the depletion of these pigments in mature chloroplasts of older leaves. Moreover, in AYP-infected *C. roseus* also the content of carotenoids in leaves diminish some weeks after inoculation. The study of the effects in the interaction between '*Ca*. P. mali' and apple showed that there is a depletion of the photosystems and photosynthetic activities (Bertamini *et al.*, 2003). A study supporting the evidence that the hormonal balance of the hosts is affected was presented by Perica *et al*. (2007). They demonstrated that in *in vitro*-grown *C. roseus* shoots separately infected with stolbur phytoplasma (SA-I) and aster yellows phytoplasma (HYDB) the presence of the pathogen influenced the auxin/cytokinin balance. Further studies on the *C. roseus*-phytoplasmas interaction showed that in general the whole plant secondary metabolism was changed after infection (Musetti *et al.*, 2000; Choi *et al.*, 2004; Favali *et al.*, 2004).

Considering the different type of metabolisms that seem to be involved in the phytoplasma-plant interaction one can argue that a broad range of changes are taking place in the plant after the phytoplasma infection. Moreover, the impossibility to grow these pathogens in axenic media further evidence the complexity of the relationship and the dependency of the phytoplasma to their plant host metabolisms. This challenges the understanding of the mechanisms that are involved in the symptom development whereas the lack of information is a limit for the design of defence strategies against these diseases.

2.3 A strategy for the development of apple proliferation resistant rootstocks

2.3.1 Use of resistance in the control of AP

In European apple orchards, the cultivars are currently mainly grafted only on one rootstock cultivar, *M.* x *domestica* cv. M9. The good productivity, the constant yield during the years and the low vigour are all characteristics that contributed to its large diffusion as optimal rootstock. Unfortunately, M9 like all other *Malus* x *domestica* cultivars, is susceptible to '*Ca*. P. mali' remaining infected for lifetime (Kartte and Seemüller, 1991).

In studies on the phytoplasma distribution in the plant organs it was observed that in infected plants the phytoplasmas disappeared from the stems during the winter due to the degeneration of the sieve tubes of which the organisms are depending (Schaper and Seemüller, 1982; Schaper and and Seemüller, 1984; Seemüller *et al.*, 1984). In susceptible plants the pathogen is able to overwinter in the root organ (of the rootstock) and to recolonise the aerial part in the next year when the new phloematical tissue is generated (Pedrazzoli *et al.*, 2008). This cyclic variation and migration of the phytoplasma distribution could be generally exploited for several phytoplasmoses affecting fruit trees in a strategy where resistant rootstocks would hinder the recolonization of the aerial part in the next season as it happens for the susceptible rootstock (Seemüller *et al.*, 2008; Bisognin *et al.*, 2008b; Seemüller *et al.*, 2009). Although the cultivars grafted on the resistant rootstock can still become infected (through new vector-mediated transmission of the pathogen), this strategy would avoid or drastically reduce (Seemüller *et al.*, 2008; Bisognin *et al.*, 2008b) the recolonization of the canopy in the following years through the effect of the resistant rootstock. The exploitation of this characteristic would be independently from the cultivar used as scion allowing to further cultivate the apple varieties present today on the market without the necessity to adopt or develop resistant cultivars. In the specific case of AP disease wild genotypes showing resistance against '*Ca.* P. mali' were individuated (Kartte and Seemüller, 1991). These genotypes have given new perspectives for a solution to AP disease and at the same time to study the phytoplasma-host interaction.

2.3.1.1 Identification of a natural resistance to AP

In a screening for AP resistance in wild *Malus* species, ornamental *Malus* species and hybrid genotypes, 5 groups were identified based on the response to the phytoplasma (Kartte and Seemüller, 1991). Resistance was found in the group formed by plants showing a low mortality and a high rate of spontaneous remission of the symptoms. These were hybrids obtained from crossings between the wild species *Malus sieboldii* with cultivars of *Malus* x *domestica*. Although these plants showed symptoms in the first year after the infection, the symptoms disappeared in the following years. The terminal growth became normal and at the end of the observation period

the phytoplasma could be detected only in few plants. Kartte and Seemüller (1991) observed that in plants that remitted the symptoms only moderate damage of the phloem tissues occurred, like in the *Malus* x *domestica* genotypes. However, the latter proved to be suitable hosts allowing the multiplication of the phytoplasmas with development of the characteristic symptoms and remaining infected lifetime. The analysis of the obtained data revealed a form of resistance associated to the wild genotype *Malus sieboldii.* However, the poor agronomic value of this genotype hindered the possibility to use it directly as rootstock in the orchards. The spontaneous remission of the symptoms, a phenomenon known as recovery, was observed also from another research group which showed that in recovered plants the phytoplasma can be still present in the roots but can not colonise the above ground part of the plant (Carraro *et al.*, 2004). In contrast, symptomatic trees were completely colonised. Variations in symptom expression are related to the presence or the absence of phytoplasmas in the aerial part of the plants (Seemüller *et al*., 1984). In recovered plants the phytoplasmas may or may not be eliminated from the bottom part (Schmid, 1965) but the diffusion of the pathogen in the phloem seems to be hindered. Little is known about the mechanism of recovery.

It is unknown if a similar mechanism is involved in the resistance shown by the *Malus sieboldii*-derived genotypes. However, the observation of Kartte and Seemüller (1991) suggested a strategy where the phloem-tolerance and good agronomic traits of the *M*. x *domestica* genotypes and the resistance associated with the wild genotypes can be combined for the breeding of new hybrids to use as rootstocks in substitution of the AP-susceptible M9. The genotypes that showed the remission of the symptoms were object of about twenty years of observations (Seemüller *et al.*, 2008; Bisognin *et al.*, 2008b). The behaviour of the resistant hybrids with regard to the disease and their classification based on different parameters was established. These parameters monitoring the type of symptoms shown by the plants, the severity of the symptoms and their manifestation during the years allowed the selection of genotypes that could be used as donors of the resistance trait (Seemüller *et al.*, 2008). In addition, a first progeny population derived from the original ancestor *M. sieboldii* was object of a study for the association of the resistance character with specific genotypic traits based on simple sequence repeats (SSR) (Bisognin *et al.*, 2008b). Based on the data of these two latter works a more complex and organised breeding program was started.

2.3.1.2 Breeding of new *Malus* hybrids for the development of a resistant rootstock

A set of interesting genotypes (Table 1) was chosen to be employed in the development of a new AP-resistant rootstock. The poor agronomic value of the hybrids (alternance in the production, fruit fall, too vigorous growth) prevented the direct use of these genotypes in the agricultural

practice. Therefore, a breeding programme was initiated crossing the plants with the M9 rootstock because of its good agronomic characteristics.

Table 1 Genotypes selected for the breeding of a new AP-resistant rootstock with good agronomic characteristics

Name	Ploidy	Apomixis
M. sieboldii	4n	yes
H0909	4n	yes
4551	3n	yes
D2212	4n	yes
4608	3n	yes
M9	2n	no

The attempt to introduce part of the M9 genome in a new crossing population with *M. sieboldii* hybrids was complicated because the latter were deriving from apomictic genotypes. Apomixis is the asexual reproduction through seeds. The fertile seeds resulting from apomictic reproduction contain embryos having a genetic constitution identical to that of the female parent (Schmidt, 1988; Koltunow, 1993). The major problem using these hybrids in the breeding programme was represented by the low vitality of their pollen requiring using them as mother in the crossings. Consequently, several plants identical to the original (motherlike) were obtained because of the high percentage of seeds in which no meiosis occurred. A schema reporting the possible genetic combinations is represented in Figure 3.

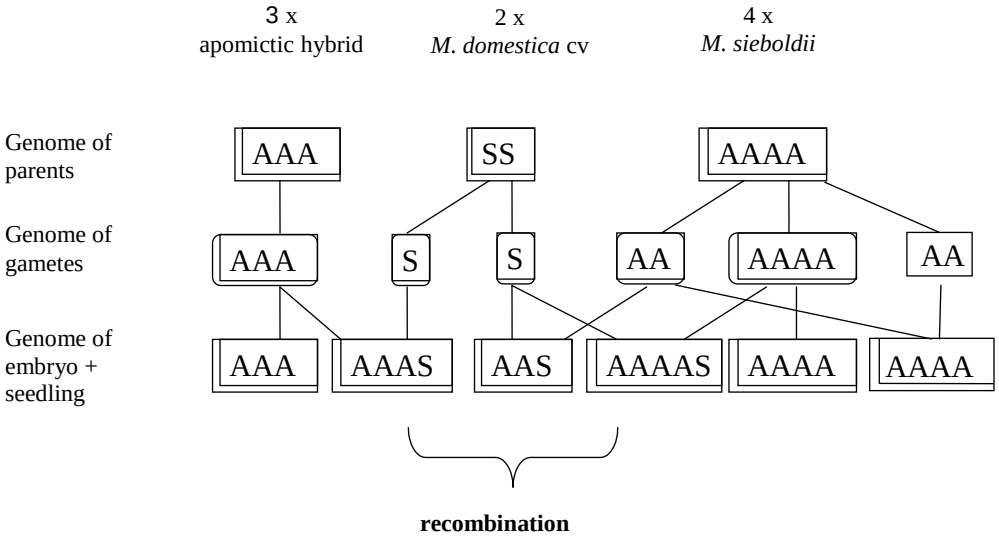


Figure 3 Inheritance of apomixis of triploid and tetraploid apomictic *Malus sp.* genotypes following pollination with diploid cultivar (A: apomixis; S: sexual). The schema is a modified version of the original suggested by Schmidt (1988) (Bisognin *et al.*, 2009)

The breeding process is even more complicated due to the polyploidy associated with the apomictic parents (Table 1) increasing the complexity of the analyses and the selection of the interesting traits.

In order to try to solve such difficulties important efforts have been started in a coordinated way to reach the objective of the development of a new rootstock with sufficient agronomic characteristics and at the same time showing resistance to AP disease. Therefore, a project (scopazzi del melo apple proliferation; SMAP) between the Istituto Agrario di San Michele a/A (IASMA; Italy) as main research partner, the AlPlanta – Institute for Plant Research (AlPlanta – Agroscience; Neustadt a.d. W.; Germany) and the Julius Kühn-Institute (JKI) (Institute for Plant Protection in Fruit Crops and Viticulture, Dossenheim, Germany) was established not only focused on the breeding but also on the study of different aspects of the disease.

The rootstock development was based on the hybrids reported in Table 1 present at the JKI and characterised in the last two decades by the prof. Seemüller. To improve the breeding process the development of genetic markers associated with the resistant traits was initiated in order to follow the segregation in the new crossing population.

The genotypes previously individuated as interesting were the wild species *Malus sieboldii*, the derivated first generation selections 4608 and 4551 and the second generation selections D2118, D2212, H0909 and H0801. These genotypes were first characterised through simple sequence repeat (SSR) analysis in order to define specific markers to be used in the study of the segregating population (Bisognin *et al.*, 2009).

Some of these genotypes (Table 1) were used in the crossings with M9 as father producing the new progenies H0909xM9, 4551xM9, 4608xM9, D2212xM9, *M. sieboldii*xM9. On the crossing population infection trials were conducted in order to observe the resistance trait (Bisognin *et al.,* 2008b). Because this work requires several years to collect enough data from the symptom observations, an *in vitro* system was established to screen the plants after the infection. The possibility to perform the infection experiments in *in vitro* plants offers several advantages.

2.4 *In vitro* system

The possibility to cultivate and maintain AP-infected *Malus* genotypes *in vitro* was shown since the end of the 1980s (Jarausch *et al.*, 1996). Moreover, the transmissibility of the pathogen through *in vitro* grafting was tested and successfully demonstrated (Jarausch *et al.*, 1999a). This was a very important step that provided the possibility to improve the studies on the phytoplasma-plant interaction. The method was shown to work also in other woody plants maintained *in vitro* like *Prunus* (Jarausch *et al.*, 1999a). Therefore, this technique opened new perspectives thanks to the possibility to obtain a high number of infected plants in controlled conditions and in a reproducible way. The studies on *in vitro* plants also showed that this system can be used to screen a large number of different genotypes to test their resistance against the phytoplasma (Bisognin *et al.*,

2008a). The most advantageous points of the *in vitro* system are the economicity of the method, the relative limited space required for the plant maintenance, the sterile conditions allowing the control of the plant infection, the short time needed to obtain plants ready for the infection trials together with the possibility to control important parameters like temperature and light conditions.

Therefore, in the SMAP project it was decided to establish *in vitro* cultures of the hybrids used as starting parental material. All the genotypes were adapted to the *in vitro* conditions obtaining a standardised population ready for the phytoplasma transmission trials (Ciccotti *et al.*, 2008b). Moreover, the system allows the fast screening of the new crossing population in order to improve the data needed to select the most interesting genotypes for the further breeding steps. The possibility to transfer the *in vitro* plants into pots through a rooting protocol complete the set of potentiality of this method allowing the production of *ex vitro* plants that can be further grown and maintained in greenhouse conditions.

Another important aspect of the *in vitro* system is the possibility to conduct gene expression studies to investigate the phytoplasma-plant interaction in standardised conditions. Therefore, in order to limit as much as possible external influences to the pathogen-host system, the *in vitro* plants represent a valid tool allowing the production of enough standardised material in a relative short time.

2.5 Individuation of differentially expressed genes in phytoplasmas diseases

At the beginning of the project the lack of information on which genes are involved in the response to the phytoplasma infection in woody fruit plants, together with the poor availability of information on the apple gene expression were the limiting factors. In the past, the differential display (DD) technique was used for the first time to study the gene expression in *C. roseus* after infection with one of three different mollicutes: *Spiroplasma citri*, stolbur phytoplasma (yellows symptoms) and '*Candidatus* Phytoplasma aurantifolia' (proliferation symptoms) (Jagoueix-Eveillard *et al.*, 2001).

The results showed that in the *S. citri*-plant interaction several genes were differentially expressed like sterol-C-methyltransferase (sterol biosynthesis pathway), transketolase (sugar transport; regeneration of ribulose-1,5-bisphosphate), wall associated kinase, desiccation protectant protein, chlorophyll a/b binding protein and other for which it was not possible to find similarity with known sequences.

In the stolbur-*C. roseus* interaction the genes individuated as differentially expressed were involved in the photosynthesis (subunit III of photosystem I and ribulose 1,5-bisphosphate carboxylase/oxygenase). In grapevine, microarray analysis of 'Blois noir' (stolbur group) phytoplasma-infected plants showed a significant reprogramming of the leaf transcriptome where genes involved in carbohydrate metabolism, in photosynthesis, in secondary metabolism as well as

genes coding for osmotin- and thaumatin-like proteins and for callose synthesis were found to be differentially expressed (Hren *et al.*, 2009).

In another study using again DD analysis carried out in '*Ca*. P. prunorum'-infected *Prunus armeniaca*, genes coding for a heat shock protein (HSP70), a metallothionein (MT) and an aminoacid transporter resulted differentially regulated (Carginale *et al.*, 2004).

Recently, the interaction between *M.* x *domestica* and '*Ca.* P. mali' was investigated through cDNA-AFLP technique showing that several genes involved in stress response, in signalling, in photosystem related pathways and in leaf metabolism were deregulated after infection with the AP-phytoplasma (Aldaghi, 2009).

Another approach for the study of the genes involved in the symptom development was the comparison of these symptoms with mutants of *Arabidopsis thaliana* that showed a similar phenotype focusing the studies on genes known to be involved in the mutant-phenotype development. In this way, Pracros *et al.* (2006) showed that the tomato flower abnormalities associated with the stolbur phytoplasma infection were determined by the differential expression of genes involved in the flower development. However, in this case the analysed genes were selected from a previously characterised *A. thaliana* mutant and not directly identified from the study of the phytoplasma-plant interaction.

2.5.1 cDNA-AFLP: a suitable technique for the selection of differentially expressed genes in '*Ca*. P. mali' infected AP-resistant and –susceptible genotypes

Although form the previous work it was not possible to obtain a clear picture on the phytoplasma-plant interaction all these studies clearly showed the high level of complexity of this system. The particular difficulty to study the pathogen-host relationships and the wide range of metabolisms and modifications that are involved after the phytoplasma infection further complicate the comprehension of the phytoplasma-plant interaction. Therefore, in this thesis a technique that allows a broad identification of differentially expressed genes after the infection was chosen for the analysis of the interaction of '*Ca*. P. mali' with susceptible and resistant *Malus* genotypes. The cDNA amplified fragment length polymorphism transcript profiling (cDNA-AFLP-TP) developed in this version by Breyne *et al* (2002) was adopted because of the potential of this technique in this type of studies (Polesani *et al.*, 2008; Wang *et al.*, 2009; Baldo *et al.*, 2010; Wang *et al.*, 2010). The procedure is based on the enzymatic restriction of the cDNA that has to be analysed followed by the ligation of specific adapters that are compatible with the sticky ends generated by the enzyme cleavage. The known adapter sequences allow the use of specific primers for the amplification of the cDNA fragments obtained. Compared to the original cDNA-AFLP procedure (Bachem *et al.*, 1996) Breyne *et al.* applied a substantial modification: after the production of the first set of cDNA fragments through the first digestion, the fragments with the 3' poly A+ termini are isolated and then used in the second digestion followed by adapter ligation. Therefore, a single

fragment per gene should be obtained giving a fingerprinting (transcript profiling) of the RNA of the organism (Figure 4).

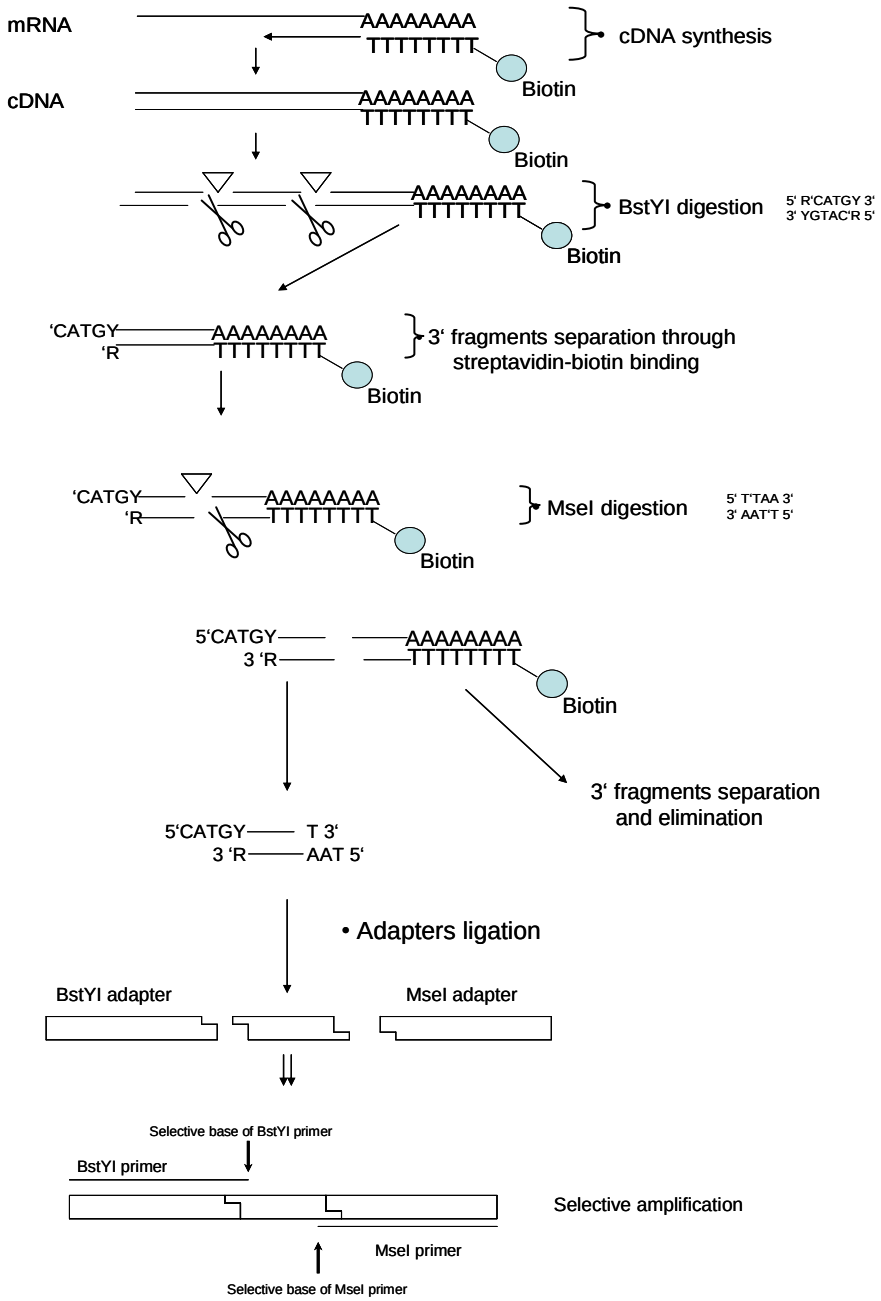


Figure 4 Flowthrough of the cDNA-AFLP-TP procedure

Thanks to the pre-amplification and the selective amplifications a unique pattern characteristic of every genotype can be produced from every primer combination used. The higher the number of primer combinations the higher is the possibility to characterise the differences between genotypes or two different states. In this way, genes that are differentially expressed after the infection because they are activated or deactivated or because they are over- or under-

expressed can be individuated based on the intensity or the presence/absence of the signal in the cDNA-AFLP pattern.

The individuation and isolation of the differentially expressed genes is only the first step of the gene expression analysis. Following, the results have to be validated to confirm the differences in the gene expression level. The most used method in the past was the northern hybridisation but this technique requires a relative high amount of mRNA, it is time consuming and its sensitivity is low compared with other techniques (Montrichard *et al.*, 2003; Gachon *et al.*, 2004). In the last years the real-time quantitative PCR (qPCR) has been used in several studies showing that this is a powerful tool for the gene expression analysis (Bustin, 2002; Nolan *et al.*, 2006) The major advantages of the qPCR are the low amount of starting material needed, the high sensitivity and rapidity of the method and the high reproducibility of the results. Therefore, the cDNA-AFLP bands individuated from the comparison analysis of the infected and healthy plants were studied using the real-time qPCR technique.

2.6 Gene expression analysis by real-time PCR

The real-time qPCR technique allows to quantify the initial amount or to calculate the relative quantity of specific targets in DNA, cDNA or mRNA samples. The principle at the base of the technique is the generation and detection of a fluorescence signal proportional to the quantity of product that is being synthesised (Lee *et al.*, 1993; Livak *et al.*, 1995). The detection of the signal during the PCR allows to follow the kinetic of the reaction. The important parameter obtained from the analysis is the cycle threshold (Ct) value that corresponds to the cycle at which the fluorescence signal reaches an intensity value equivalent to a predefined threshold chosen in the linear phase of the PCR reaction (Figure 5).

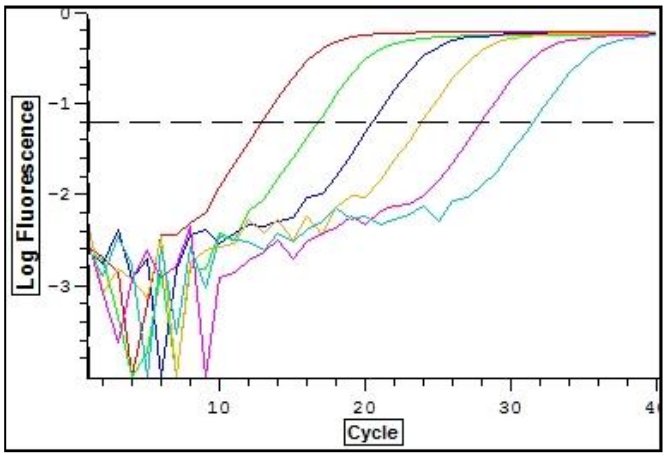


Figure 5 Real-time qPCR plot. The PCR cycles are reported on the X axis while the fluorescence intensity signal is reported on the Y axis. The curves represent different samples. The horizontal discontinued line corresponds to the threshold chosen for the determination of the Ct values. The Ct values correspond to the cycle at which the threshold line meets the real-time curves. Higher Ct values correspond to lower initial amount of target in the sample.

Comparing the Ct value to those of an internal control or of a standard curve, the initial amount of product or the relative amount referred to a defined gene can be determined.
Several methods are available for the generation and detection of the signal in the reaction. The most used are Taqman® probes, molecular beacons and SYBR® Green (Dorak, 2006). The first two solutions have a greater sensitivity and specificity to the target but on the other side they are difficult to design and optimise and therefore not applicable to studies in which several different sequences have to be analysed. The real-time PCR technique is successfully applied for the gene expression studies in several different organisms. In the latest years the availability of instruments at a more affordable price increased the diffusion of the real-time PCR approach (Ginzinger, 2002). Nevertheless, the major part of the gene expression studies carried out using this technique are applied to animal systems for which several tools are available on the market and the treatment of the tissues and materials is easier because of the absence of a cell wall and products like lignine, chitine, cellulose like in plants and fungi. Moreover, if in non-woody plants there is a diffused use of this technique and therefore availability of knowledge on real-time PCR methods, in woody plants the northern approach or a microarray approach were still the major source of information in gene expression studies. Therefore, in this thesis a method for the gene expression analysis was developed in order to investigate the phytoplasma-apple interaction that could be applied also in other gene expression studies in woody plants. Since the availability of a fast method for the analysis of several genes was necessary, it was chosen to adopt the quantitative reverse transcriptase PCR approach (qRT-PCR) (Gachon *et al.*, 2004). The advantages are that there is no need to produce total cDNA and the reaction can be performed directly on RNA specifically for the genes chosen as targets (Montrichard *et al.*, 2003). In addition, lower RNA quality can be overcomed with this method because the production of full length cDNA is not necessary and relatively degraded RNA can still deliver reliable data comparable to those obtained from fully conserved RNA (Fleige and Pfaffl, 2006; Fleige *et al.*, 2006).
Another important aspect in the real-time analysis is the processing of the huge amount of data generated. The raw values obtained from the experiment have to be statistically analysed in order to obtain normalised and reliable data. In this thesis a modified procedure of the method decribed by Peirson *et al.* (2003) (DART) was developed for the data processing obtained from the gene expression analysis.

2.7 Aims of the work

The SMAP project was established as a coordinated research project involving different research groups and institutions. The integrated contributions of different partners were the basis for other research activities in the project. The themes on which it focused were the development of a new '*Ca*. P. mali'-resistant rootstock with sufficient agronomic values, the study of the parameters of AP disease spread with the identification and further investigation of the possible

phytoplasma vectors, and the analysis of the interaction between the phytoplasma and the plant host.

This thesis was part of this larger project focusing on one main objective: the study of the gene expression in the pathogen-plant interaction.

- The main objective was the study of the plant-gene expression in susceptible and resistant genotypes after the infection with '*Ca*. P. mali'. The results should contribute to the elucidation of the so far unknown nature of the resistance mechanism.
- As standardised conditions which limit the influence of parameters not involved in the disease development are a crucial point to perform such analysis, a system that allows the control of several parameters should be adopted. The *in vitro* system is a promising technique that can fulfil the conditions necessary to have a good platform for the gene expression analysis. A set of infected plants should be produced in order to standardise the starting material.
- The power of the cDNA-AFLP method demonstrated in previous studies on gene expression makes this technique an optimal tool to investigate the gene expression after the infection with the phytoplasma. A cDNA-AFLP analysis should be performed comparing healthy and '*Ca*. P. mali'-infected *in vitro* plants.
- The plant-phytoplasma is a complex system dependent on several environmental parameters. In order to obtain reliable information on the gene expression plants maintained in conditions similar to those present in the field should be analysed. Therefore, a set of *ex vitro* plants should be produced in order to confirm the genes that will be individuated in the *in vitro* system.
- The real-time PCR technique is widely employed for the gene expression analysis. Studies on plant/pathogen interaction in other systems demonstrated the power of this method. The possibility to analyse several gene candidates in a fast and reliable way makes the real-time PCR technique an ideal tool for the aim of this thesis.
- Because at the beginning of this thesis there were no reports on the gene expression analysis of the phytoplasma/apple interaction using the real-time PCR technique, a system should be set up for the rapid screening of several genes to analyse their expression after the infection.

3 Materials and methods

3.1 Materials

3.1.1 Plant material

The plant material used in this study derived from a plant collection present at the Julius Kühn-Institute (JKI) Dossenheim (Germany) (Table 2). The gene expression study of the '*Ca*. P. mali'-apple interaction was carried out on plants that were obtained from *in vitro* cultures. The genotypes maintained *in vitro* for the infection trials are reported in Table 3 and are referenced in Bisognin *et al.,* 2008a and Ciccotti *et al.,* 2008b.

Table 2 Genotypes used as plant material

Genotype	Ploidy
Malus sieboldii	**4n**
4551, **4556**: *Malus x domestica* cv Laxton's superb x *Malus sieboldii* selection 4551 and 4556	**3n**
H0909: selection 4556 x M9	**4n**
D2212: *Malus sieboldii* x *Malus x domestica* cv Husmoder x open pollinated	**4n**
GD:*Malus x domestica* cv Golden delicious	**2n**
M9: *Malus x domestica* rootstock cultivar	**2n**

As inoculum source for '*Ca*. P. mali' infected tissue cultures of *Malus* x *domestica* cvs Golden Delicious and Rubinette were used (Bisognin *et al.*, 2008a) (Table 3). Moreover, *ex vitro* plants were generated from these cultures for the analysis of the pathogen-plant interaction at conditions more similar to those present in the orchards.

Table 3 Genotypes maintained *in vitro* for the generation of standardised infected and healthy material

Genotypes used for the study of the '*Ca*. P. mali'-plant interaction	'*Ca*. P. mali' inoculum sources (genotype/strain subtype)
Malus sieboldii	**GD/AT2**
H0909	**Rubinette/AP**
D2212	
GD	
M9	

3.1.2 Chemicals and radioisotopes

The chemicals, biochemicals, tissue culture media and phytohormones used were supplied by the manufacturers listed below:

Applichem (Darmstadt, Germany), Duchefa (Haarlem, The Netherlands), Flucha (Sigma-Aldrich Chemie GmbH, Munich, Germany), Gibco BLR (Invitrogen, Karlsruhe, Germany), Invitrogen (Karlsruhe, Germany), National diagnostics (Biozym Scientific GmbH, Hessisch Oldendorf, Germany), Roth (Carl Roth GmbH, Karlsruhe, Germany), Sigma-Aldrich (Sigma-Aldrich Chemie GmbH, Munich, Germany).

Radiochemical (γ^{33}P ATP 9.25 MBq; 250 µCi; 1,95 MBq/mmol; product code: AH9968-250UCI) and blotting membranes, Hybond N^+-membrane, were supplied by Amersham Biosciences (GE Healthcare Europe GmbH, Freiburg, Germany).

3.1.3 Enzymes

The restriction enzymes, Taq DNA polymerase, mesophilic DNA polymerase and DNA modifying enzymes used were supplied from Eppendorf (Hamburg, Germany), Fermentas-MBI (St.Leon-Rot, Germany), Invitrogen (Karlsruhe, Germany), New England Biolabs (NEB) (Frankfurt am Main, Germany), USB (GE Healthcare Europe GmbH, Freiburg, Germany).

3.1.4 Bacterial strains

The cloning procedures were carried out using the *Escherichia coli* (*E. coli*) bacteria strain:

DH5α: F^- φ80d*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 λ^- *thi*-1 *gyr*A96 *rel*A1 (Invitrogen, Karlsruhe, Germany)

3.1.5 Vectors

pUC19	Yanisch-Perron et al 1985 (Fermentas, St.Leon-Rot, Germany)
pNEB205A	pUC19 derived vector, (NEB, Frankfurt am Main, Germany)

3.1.6 Buffers and media

LB:

10 g/L tryptone, 10 g/L NaCl, 5 g/L yeast extract, pH 7.5 adjusted with NaOH, 15 g/L agar were added for solid medium

SOC:

20 g bacto-tryptone, 5 g bacto-yeast extract, 0.5 g NaCl, 2.5 ml 1M KCl, ddH2O to 1 L pH 7.0 adjusted with NaOH, autoclave to sterilize, add 20 ml of sterile 1M glucose immediately before use

MLO6 pH 5.6 (Jarausch *et al.*, 1999a; Ciccotti *et al.*, 2008b):

4.46 g/L Murashige and Skoog (MS) modified as described by Van der Salm *et al*. (1994) with macro & microelements + vitamins; 1 mg/L thiamine, 30 g/L sucrose, 8 g/L micro agar

The following hormones are added at a temperature of 50°C:

0.05 mg/L auxine IBA
0.1 mg/L gibberelline GA3
1 mg/L cytokine BAP

XT buffer:

0.2 M $NaBO_4$ • 10 H2O, 30 mM EDTA, 1% (w/v) SDS (w/v), 1% (w/v) sodium deoxycholate

TAE buffer:

40 mM Tris-acetate pH 8.0, 0.05 M EDTA

TBE buffer:

89 mM Tris-HCl pH 8.0, 89 mM boric acid, 2 mM EDTA

TE buffer:

10 mM Tris-HCl pH 8.0, 1 mM EDTA

TEXdir buffer:

CTAB 2 %, 1.4 M NaCl, 20 mM EDTA pH 8.0, 100 mM Tris-HCl pH 8.0, 2 % PVP25

P1:

25 mM Tris-HCl pH 8.0, 50 mM glucose, 10 mM EDTA pH 8.0

P2:

0.1 M NaOH, 1% SDS

P3:

3M KOAc, 5M glacial acetic acid

3.2 Methods

3.2.1 Plant tissue culture

The tissue cultures used in this thesis were those described by Bisognin *et al.,* 2008a and Ciccotti *et al.,* 2008b (Table 3). The material necessary for the experiments was prepared from these cultures that were maintained and propagated into tubes containing ten ml of MLO6 media (3.1.6) and placed in usual growth chamber at 23/18 °C ± 1°C day/night and 16h photoperiod under cool-white fluorescent lights (60 µE m^{-2} s^{-1}). Plantlets developed from the transferred shoot were subcultured every 4 - 6 weeks (Jarausch *et al.*, 1996; Ciccotti *et al.*, 2008b).

3.2.1.1 *In vitro* grafting

The phytoplasma-infected *in vitro* plants were produced by *in vitro* graft-inoculation as described in Jarausch *et al.* (1999a). For each genotype a minimum of 10 healthy shoots of adequate size were chosen for the grafting with phytoplasma-infected tips. The latter were selected from '*Ca*. P. mali'-infected *in vitro* plants that were propagated and maintained as inoculums reservoir (Table 3). The grafting procedure used to infect the plants was carried out joining the infected tip to the healthy shoot. Tip and shoot had to be prepared producing a small cut in the middle of the shoot and modifying the extremity of the tip in a wedge form in order to produce a compatible union (graft) between the two plants (Figure 6). After one month of graft contact, the tip was separated from the shoot and for each sample the graft-union strength was annotated. The tips were stored at -20°C while the shoots were further subcultured for another month in MLO6 medium (3.1.6). After this period, the plants were tested through PCR (3.2.4) for the presence of '*Ca*. P. mali'. For those plants that resulted PCR negative and that had a strong graft-union the tip was PCR-tested for the presence of the phytoplasma. In this way, it was possible to calculate the transmission rate excluding from the analysis the plants that were negative because grafted with a healthy tip.

a) b)

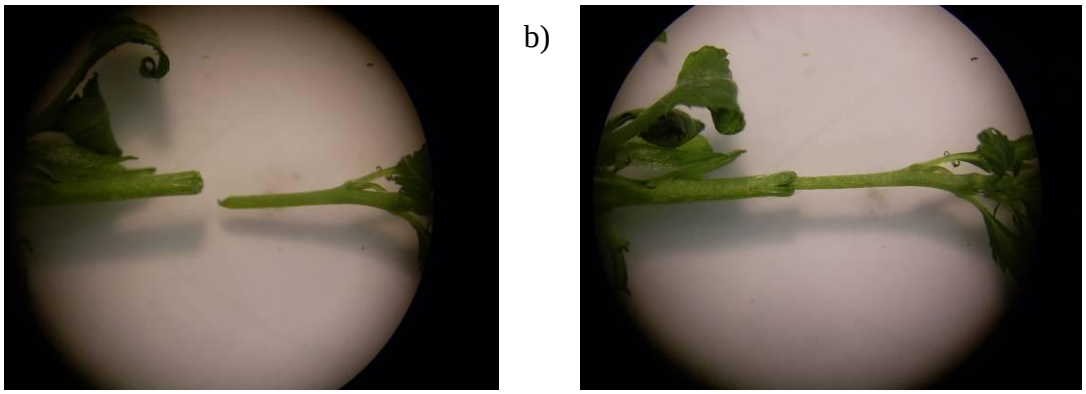


Figure 6 a) preparation of the plant shoot and the '*Ca*. P. mali'-infected tip; b) union of the plant shoot and the tip

3.2.1.2 Rooting and acclimatization

Homogeneous shoots, longer than 2 cm, were excised from proliferating cultures: 10-30 explants per genotype were incubated in a solution containing 30 g/l sucrose and 5 mg/l IBA in darkness at 24 ± 1 °C for 4 days. After dark induction, explants were transferred to the light on agarised (0.6%) auxin-free medium with half strength MS salts, without vitamins and with 58.4 mM sucrose (Ciccotti *et al.*, 2008a; Ciccotti *et al.*, 2008b). After 25 days rooted plantlets were removed, rinsed in tap water, transplanted in sterilised potting mixture of peat and agriperlite (15%) and grown in greenhouse benches under closed environments and under natural daylight conditions.

Humidity (RH), maintained near saturation, was reduced over a period of 4 weeks by gradually opening the lid. A weak solution of fungicide was sprayed on the plants to prevent fungal contamination. Acclimatized plantlets were then transplanted to plastic square pots and maintained in greenhouse. The plants obtained with this procedure were defined as *ex vitro* plants.

3.2.2 RNA isolation from plant tissue

3.2.2.1 Preparation of phloem tissue from roots and branches

The phloem tissue used for the RNA extraction was prepared from healthy in field-grown plants. Opposing excisions were made in the cortex of roots and branches and then peeled away. The phloem tissue was then removed by scratching with a scalpel and the material immediately frozen in liquid nitrogen and stored at -70°C.

3.2.2.2 Collection of the material from *in vitro* and *ex vitro* plants

For each genotype and healthy state 7-8 *in vitro* shoots with homogeneous size were respectively collected and pooled and used as material in the RNA extraction. For the analysis in *ex vitro* plants, 4-5 fully developed leaves where collected from 3-4 different plants for each genotype and healthy state respectively. As root material several primary and secondary roots were collected and pooled from 3-4 different plants for each genotype and healthy state. After sampling the material was immediately frozen in liquid nitrogen and stored at -70°C.

3.2.2.3 RNA extraction procedure

Total RNA extraction was performed with a modified hot borate method as described in Moser *et al.* (2004).

The following reagents have to be added to the XT buffer (3.1.6) immediately before its use in the extraction:

- 2% ß-mercaptoethanol (or DTT 10 mM)
- 0,5% spermidin
- 1% Nonidet P-40 (IGEPAL) (or Tween 20)
- 2% (w/v) PVP

3.5 ml complete buffer/g tissue is required.

The extraction is carried out in 2 ml Eppendorf tubes.

Procedure:

1) Pre-warm the buffer to 80 °C in waterbath
2) grind the tissue with liquid nitrogen, as fine as possible. Transfer 0.4 g of powder in pre-chilled 2 ml Eppendorf tubes

3) immediately add the hot buffer to the powder by slow flow to avoid foam formation. Vortex briefly to favour the contact of the powder with the buffer. Incubate the Eppendorf tube for 5 minutes at 80°C in a thermoblock
4) add 2.5 mg/g tissue of proteinase K (Sigma 100 mg; 20 mg/ml solution). Incubate 1 h at 42°C in the thermoblock. Mix every 10 minutes
5) add 2 M KCl at final concentration of 160 mM. Mix gently. Place the Eppendorf tube 45 min in ice (formation of an insoluble precipitate)
6) centrifuge 15 min at 12000 rpm (15000xg) at 8°C
7) transfer clarified upper liquid in a new Eppendorf tube and add 8 M LiCl at final concentration of 2 M and 1% ß-mercaptoethanol
8) incubate overnight at 4°C (on ice) for RNA precipitation
9) centrifuge 25 min at 12000 rpm (15000xg) at 8°C
10) discard liquid and wash the pellet (generally relatively pigmented) once with 1 ml cold 2M LiCl
11) centrifuge at 13000 rpm (21000xg) at 8°C for 15 minutes (the pellet must be compact)
12) if the pellet is not sufficiently clean, repeat the steps 10-11 with 1 ml of cold 2 M LiCl
13) dissolve the pellet in 600 µl of 10 mM Tris-HCl pH 7.5
14) add 1/10 volume of 2 M KAc pH 5.5. Put the Eppendorf tube 10 minutes on ice
15) centrifuge 15 min at 13000 rpm (21000xg) at 8°C
16) transfer the liquid (clean and transparent) in a new Eppendorf tube, and add 0.9 volumes of cold isopropanol. Incubate 1 h at –20°C
17) centrifuge 20 min at 13000 rpm (21000xg) at 8°C
18) wash the pellet with 1 ml cold 80 % EtOH
19) centrifuge 15 min at 13000 rpm (21000xg) at 8°C
20) dry the pellet on ice
21) resuspend the pellet in 100 µl of RNAse free water

3.2.3 DNA isolation from plant tissue

The total DNA of plant tissues (leaves, petiols, root phloem, branch phloem or *in vitro* plant) was extracted with the method described below (W. Jarausch, personal communication):

Plant material: about 0.5 g plant tissue + 3 ml TEXdir buffer (3.1.6)

1) homogenise the material in ELISA bag in a homogenisator (Bioreba)
2) transfer 1.5 - 2 ml of liquid in 2 ml Eppendorf tube
3) centrifuge 4 min at 2000 rpm (2200xg)
4) transfer 1 ml of the supernatant in a new 2 ml safe cap Eppendorf tube
5) incubate 30 min – 60 min at 60°C (waterbath/heatblock), spin to take down the condensed water at the cap
6) add 1 ml chloroform

7) mix by inverting the tube 1 min
8) centrifuge 5 min at 13000 rpm (16000xg)
9) transfer 900 µl of aqueous phase in a new 1.5 ml Eppendorf tube
10) add 0.8-volumes isopropanol and vortex
11) incubate 15 min at RT
12) centrifuge 15 min at 13000 rpm (16000xg)
13) discard the supernatant
14) wash the pellet in 500 µl of 70 % EtOH, vortex
15) centrifuge 10 min at 13000 rpm (16000xg)
16) discard the supernatant and dry the pellet in a speedvac
17) dissolve the pellet in 100 µl H_2O

3.2.4 General PCR

Several PCR reactions were optimised for the specific procedure or protocol carried out in this thesis. However, when not specified in the method descripted a standard PCR procedure was performed as follows:

PCR mix:

1 µl primer forward	(10 µM)
1 µl primer reverse	(10 µM)
0.5 µl dNTPs	(5 mM each)
2 µl 10X Taq buffer	(Eppendorf)
0.1 µl Taq	(5 U/µl;Eppendorf)
X µl template	(10 ng plasmid DNA- 100 ng genomic DNA or RNA)
H_2O to 20 µl	

PCR reaction parameters:

95°C 4 min	
95°C 30 sec	} 35 cycles
X°C 30 sec	
72°C Y sec	
72°C 4’	
16°C oo	

X is the optimal annealing temperature of the primer pair used
Y is the elongation time that depends on the insert size (1000 bp/minute was used as parameter for the calculation of Y)

3.2.5 Gradient PCR

The optimal annealing temperature of each primer was empirically defined through the application of a gradient PCR. A specific thermal cycler was used (Epgradient; Eppendorf) in which different temperatures are applied depending on the position of the sample in the thermoblock. The temperature gradient, increasing from left to right, was applied using a range between 52°C and 64°C. Each position in the block has its own temperature. The same reaction mix and the cycle parameters described in 3.2.4 were used.

3.2.6 RT-PCR

The primers developed for the real-time PCR analysis were tested on RNA using a one-step RT-PCR protocol (one-step RT-PCR kit; Invitrogen) to verify presence of the target in the RNA and to evaluate the quality of the primers. Moreover, with this method were produced the PCR products that had to be sequenced to control if the right target was amplified (kit containing the Platinum Taq High Fidelity; Invitrogen).

Reaction mix:

0.5 µl enzyme mix (Superscript III and Platinum Taq or Platinum Taq High Fidelity; Invitrogen)
0.5 µl dNTPs (5 mM each)
1 µl primer forward (10 µM)
1 µl primer forward (10 µM)
2 µl 10X RT buffer (Invitrogen)
X µl RNA (100 ng)
H_20 to 20 µl

PCR parameters

RT step 55°C 10 min
95°C 5 min
95°C 30 sec
60°C 30 sec } 40 cycles
72°C 30 sec
72°C 4 min
16°C oo

3.2.7 Phytoplasma detection

The presence of the phytoplasma in the infected plants was detected through PCR as described in Jarausch *et al*. (1994a) using the specific primers AP3 (5'-GAAACATGTCCTATTGGTGG-3') / AP4 (5'-CCAATGTGTGAAATCTGTAG-3').

The DNA extracted from the plant tissue was used as template in the PCR mix prepared as follows:

2 µl primer AP3 (10µM)
2 µl primer AP4 (10µM)
0.5 µl dNTPs mix (5 mM each)
2 µl 10x PCR buffer (Eppendorf)
0.2 µl Taq DNA polymerase (5 U/µl; Eppendorf)
12.3 µl H_2O
1 µl of DNA
20 µl final volume

PCR reaction conditions:

95°C 5 min
95°C 15 sec
57°C 15 sec } 40 cycles
72°C 30 sec
72°C 4 min
16°C oo

The PCR products were loaded on a 2 % agarose gel (3.2.21) and the electrophoresis performed at 100 V for 30 min

3.2.8 Restriction reactions

The restriction reactions were performed as follows:

5 unit restriction enzyme
2 µl 10X enzyme buffer
300 ng DNA
Final volume 20 µl

For double digestion if the enzymes are compatible:

5 unit restriction enzyme A
5 unit restriction enzyme B
2 µl 10X enzyme buffer (optimal for both enzymes)
0.2 µl 100X BSA
300 ng DNA
Final volume 20 µl

Digestion overnight at the temperature recommended by the supplier.

3.2.9 Cloning of PCR products

3.2.9.1 TA cloning

The pUC19 vector was modified for the TA cloning of PCR bands. The preparation of the linearised pUC19 with T extremities was carried out as follows:

5 µl pUC19 (1 µg/µl)
2 µl EcoRV (10 U/µl)
4 µl 10x EcoRV buffer
29 µl H_20
Final volume 40 µl

Incubate 4 h at 37°C

3 µl of the reaction mix were loaded on a 0.8 % agarose gel (3.2.21) in order to control if the vector was linearised.
The addition of the dTTP at the 3' extremities of the vector was performed as follows:

0.4 µl Taq DNA polymerase (5 U/µl Eppendorf)
10 µl dTTP (10 mM)
Final volume 50 µl

The modified vector was then purified using the QIAquick PCR purification kit (Qiagen)
For PCR products destinated to be cloned a final step of 10 minutes at 72°C in the PCR cycles was added.
The ligation reaction was then performed using:

50 ng of T modified pUC19 vector
PCR product in a molar ratio 3:1
1.5 µl 10X T4 ligase buffer (NEB)
1 µl T4 ligase (400 NEB units/µl; NEB)
Final volume 15 µl

3.2.9.2 USER PCR product cloning method

In order to have a rapid and efficient cloning of PCR products the USER (uracil-specific excision reagent) Friendly Cloning Kit from NEB was chosen (NEB #E5500). This method combines the use of the plasmid pNEB206A (pUC19 derived vector) with the 5' USER modified PCR primers. The modification consists in a clamp at the 5' end of the primers conferring this final structure:

Forward Primer: 5´-[GGAGACA**U** + (restriction site) + priming sequence]

Reverse Primer: 5´-[GGGAAAG**U** + (restriction site) + priming sequence]

The PCR reaction is performed at the conditions normally used for the unmodified specific primer. The uracil present in the primer sequence allows the enzymatic modification by the USER enzyme (a mix of uracil DNA glycosylase and DNA glycosylase-lyase Endo VIII). After the USER reaction the PCR products present sticky ends because of the excision activity of the enzyme mix. The extremities are complementary to 5' and 3' extremities of the linearised pNEB206A allowing the cloning of the PCR products.

The protocol of the USER cloning is carried out as follows:

10 µl of USER-PCR product (minimum 50 ng)
1 µl of USER enzyme mix
1 µl of pNEB206 A vector
15 min at 37°C
15 min at RT

3.2.9.3 Preparation of chemically competent *E. coli* cells

Chemically competent *E. coli* cells were prepared as follows:

1) a single *E. coli* colony is grown overnight in LB media
2) the overnight culture is inoculated in a 100 times volume of LB media and grown until OD_{600} value reaches 0.4-0.6
3) centrifuge the culture at 5000 rpm for 5 min at 4°C
4) gently resuspend the pellet in 1/4 original volume ice cold $CaCl_2$
5) chill on ice for 15 min
6) centrifuge the suspension at 5000 rpm for 5 min at 4°C
7) gently resuspend the pellet in 1/25 volume ice cold $CaCl_2$ containing 15% glycerol
8) freeze the cells in 100 µl aliquots directly in liquid nitrogen
9) keep the aliquots at -70°C

3.2.9.4 Heat shock transformation of *E. coli*

2 µl of the ligation reaction is added to 50 µl of chemically competent *E. coli* that are transformed using a heat shock procedure:

1) incubate the competent cells 30 min on ice
2) heat shock at 42 °C for 45 sec
3) add 600 µl LB medium to the cells
4) incubate 1h at 37°C in agitation (300 rpm)
5) plate 100 µl on previously prepared LB agar plates containing 100 µg/ml ampicillin and 40 µl of X-gal (40 mg/ml)

3.2.10 Colony PCR

The colonies were picked with a toothpick and replicated on a new LB agar plate containing 100 µg/ml ampicillin. In order to check the presence of the insert the toothpick was then swirled for 5 seconds in a PCR mix prepared as described in 3.2.4 using the specific primers of the insert that had to be checked.

The PCR conditions were :

95°C 10 min
95°C 30 sec
X°C 30 sec } 35 cycles
72°C Y min
72°C 4 min
16°C oo

X is the optimal annealing temperature of the primer pair used
Y is the elongation time that depends on the insert size (1000 bp/minute was used as parameter for the calculation of Y).

3.2.11 Plasmid miniprep

The plasmid minipreps were performed starting with an overnight culture prepared by inoculating a colony in 2 ml of LB + antibiotic (ampicillin 50 µg/ml).
The miniprep procedure was carried out as follows:

1) centrifuge the bacteria 5 min at 13000 rpm (16000xg) at RT
2) discard the supernatant and add 100 µl of P1 buffer. Vortex and wait 5 min
3) add 200 µl of P2 buffer. Mix gently but thoroughly and wait 5 min at RT
4) add 150 µl of P3 buffer. Mix gently but thoroughly and incubate 5 min on ice
5) centrifuge 5 min at 13000 rpm (16000xg)
6) transfer 400 µl of the supernatant to a new 1.5 ml Eppendorf tube and add 800 µl of cold 100 % EtOH. Incubate 2 min at RT
7) centrifuge 5 min at 13000 rpm (16000xg) at RT
8) remove the supernatant and add 1 ml of cold 70 % EtOH
9) centrifuge 5 min at 13000 rpm (16000xg) at RT
10) dry the pellet and resuspend in 20 µl of H_2O
11) the presence of the insert is checked through restriction analysis using appropriates restriction enzymes (3.2.8).

3.2.12 Poly $(A)^+$ mRNA isolation

Poly $(A)^+$ mRNA was isolated following a modified procedure of the NEB protocol for streptavidin magnetic beads (NEB #S1420S):

Elution buffer (EB):	10 mM Tris-HCl (pH 7.5), 1 mM EDTA
Low salt buffer:	0.15 M NaCl, 20 mM Tris-HCl (pH 7.5), 1 mM EDTA
Wash/binding buffer:	0.5 M NaCl, 20 mM Tris-HCl (pH 7.5), 1 mM EDTA

1) prepare a 65°C water bath
2) pre-warm the elution buffer in a 70°C water bath
3) place the low salt buffer in an ice bath
4) dissolve 1.0 A_{260} unit of biotin-$(dT)_{18}$ (NEB #S1325S) in 500 µl of 1x wash/binding buffer. Final concentration 8 pmol/µl
5) aliquot 125 µl (500 µg) of streptavidin magnetic beads per 100 µg of total RNA into a clean RNase-free microcentrifuge tube. Add 100 µl of 1x wash/binding buffer and vortex to suspend the beads. Apply the magnetic separator to the side of the tube for approximately 30 seconds. Remove and discard the supernatant
6) add 25 µl of biotin-$(dT)_{18}$ solution to the magnetic beads and vortex to suspend the beads. Incubate at room temperature for 5 minutes with occasional agitation by hand. Apply the magnetic separator then remove and discard the supernatant
7) wash the beads by adding 100 µl of 1x wash/binding buffer. Vortex to suspend, then apply the magnetic separator and discard the supernatant. Repeat the wash
8) add 5 µl of 10x wash/binding buffer to the volume containing 100 µg of total RNA and adjust the volume to 50 µl with RNAse free H_2O. Heat at 65°C for 5 minutes then quickly chill in an ice bath for 3 min
9) add the total RNA sample to the previously prepared magnetic beads. Vortex to suspend the particles then incubate at room temperature for 10 minutes with occasional agitation by hand
10) apply the magnetic separator and remove the supernatant. Add 100 µl of wash/binding buffer, vortex to suspend the beads. Apply the magnetic separator then remove and discard the supernatant. Repeat the washing step with fresh wash/binding buffer
11) add 100 µl of cold low salt buffer to the beads, vortex to suspend. Apply the magnetic separator then remove and discard the supernatant
12) add 25 µl of pre-warmed elution buffer, vortex to suspend the beads then incubate at room temperature for 2 minutes
13) apply the magnetic separator then transfer the supernatant to a clean RNase-free microcentrifuge tube

14) repeat the elution step with 25 µl of fresh elution buffer. Apply the magnetic separator and add the supernatant to the previously eluted mRNA.

3.2.13 cDNA synthesis

For the cDNA synthesis and the cDNA-AFLP analysis the procedures described by Breyne *et al.* (2003) were followed with some modifications and adaptations.

- **First strand cDNA synthesis**

20 µl poly(A)$^+$ mRNA (~1.0 µg)
1 µl oligo-dT$_{25}$-biotynilated (700 ng/µl; Invitrogen)
4 µl 10x SuperscriptII buffer (Invitrogen)
4 µl 0.1 M DTT
2 µl dNTPs 10 mM (2.5 mM each)
1 µl Superscript II (200 U/µl; Invitrogen)
8 µl H_2O

Incubate 2 hr at 42°C

- **Second strand synthesis**

40 µl first strand reaction mixture
16 µl 10x *E. coli* ligase buffer (NEB)
3 µl dNTPs 10 mM (2.5 mM each)
6 µl 0.1 M DTT
15 U *E. coli* ligase (NEB)
50 U *E. coli* DNA polymerase I (NEB)
1.6 U RNAse-H (Fermentas-MBI)
H_2O to 160 µl

Incubate 1 hr at 12°C and 1 hr at 22°C

Extract the reaction mixture once with phenol/chloroform and purify the cDNA using a Qiaquick spin column (Qiaquick PCR purification kit, Qiagen). Check the quality and yield of cDNA by agarose gel electrophoresis.

3.2.14 cDNA-AFLP

AFLP template preparation (using BstYI and MseI)

- First digestion

30 µl cDNA (500 ng)
10 U BstYI
4 µl 10X BstYI buffer
H_2O to 40 µl

Incubate 2 hrs at 60 °C

- Immobilization of 3'-terminal cDNA fragments on streptavidin magnetic beads

The biotinylated 3'-end fragments are separated from the non-biotinylated fragments by binding to streptavidin magnetic beads (S1420S; NEB) as follows:

STEX buffer (1X): 1 M NaCl, 10 mM Tris, 1 mM EDTA, 0.1 % Triton X-100, pH 8.0

1. wash 10 µl streptavidin magnetic beads once in 100 µl 2x STEX and resuspend in 40 µl 2x STEX (the same volume as the restricted cDNA)
2. add the beads to the restricted cDNA to give a final volume of 80 µl
3. incubate the mixture at room temperature for 30 minutes with gentle agitation
4. collect the beads with the magnetic separator (S1506S; NEB) wash once with 100 µl 1x STEX and transfer to a fresh tube
5. wash 4 additional times with 100 µl 1x STEX
6. resuspend the beads in 30 µl TE buffer and transfer again to a fresh tube.

- Second digestion

Add to the 30 µl beads-suspension:

10 U MseI
4 µl 10X MseI buffer
H_2O to 40 µl

Incubate 2 hrs at 37 °C with gentle agitation

Collect the beads using the magnetic separator and transfer the supernatant, containing the liberated template fragments, to a clean tube for the subsequent adapter ligation reaction.

- Adapter preparation

All adapters (BstYI, MseI) are non-phosphorylated oligonucleotides

For adapter BstYI, mix oligo Bst-F and Bst-R such that an adapter is obtained at a concentration of 5 pmol/µl.

Bst-F: 5' - CTCGTAGACTGCGTAGT - 3'
Bst-R: 5' - GATCACTACGCAGTCTAC - 3'

For adapter MseI, mix oligo Mse-F and Mse-R such that an adapter is obtained at a concentration of 50 pmol/µl.

Mse-F: 5' - GACGATGAGTCCTGAG - 3'
Mse-R: 5' - TACTCAGGACTCAT - 3'

- Adapter ligation

Add to the supernatant (containing the template fragments; 40 µl) a mix of 10 µl containing:

1 µl BstYI adapter (5 pmol)
1 µl MseI adapter (50 pmol)
1 µl 10 mM ATP (Fermentas-MBI)
2 µl 5 x DNA ligase buffer (Invitrogen)
1 U T4 DNA ligase (Invitrogen)
10 U BstYI
H_2O to 10 µl

Incubate >= 3 hrs at 37°C

Pre-amplification with a selective Bst primer

Dilute the adapter ligation 10-fold with TE buffer and use 5 µl as a template in the pre-amplification. BstT0 and BstC0 are combined with the Mse0 primer

Primers:

BstT0:	5' - GACTGCGTAGTGATCT - 3'
BstC0:	5' - GACTGCGTAGTGATCC - 3'
Mse0:	5' - GATGAGTCCTGAGTAA - 3'

The reaction contains:

5.0 µl 10-fold diluted adapter ligation
1.5 µl BstT0 or BstC0 primer (50 ng/µl = 75 ng)
1.5 µl Mse primer (50 ng/µl = 75 ng)
2.0 µl 5 mM dNTPs (0.2 mM final concentration of each dNTP)
1 U Taq DNA polymerase (5 U/µl; Eppendorf)
5.0 µl 10x PCR-buffer (Eppendorf)
H_2O to 50 µl

PCR amplification:

94°C 30 sec.
56°C 60 sec.
72°C 60 sec
} 20 cycles

The pre-amplification is checked by running 10 µl of the reaction mixture on a 1 % agarose gel (3.2.21).

Selective amplification reaction using a labeled primer

- Primer labeling

For one single AFLP reaction, 5 ng of labeled primer is needed.

0.1 µl of a 50 ng/µl selective Bst(C orT)+N-primer solution
0.1 µl ^{33}P-g-ATP (»2000 Ci/mmol » 50 pmol; Amersham)
0.05 µl 10x T4-kinase buffer (USB)
0.2 units T4-kinase (USB)
H_2O to 0.5 µl

0.5 µl final volume yields a labeled primer at a concentration of 10 ng/µl

Incubate 60 min at 37°C and, subsequently, 10 min at 70°C (inactivation of the kinase).
The actual reaction volumes depend on the number of AFLP amplifications that need to be performed with the selective primer.

- PCR reaction

For each amplification, 5 µl of a 600-fold dilution of the pre-amplification mixture is used as template.

A single PCR reaction contains:

5.0 µl pre-amplification mix (diluted 600-fold in $T_{10}E_{0.1}$)
0.5 µl labeled Bst(C or T)+N-primer (10 ng/µl --> 5 ng)
0.6 µl unlabeled MseI+N-primer (50 ng/µl --> 30 ng)
0.8 µl 5 mM dNTPs
2.0 µl 10x PCR-buffer (Eppendorf)
0.6 U Taq DNA polymerase (5 U/µl; Eppendorf)
H_2O to 20 µl

PCR amplification:

94°C 5 min
94°C 30 sec
65°C 30 sec
72°C 60 sec
} cycle 1

cycle 2 to 13: lower the annealing-temperature each cycle by 0.7°C during 12 cycles

94°C 30 sec
56°C 30 sec
72°C 60 sec
} 23 cycles:

Polyacrylamide gel analysis

After the selective PCR amplification, the samples were analyzed on 6% polyacrylamide gels (3.2.21) using the Sequigel system (Bio-Rad). Gel electrophoresis was performed as described in the common AFLP analysis (Vos *et al.*, 1995). After the run, the gel was vacuum-dried on a Whatman paper and exposed on a Biomax MR autoradiogram film (Kodak).

cDNA-AFLP band isolation

The bands were selected on the autoradiographic film that was then used as guide for the isolation of the cDNA-AFLP fragments. Therefore, the position of the band on the blot was individuated overlapping the film and excising the correspondent zone from the Whatman paper (and therefore the gel slice). The obtained fragment was then eluted overnight in 100 µl of ddH_2O.

3.2.15 Sequencing

The sequencing of the cloned inserts was performed using the sequencing kit BigDye (ABI). The capillary electrophoresis analysis was carried out by the company GENterprise (Gesellschaft für Genanalyse und Biotechnologie mbH; Mainz; Germany).

Sequence reaction mix:

2 µl BigDye v 3.0
2 µl 5X BigDye buffer
1 µl primer for sequencing (10 µM)
0.5-5 µl DNA (between 170-350 ng)
0-4.5 µl H20
Final volume 10 µl

PCR parameters:

95°C 1 min
95°C 10 sec
55°C 4 min } 30 cycles
16°C oo

For some sequences, the capillary electrophoresis analysis was self performed using ABI prism 3100 sequencer from ABI.

The sequencing of PCR products was carried out from the company 4BaseLab (Reutlingen; Germany) directly on 10 µl of reaction using the reverse or forward specific primers.

3.2.16 Sequencing data analysis and software

The sequences were analysed using the BLAST web form and the tools present at the NCBI (www.ncbi.nih.gov), Expasy (www.expasy.org) and TIGR (www.tigr.org). Analysis of the sequence chromatograms was performed with the software Chromas (Technelysium Pty Ltd) and the freeware suite BioEdit 7.0.5 (developed by Tom Hall; Ibis Therapeutics). The development of SNP markers was carried out using the software Staden package (Staden *et al.*, 2003).

3.2.17 Macro-array production

3.2.17.1 Low density macro-array production for the first hybridisation screening experiment

The low density macro-array was produced spotting the PCR products obtained from the re-amplification of the cDNA-AFLP fragments. The specific primers associated with the corresponding isolated band were used for the PCR reaction carried out as reported in Diehl *et al.*(2002) and performed in 96 well PCR plates:

10 mM	Tris- HCl pH 8.3 (stock : 100 mM)
2.25 mM	$MgCl_2$ (stock : 50 mM)
50 mM	KCl (stock : 500 mM)
0.2 mM	dNTPs (stock : 10 mM)
1.5 M	Betaine (stock : 3 M)
0.25 µM	Specific Bst-C+N primer
0.25 µM	Specific MseI+N primer
0.1 mM	Cresol Red (stock : 50 mM)
0.2 µl	Taq Polymerase (Eppendorf 5 U/µl)
1 µl	cDNA-AFLP band

Final volume 40 µl

PCR parameters:

94°C 10 min	
95°C 30 sec	} 35 cycles
56°C 30 sec	
72°C 40 sec	
72°C 4 min	
16°C oo	

Spotting procedure for the low density macro-array production

The spotting procedure was carried out using a 96 pins tool (V&P scientific; product #VP407) with a spotting volume of 1.5 µl per pin. The presence of the cresol red allowed the control of the procedure through the red coloration of the spot.

The production of the macro-array was carried out using a 4x4 geometry as spotting grid with 5 positions spotted in double. A total of 96*10 spots were produced on a 10 cm x 14 cm filter membrane. The geometry is reported below:

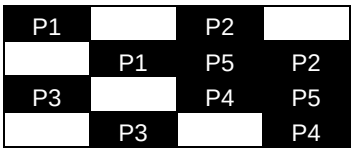


Figure 7 Base geometry of the macroarray. P1= plate 1; P2 = plate 2; P3= plate 3; P4= plate 4; P5= plate5; the white squares are empty

The preparation of the membrane and the spotting procedure were carried out as follows:

Denaturation solution: 1.5 M NaCl, 0.5 M NaOH

Neutralisation solution: 1.5 CaCl, 0.5 M Tris-HCl pH 7.2

1) place the filter membrane (Hybond N+; Amersham) in the denaturation solution for 5 min
2) wet 3 layers of whatman paper with the denaturation solution discarding the excess of liquid
3) place the membrane on the whatman paper avoiding the formation of air bubbles under the membrane. Avoid the presence of an excess of liquid on the membrane
4) perform the spotting procedure
5) place the membrane for 5 min on 3 layers of whatman paper previously immerged in neutralisation solution and eliminating the excess of liquid
6) transfer the membrane on a dryed whathman paper
7) leave the membrane dry
8) the DNA is fixed on the membrane through backing 2 h at 80 °C

The hybridisation procedure is described in 3.2.18.

3.2.17.2 High density macro-array production for the second hybridisation screening experiment

Cloning of cDNA-AFLP bands

The eluted cDNA-AFLP bands were cloned for the preparation of a macro-array. The procedure used was carried out as follows:

1) re-amplification of all isolated cDNA-AFLP bands with the correspondent USER modified selective primers (3.2.9.2). For all the primers one common PCR cycle is used:

94°C	5 min	
94°C	30 sec	} 30 cycles
56°C	30 sec	
72°C	60 sec	

2) check the PCR products on a 1 % agarose gel (3.2.21)
3) mix 5 µl of each PCR (pooling) carried out with the same primer combination
4) take 10 µl of the mix and add 1 µl of USER enzyme and 1 µl of pNEB206a vector
5) incubate 15 min at 37°C and then 15 min at room temperature
6) use 5 µl of the reaction mix to transform chemically competent cells as described in 3.2.9.4
7) plate 70-200 µl of the transformed cells

Every cloned pool corresponded to a selective combination. A number of colonies 10-fold the number of the cDNA-AFLP bands forming the pool were picked and grown overnight in 384 well plates containing 60 µl LB + ampicillin 120 µg/ml per wells.

Colony PCR of the cloned cDNA-AFLP bands

The 96 pins tool (VP scientific) was used to perform a colony PCR allowing transferring per each pin about 1 µl of overnight culture in the PCR mix previously aliquoted in 96 well PCR plates. The mix used is a modified version of that described in a previous work by Diehl *et al.*(2002).

10 mM	Tris- HCl pH 8.3 (stock: 100 mM)
2.25 mM	$MgCl_2$ (stock: 50 mM)
50 mM	KCl (stock: 500 mM)
0.2 mM	dNTPs (stock: 10 mM)
1.5 M	Betaine (stock: 3 M)
0.25 µM	primer BstC0
0.25 µM	primer MseI0
0.1 mM	Cresol Red (stock: 50 mM)
0.2 µl	Taq Polymerase (Eppendorf 5 U/µl)
Final volume 40 µl	

PCR parameters:

94°C 10 min
95°C 30 sec
56°C 30 sec
72°C 40 sec } 35 cycles
72°C 4 min
16°C oo

The PCR products were controlled loading 5 µl on an agarose gel using two samples per plate-lane randomly chosen.
Then the PCR products were evaporated at 37°C until obtaining 1/2 of the initial volume and used for the spotting procedure.

Spotting procedure for the high density macro-array production

The production of the macro-array was carried out with the robot Biomek 2000 (Biomek) using a 4x4 geometry as spotting grid. A total of 6144 spots can be produced on a 14 cm x 10 cm filter membrane. The geometry is reported below:

P1	P2	P3	P4
P5	P6	P7	P8
P4	P3	P2	P1
P8	P7	P6	P5

Figure 8 Base geometry of the macroarray. P1= plate 1; P2 = plate 2; P3= plate 3; P4= plate 4; P5= plate5; P6=plate 6; P7=plate 7; P8=plate 8

The preparation of the membrane and the spotting steps were the same as described above for the low density macro-array procedure.

3.2.18 Hybridisation and differential screening

Preparation of the DIG labelled probe

The PCR DIG probe synthesis Kit (Roche) was used for the labeling of the PCR products. The kit allows the production of 25 labeled probes. The ideal ratio of the labeled dUTP-DIG:dTTP is 1:2 to perform a hybridisation with a single copy gene in a Southern hybridisation analysis.
For the array application, the ratio was changed and a 1:19 ratio was found to be an optimal condition to obtain good hybridisation signals. In this way, the number of reactions per kit can be increased to 437.

The PCR reaction mix was prepared as follows:

For the low density macro-array

2 µl of 1:19 dTTP:dUTP-DIG (2mM)
2 µl of dATP (2 mM)
2 µl of dCTP (2 mM)
2 µl of dGTP (2 mM)
2 µl 10X PCR buffer
0.2 µl Taq polymerase (Eppendorf 5 U/µl)
1 µl Bst-C0 primer
1 µl MseI0 primer
2.8 µl H_20
5 µl adapted cDNA-AFLP (diluted 1:500)
Final volume 20 µl

For the high density macro-array

2 µl of 1:19 dTTP:dUTP-DIG (2mM)
2 µl of dATP (2 mM)
2 µl of dCTP (2 mM)
2 µl of dGTP (2 mM)
2 µl 10X PCR buffer
0.2 µl Taq polymerase (Eppendorf 5 U/µl)
1 µl Specific BstYI-C+N primer
1 µl Specific MseI+N primer
2.8 µl H_20
5 µl pre-amplified cDNA-AFLP
Final volume 20 µl

The PCR reaction parameter used were the same as described in 3.2.14.

Hybridisation procedure

The protocol for the hybridisation procedure, was derived from the DIG manual (DIG application manual; Roche) and optimised with different conditions.
Composition of the used buffer:

Maleic acid buffer 10X:	1 M Maleic acid, 1.5 M NaCl; adjust with NaOH to pH 7.5 Autoclaved. RT stable
Blocking solution:	dilute 10X blocking solution 1:10 with 1X maleic acid buffer. Prepared fresh for every hybridisation
Blocking solution 10X:	10% w/v blocking reagent in 1X maleic acid buffer. Autoclaved and stored at 4°C
Detection buffer:	0.1 M Tris-HCl, 0.1 M NaCl, pH 9.5 (20 °C). Autoclaved. Stable at RT
20X SSC:	3 M NaCl, 0.3 M Na-citrate, pH 7.0.

1) pre-hybridisation of the macro-array filter for 1.5-2 h with 10 ml DIG easy hyb buffer
2) add the denatured (5 min, 95°C) labeled probe directly into the hybridisation tube
3) overnight hybridisation with the optimal temperature (47°C) calculated as suggested in the DIG manual
4) after the hybridisation, the probe can be maintained at -20° and reused for further hybridisation
5) place the filter in a plastic tray and add the low stringency (LS) buffer (2X SSC, 0.1 % SDS). Incubate 5-10 min in agitation at RT. Repeat this step with fresh LS buffer.
6) add the high stringency (HS) buffer (0.1X SSC, 0.1% SDS) pre-warmed at the 65°C
7) incubate the filter in agitation at the selected wash temperature (65°C) for 15 min. Repeat the wash step with fresh HS buffer.
8) wash once with wash buffer (Maleic acid buffer + 0.3 % v/v Tween 20) for 2 min
9) discard the wash buffer, add the blocking buffer and incubate 2 h in agitation
10) discard the blocking buffer and add 20 ml antibody solution freshly prepared (20 ml blocking buffer + 2 µl 10000X AP-Fab DIG)
11) incubate 30 min in agitation
12) wash twice for 15 min with wash buffer
13) add 20 ml detection buffer. Incubate 3 min in agitation
14) place the filter on a saran wrap with the hybridised side facing up
15) add 1 ml CDP-Star (1:200 diluted with detection buffer: 5 µl concentrated CDP-Star + 995 µl detection buffer) and cover the filter with the saran wrap. The liquid will be distributed over the filter surface. Avoid air bubbles formation
16) incubate the filter 5 min in the saran wrap envelope

17) place the filter over another piece of saran wrap (eliminate the exceeding liquid) and seal the contour forming a bag containing the filter. Avoid the formation of air bubbles
18) incubate 5 min in dark and then place the film (Kodak MR light) over the sealed filter
19) 5 -10 min exposure
20) develop the film

Stripping step

1) wash with distilled water 1-5 min
2) wash twice for 15 min with NaOH 0.2M + 0.1 % SDS at 37° C
3) wash with 2X SSC 5 min
4) filter can be maintained in 2X SSC for further hybridisations

3.2.19 Primer design

The design and development of all new primers used in this study was performed with the online web service of primer3 (http://biotools.umassmed.edu/bioapps/primer3_www.cgi) using the following standard parameters:

Table 4 Parameters used for the primer development

Parameters	Minimum	Maximum
Primer size	18	25
Primer GC content	45	65
Melting temperature	57	65
PCR product	120	220
3' clamp GC content	1	2
Max poly X	-	4

3.2.20 Real-time PCR analysis

The real-time analysis was performed using SYBR Green as detection method. The SYBR Green I® molecule binds specifically double-strand DNA. In this state, after excitation from a 492 nm light-source, it emits a fluorescence signal with a 513 nm wavelength. The emitted signal is proportional to the quantity of SYBR Green I® bound and therefore to the DNA quantity and length. In association with a PCR the build-up of the product and the reaction kinetic can be followed measuring the intensity of the emitted fluorescence proportional to the quantity of DNA.

3.2.20.1 Real-time PCR on RNA

The gene expression analysis was performed directly on the plant total RNA extracted with the method described in 3.2.2.
20 µg of total RNA were treated with RNAse free DNAse as follows:

20 µg total RNA
10 µl 10x RNAse free DNAse buffer
1 µl RNAse free DNAse
DEPC water to 100 µl

15 min at 37°C

Add 1 µl EDTA 0.5 M (final concentration 5 mM)

15 min at 70°C

The treated total RNA was directly used as template in the one-step qRT-PCR (Invitrogen kit #11736-59). The reaction mix was prepared as follows:

0.3 µl gene specific primer forward (10 µM)
0.3 µl gene specific primer reverse (10 µM)
10 µl 2X SYBR Green II PCR mix (Invitrogen)
0.4 µl RT/Platinum Taq mix (Invitrogen)
0.5 µl dNTPs (5 mM each) (Fermentas)
7.5 µl H_2O
1 µl DNAse treated total RNA
20 µl final volume

The parameters for the qRT-PCR were:

55°C 5 min
95°C 30 sec
60°C 25 sec
72°C 23 sec
Plate read 10 sec
} 40 cycles
72°C 4 min
Melting curve analysis 55 to 95°C, 1°C/min 10 sec read
16°C oo

3.2.20.2 Real-time PCR on cDNA

The real-time PCR analysis on the cDNA was performed using the cDNA synthesized with the method described in 3.2.13. The reaction mix for the real-time PCR was prepared as follows:

0.3 µl primer forward (10µM)
0.3 µl primer reverse (10µM)
2 µl 10x PCR buffer (Eppendorf)
0.5 µl dNTPs (Fermentas)
0.6 µl SYBR Green II (Biorad)
15.2 µl H_2O
0.1 µl Taq polymerase (5 U/µl; Eppendorf)
1 µl cDNA

The PCR reaction parameters were :

95°C 3 min
95°C 30 sec
60°C 25 sec
72°C 23 sec
Plate read 10 sec
} 40 cycles
72°C 4 min
Melting curve 55 to 95°C, 1°C/min 10 sec read
16°C oo

3.2.20.3 Preparation of a dilution series for the real-time qPCR primer efficiency test

For each gene tested in the real-time PCR a dilution series of the plasmids containing the corresponding cloned insert was prepared. A stock with a concentration of 10^9 molecules/µl was set for each plasmid based on their initial concentration and size. Subsequently a ten-fold dilution series was prepared until obtaining a single molecule/µl as last aliquot ($10^9 \rightarrow 10^0$).

3.2.20.4 Real-time PCR data analysis

The analysis of the data generated in the real-time PCR is an important part of the gene expression study. The method used in this thesis for the elaboration of the raw data is a modified version of the method described in Peirson *et al.* (2003). In each real-time PCR run the genes were investigated in double and 2 housekeeping genes were used as internal control for the healthy and infected samples. Three runs are performed for each gene investigated obtaining 3 values in double (6 raw values). The collection of the raw data was carried out with the Opticon Monitor 3.1 software (Bio-Rad Laboratories) while their elaboration and statistical analysis (one-factor analysis of variance - ANOVA with α=0.05; two tailed t-test with α=0.05) were performed with Excel®.

3.2.21 Agarose gel, polyacrylamide gel preparation and electrophoresis parameter

Agarose gel

The agarose gels were prepared using the quantity reported in the following table:

Table 5 Quantity of agarose depending on concentration and size of the gel

Concentration	Small size gel (55 ml TBE 1 X)	Medium size (110 ml TBE 1 X)	Large size (160 ml TBE 1 X)
	Grams of agarose (Roth)		
0.8 %	0.44	0.88	1.28
1 %	0.55	1.1	1.6
1.5 %	0.825	1.65	240
2 %	1.1	2.2	3.2

The agarose gels were run on TBE 1X buffer at 100 V. Gels were stained in 1% ethidium bromide and bands were visualised on an UV transilluminator.

Polyacrylamide gel

The polyacrylamide gels used in the cDNA-AFLP procedure were prepared using the sequencing gel electrophoresis apparatus (Model S2; Biometra, Göttingen, Germany). After the assembly of the glasses for the gel casting the polyacrylamide solution was prepared as follows:

80 ml SequaGel Monomer Solution 6 % (National diagnostics)
20 ml SequaGel Complete Buffer (National diagnostics)
0.8 ml APS 10 % (fresh prepared)

Mix with a magnetic stirrer and purge carefully between the two glasses of the sequencing apparatus avoiding the inclusion of air bubbles. Wait 1-1.5 hours for an optimal polymerisation reaction.

Gel electrophoresis parameters

The parameters for the electrophoretic run of the polyacrylamide gels were:

Voltage: 2500 V
Power: 80 W
Time: 3 hours or longer in case of the optimised step for the better redolution of the cDNA-AFLP bands

4 Results

At the beginning of this thesis nothing was known about the possible mechanism of the resistance to apple proliferation identified in *Malus sieboldii* and its hybrids. The trait of resistance to AP might be associated to constitutive differences in the morphology or in the type and quality of particular gene products of the phloematic tissues or, the resistance response of the plant might be induced after infection with '*Ca*. Phytoplasma mali'.

Therefore, the two main objectives of this study were the individuation of constitutive differences between AP-resistant and –susceptible *Malus* genotypes for the development of molecular markers associated with the trait of resistance to apple proliferation and the individuation in the resistant genotypes of genes involved in the plant response after infection with '*Ca*. P. mali'. The experimental work was structured in two main parts: in the first part analyses were performed for the development of genetic markers associated with expressed sequences while the second part was focused on the study of the variation in the plant-gene expression after the phytoplasma infection. Thanks to its versatility, the cDNA-AFLP TP technique was chosen as main procedure in both approaches for the analysis of the gene-expression patterns.

4.1 Analysis of constitutively expressed genes in healthy AP-resistant and -susceptible *Malus* genotypes

4.1.1 cDNA-AFLP on healthy field-grown plants

The aim of this first experimental part was the comparison of the expression patterns in the healthy state of the resistant and susceptible genotypes. The analysis of the constitutively expressed genes was performed through the application of the cDNA-AFLP TP technique. With this procedure, differences related to the nucleotide sequence like single nucleotide polymorphism (SNP), or differences in the expression of the genes can be individuated. In order to maximise the possibility to find such type of differences, three genotypes were chosen and compared because of their particular genetic interrelation: H0909, 4551 and M9.

As described in the introduction H0909 and M9 are directly related since M9 was used as father in the crossing that had produced H0909. On the other side, 4551 was derived from the same crossing trial (*Malus x domestica* cv Laxton's superb x *Malus sieboldii*) that produced the genotype 4556 used as mother in the cross yielding H0909. The genotypes 4551 and H0909 were shown to be resistant to AP while *Malus* x *domestica* cv. M9 is susceptible (Seemüller *et al.*, 2008; Bisognin *et al.*, 2008b). The attempt to compare the constitutive differences was based on these relationships because M9 and H0909 share part of their genome while 4551 and H0909 have part of their genome in common to the '*Ca*. P. mali'-resistant ancestor *Malus sieboldii*. In this way, we wanted to define constitutively expressed differences in the expression of genes related to M9 or to *Malus sieboldii*. This part of the thesis was directly associated with a more extended work inside of

the main project SMAP (2.3.1.2) in which a series of resistance screening-trials were programmed for the collection of data related to the behaviour of the genotypes against AP disease. The availability of these observations together with the information about the constitutively expressed differences would allow the linkage of the phenotypic data derived from the infection-trial with the gene expression data of the cDNA-AFLP analysis. This would give the possibility to individuate constitutively expressed genes or genetic markers associated to the resistance.

The plant material used for the analysis was collected from phloem tissue preparations (3.2.2.1) of branches or roots obtained from healthy plants maintained in the experimental orchard at the JKI Dossenheim (Germany) (Table 6).

Table 6 Genotypes and tissues used for the cDNA-AFLP TP analysis

Genotype	Tissue	
H0909	Root phloem	Phloem of branches
4551	Root phloem	Phloem of branches
M9	Root phloem	Phloem of branches

The samples were immediately frozen in liquid nitrogen and conserved at -80°C. The total RNA was extracted from the phloem tissues as described in the paragraph 3.2.2 followed by the cDNA-AFLP TP analysis (3.2.13; 3.2.14)

In the cDNA-AFLP TP procedure every selective amplification produces a specific pattern associated to the expressed genes of the genotype. This technique was originally applied for the comparison of differences in the expression level between developmental stages of a chosen organism (genotype) (Breyne and Zabeau, 2001). Therefore, the validity of this approach for the comparison of constitutively expressed genes between diploid and polyploid plants had to be proven. In order to check if reliable results could be obtained it was decided to test initially only the half of the possible primer combinations. Because the restriction enzyme BstYI used in the first digestion recognise the degenerated hexanucleotide sequence 5'-RGATCY-3' (R=A or G and Y=C or T) two types of cDNA fragments are generated: one type with a dCTP and one type with a dTTP as first nucleotide at the 5' end. Therefore, after the adapter ligation two different sets of BstYI primers are available: BstYI-C and BstYI-T. Each primer amplifies a different pool of cDNA generated from the BstYI digestion. In addition to this first differentiation, the usual selective amplification of the AFLP technique is carried out combining the four variants of the BstYI specific primers with the four variants of the MseI specific primers giving 16 primer combinations available for each set of BstYI primers.

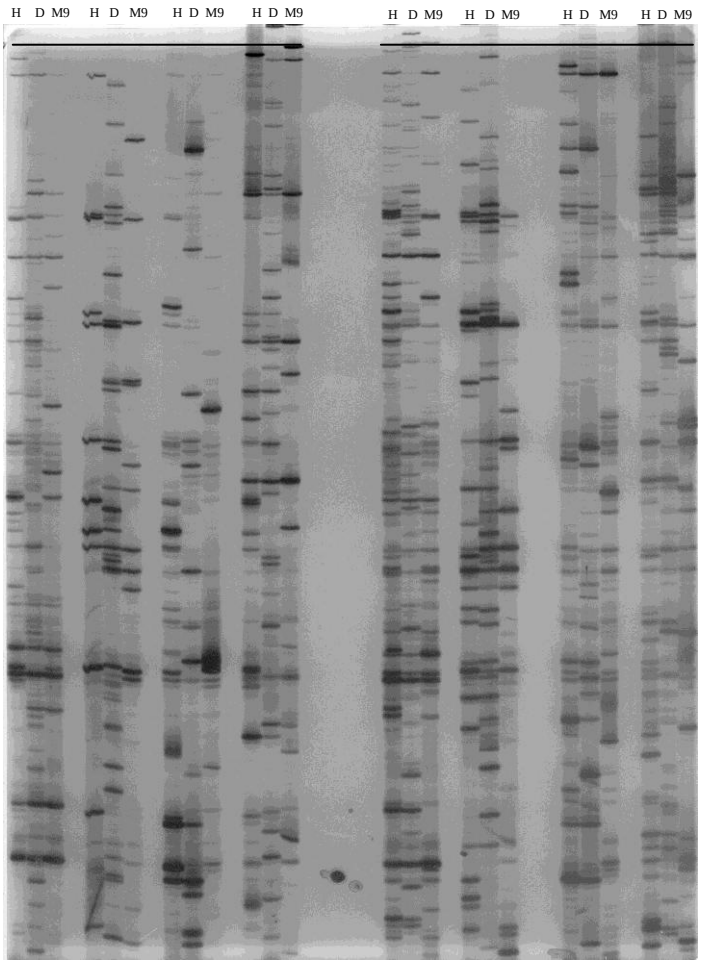


Figure 9 Example of a cDNA AFLP TP autoradiogram. Every lane corresponds to the pattern obtained from the amplification of the adapted cDNA of a genotype with the selective primer combination. Every group of three lanes corresponds to the three genotypes H=H0909 (left lane), D=4551 (middle lane) and M9 (right lane) respectively, amplified with the same primer combination. The four groups in the left part of the autoradiogram represent four different primer combinations tested on the samples obtained from the branch phloem tissue and the four groups in the right part are the same primer combinations tested on the samples obtained from the root phloem tissue.

Therefore, for the initial analysis the primer combination set associated to the BstYI-C primer was chosen (Table 7). The selective PCRs were carried out on the adapted fragments generated in the cDNA-AFLP procedure performed on branch and root phloem tissues (Figure 9). Then the expression patterns of the constitutively expressed genes were compared between the two resistant genotypes and the susceptible one. In the analysis of the pattern obtained from the same primer combination, those bands that showed a signal at the same height (bands with the same size) in both H0909 and 4551 but no signal in M9 were isolated from the blot (Figure 10).

Table 7 The four BstYI-C selective primers and the four MseI selective primers were combined to give 16 primer pair combinations.

BstYI-C selective primers	MseI selective primers
BstYI-C A	MseI A
BstYI-C C	MseI C
BstYI-C G	MseI G
BstYI-C T	MseI T

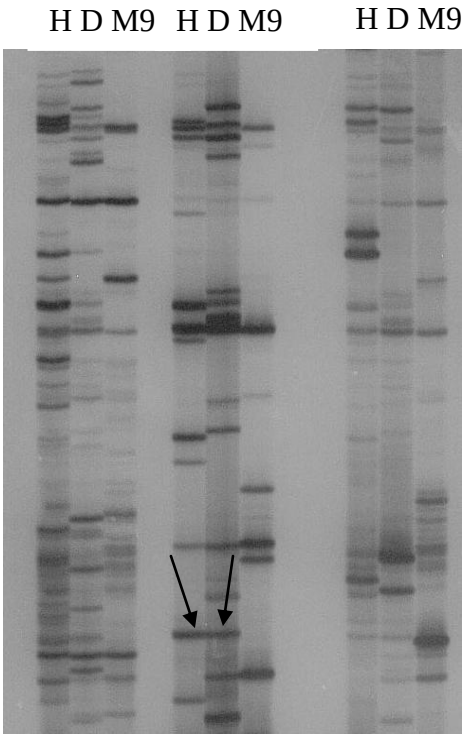


Figure 10 Zoomed region of Figure 9. The arrows show the type of bands individuated and isolated from the blot. H=H0909, D= 4551 and M9 indicate the genotypes. Every group of three patterns corresponds to a different primer combination used on the same tissue type.

From the comparison of the cDNA-AFLP patterns of the 2 resistant genotypes a total of 77 pairs of bands in the branch phloem tissue and 155 pairs of bands in the root phloem tissue were individuated (Table 8). In total, 464 bands were excised from the blot (3.2.14), 232 for H0909 and 232 for 4551. One code was associated to every band based on the primer combination used for the amplification and on the relative position that the band had in the pattern lane starting from the top of the gel.

Table 8 Number of selected pairs of bands for branch and root phloem tissue associated with the different primer combination applied in the first cDNA AFLP TP experiment. Every pair of bands corresponds to 2 bands effectively excised from the blot (one for each resistant genotype).

Selective base		N°of individuated bands present in H0909 and 4551	
of BstYI-C primer	**of MseI primer**	Branch phloem tissue	Root phloem tissue
A	*A*	0	0
	C	2	11
	G	4	8
	T	7	8
C	*A*	5	13
	C	2	6
	G	2	6
	T	6	9
G	*A*	6	13
	C	4	8
	G	5	11
	T	7	16
T	*A*	12	11
	C	4	16
	G	3	6
	T	8	13
TOTAL	**16 combinations**	**77 pair of bands**	**155 pair of bands**
TOTAL	**Branch + Root**	**232 pair of bands**	

A relative high number of bands was individuated and selected through a non-classical application of the cDNA-AFLP TP protocol. Therefore, to reduce the risk to analyse many uninteresting bands a preliminary hybridisation screening was performed in order to better identify those bands associated with cDNA fragments that could give valuable information for a further analysis.

4.1.2 Macroarray production and hybridisation

The absence of the band in M9 at the same height of the pair of bands found in the two resistant genotypes can be due to different possibilities:

- presence of a SNP in the M9 gene product at the BstYI or MseI restriction site that would produce a fragment that differ in length compared to those of H0909 and 4551
- presence of a SNP in the M9 gene-product sequence at level of the nucleotide (T or C) selected in the pre-amplification from the BstYI primer (BstYI-C or BstYI-T)

- presence of a SNP in the M9 cDNA fragment at level of the nucleotide selected from the BstYI or/and MseI selective primers
- the absence of the gene product in M9: I) not present in the genome
  II) not expressed
- very low concentration of the gene product in M9

Considering these points and in order to focus the attention as much as possible on those bands that can result informative, a macroarray with the PCR products of the re-amplified cDNA-AFLP-isolated bands was produced. The purpose was to individuate through hybridisation those cDNA fragments that were associated to the two resistant genotypes but not to M9. Therefore, each band originating from H0909 and 4551 was re-amplified with its respective selective primer combination from which the bands where individuated (Table 8). From the 464 bands, 5 x 96 well PCR plates were prepared and used for the spotting procedure (see 3.2.17 for low density macroarray preparation). The macroarray was produced using a fixed 96-pins pin tool (3.2.17.1). Every pin was immerged in the corresponding well of the PCR plate allowing to spot a standard volume of 1.5 µl of PCR products. The macroarray was constructed based on a 4x4 geometry. The five plates were spotted in double (3.2.17; Figure 7) on a 10 cm x 14 cm nylon membrane filter (3.2.17.1) giving a final number of 464x2 (928) spots. Two identical copies of the macroarray filter were produced.

The hybridisations of the filters (3.2.18) were performed using DIG-labelled PCR products as probe. The probes were obtained from the pre-amplification of the same adapted cDNA fragments used for the cDNA-AFLP TP analysis of H0909, 4551 and M9. The PCR were performed with the primer pair BstYI-C/MseI in order to obtain all the possible cDNA-AFLP fragments of the BstYI-C set (the filter were produced with the bands derived from the selective amplification of the BstYI-C set). For every genotype, 2 probes (branch and root) were produced giving a total of 6 probes (Table 9) used separately in the hybridisation experiments that were done in double (3.2.17.1).

Table 9 DIG-labelled BstYI-C/MseI probes produced for the hybridisation experiments

Probe name	Genotype	Tissue	
DIG-labelled probe 1	H0909	Branch phloem	
DIG-labelled probe 2	H0909		Root phloem
DIG-labelled probe 3	4551	Branch phloem	
DIG-labelled probe 4	4551		Root phloem
DIG-labelled probe 5	M9	Branch phloem	
DIG-labelled probe 6	M9		Root phloem

In the analysis only those spots were considered that showed a strong signal with the probes of H0909 and 4551 but not of M9, in both the replicated spots (Table 10). After the hybridisations, six spots were individuated and the corresponding cDNA-AFLP bands were used in the further analysis.

Table 10 Result of the macroarrays: cDNA-AFLP bands corresponding to those spots that showed a strong signal after the hybridisation experiments with probes of H0909 and 4551.

Primer combination		Genotype	Tissue	Name of band
BstYI-C	MseI			
G	T	H0909	Root	CGTHw3
		D4551	Branch	CGTDs2
		H0909	Root	CGTHw1
	C	D4551	Root	CGCDw3
C	T	H0909	Root	CCTHw8
T	G	H0909	Root	CTGHw5

4.1.2.1 Cloning and sequencing of the PCR products of the selected cDNA AFLP fragments

The selected bands (Table 10) were re-amplified with the specific selective primers (3.2.4) and loaded on a 1 % agarose gel (3.2.21). Each band except for CGTHw1 showed more than a single PCR product and up to 15 bands were detected (Figure 11) (Table 11) although several of them were very weak. This complicated the individuation of the bands that had hybridised with the probes. Therefore, the PCR products were separately eluted from the agarose gel (Qiagen Gel extraction kit) except the bands that were too weak. Respectively three eluted bands for CGTHw3, two for CGTDs2, one for CGTHw1, two for CCTHw8, three for CGCDw3 and one for CTG5Hw (Table 11) were purified. Again, the bands were re-amplified with the corresponding selective primers and loaded on a 1.5% agarose gel (3.2.21) to check their size. In some cases multiple PCR products were still present whereas in other no amplification was seen. However, a single product was obtained for the following bands: CGTHw3 U and B, CGTDs2 B, CGTHw1, CCTHw8 U, CGCDw3 U and M1, and CTGHw5 U (Table 11).

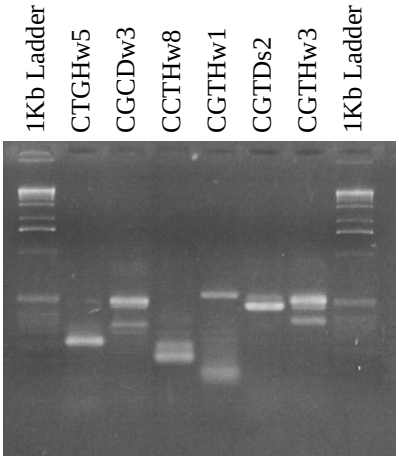


Figure 11 Agarose gel electrophoresis of re-amplified cDNA AFLP TP bands identified in the first hybridisation experiment showing that from each band more than one PCR product was obtained

Table 11 Number of PCR products obtained from the re-amplification of the cDNA-AFLP TP bands corresponding to those individuated after the hybridisation experiment

Band name	Number of bands from the first re-amplification	Name of the band eluted from the agarose gel	Band size	Products obtained after the re-amplification of the agarose gel-eluted bands
CGTHw3	3	U	510 bp	Single
		M	480 bp	Multiple products
		B	380 bp	Single
CGTDs2	2	U	510 bp	No product
		B	480 bp	Single
CGTHw1	1	S	520 bp	Single
CCTHw8	2	U	220 bp	Single
		B	200 bp	No product
CGCDw3	4	U	510 bp	Single
		M1	490 bp	Single
		M2	350 bp	Multiple products
		B	300 bp	Not analysed because too weak
CTGHw5	3	U	320 bp	Not analysed because too weak
		M	300 bp	Not analysed because too weak
		B	250 bp	Single

These PCR products were separately TA-cloned in pUC19 vector (3.2.9.1). Moreover, the CGTHw3 M and CGCDw3 M2 bands were cloned, although they showed more than one product, in order to obtain more information about the sequences that were represented by these bands. For every cloning 12 colonies were picked and after miniprep (3.2.11) the size of the different cloned inserts was checked through restriction analysis with EcoRI and HindIII (3.2.8). The bands CGTHw U, CGTHw B, CGTDs2 B, CGTHw1, CGCDw3 U and CGCDw3 M1 showed only inserts of the size corresponding to the respective PCR product. For each band three inserts were

sequenced (3.2.15). On the contrary, the clones of the bands CCTHw8 U and CTGHw5 U contained inserts with different length. Therefore, only those inserts that showed a size corresponding to the respective PCR band were sequenced. As expected, also for the bands CGTHw3 M and CGCDw3 M2 inserts with different size were obtained but in this case all were sequenced.

In total 30 inserts were selected from all the cloned bands and out of these 17 different unique sequences were obtained (Table 12). The sequences were compared with the expressed sequence tags (EST) database at NCBI and with the protein sequence database at Expasy.

Table 12 Results of the similarity analysis of the sequences obtained from 17 cloned cDNA AFLP TP fragments through blast with EST-NCBI and Expasy database. NR=no results

	Homology		Putative function			% of similarity		Identities		E value	
Band	NCBI EST	Expasy	NCBI EST	Expasy	Length bp	NCBI EST	Expasy	NCBI EST	Expasy	NCBI EST	Expasy
CGTHw3 U	*Malus x domestica*	*Citrus unshiu*	Plastid lipid associated protein (PAP)	Plastid lipid associated protein (PAP), chloroplast precursor	515	99%	69%	483/485	114/163 positives 124/163 (76%)	0	3,0E-54
CGTHw3 M	*Malus x domestica*	*Cucumis sativus*	High mobility group protein 1 (HMG1)	High mobility group protein (HMG1)	471	99%	82%	437/439	56/68 positives 64/68 (94%)	0	2,0E-27
CGTHw3 B	*Malus x domestica*	NR	AT2G31490 protein	NR	366	98%	NR	326/330	NR	2,0E-175	NR
CGTDs2 B	*Malus x domestica*	*Lycopersicon esculentum*	Ser/threo protein kinase ATPK1/ATPK6	Ser/threo protein kinase AtPK19 putative	284	98%	81%	259/263	71/87 positives 80/87 (91%)	2,0E-135	2,0E-41
CGTHw1	*Malus x domestica*	*Medicago truncatula*	Eukaryotic translation initiation factor 3 subunit 9 (eIF-3 beta)	Translation initiation factor eIF-3b	561	98%	83%	516/522	101/121 positives 108/121 (89%)	0	3,0E-55
CCTHw8 U 1	*Malus x domestica*	*Medicago truncatula*	Elongation factor EF-2	Elongation factor G, III and V translation factor	163	98%	96%	120/122	24/25 positives 24/25	5,0E-54	4,0E-11
CCTHw8 U 2	*Malus x domestica*	NR	Unknown	NR	170	99%	NR	132/133	NR	2,0E-65	NR
CCTHw8 U 3	*Malus x domestica*	NR	Probable ubiquitin-conjugating enzyme family	NR	158	95%	NR	111/116	NR	6,0E-41	NR
CCTHw8 U 4	*Malus x domestica*	NR	Unknown	NR	158	98%	NR	122/124	NR	3,0E-58	NR
CGCDw3 U	*Malus x domestica*	ARATH	Unknown	Hypothetical protein	502	98%	87%	456/464	114/130 positives 127/130 (97%)	0	4,0E-63
CGCDw3 M1	*Malus x domestica*	*Malus x domestica*	Zeta carotene/neurosporene dehydrogenase (Zcar desaturase)	Zeta carotene desaturase ZDS1	243	97%	90%	202/208	69/76 positives 71/76 (93%)	5,0E-98	8,0E-40
CGCDw3 M2 1	*Malus x domestica*	NR	Unknown	NR	335	99%	NR	107/108	NR	4,0E-50	NR
CGCDw3 M2 2	*Malus x domestica*	ARATH	Unknown	Unknown	318	99%	46%	220/221	41/89 positives 56/89 (62%)	1,0E-117	9,0E-16
CGCDw3 M2 3	*Malus x domestica*	*Cucumis sativus*	High mobility group protein 1	High mobility group protein	292	99%	77%	251/253	69/89 positives 81/89 (91%)	2,0E-134	4,0E-35
CGCDw3 M2 4	Malus x domestica	*Malus x domestica*	Zeta carotene/neurosporene dehydrogenase (Zcar desaturase)	Zeta carotene desaturase ZDS1	242	93%	82%	190/204	57/69 positives 60/69 (86%)	9,0E-72	2,0E-31
CTGHw5 B 1	Malus x domestica	NR	Expressed protein	NR	231	97%	NR	188/192	NR	3,0E-93	NR
CTGHw5 B 2	Malus x domestica	ARATH	14-3-3 protein	14-3-3- like protein GF14 kappa	217	99%	98%	179/180	58/59 positives 59/59	3,0E-93	2,0E-33

As shown in Table 12 a unique sequence was found for CGTHw U, CGTHw B, CGTDs2 B, CGTHw1 and CGCDw3 U. Moreover, from the cloned CGTHw3 M inserts only one type of sequence was detectable. However, in this case the small inserts (<120 bp) were not analysed and those inserts that were sequenced had the same size resulting to be the same product. From the clones chosen for CCTHw8 U and CTGHw5 B, different sequences were obtained although both showed a single band in the re-amplification of the gel-eluted PCR product. Finally, as expected, different sequences were identified from the CGCDw3 M2 clones.

For all sequences, it was possible to find a similarity from the BLAST analysis with the NCBI or the Expasy databases. For seven sequences, no putative function was found although they showed similarity with sequences present in the database. The other sequences showed similarity to genes or proteins that have a putative function involved in different types of metabolisms.

4.1.2.2 Development of specific primers for the evaluation of the first set of candidate genes

In order to check if some of these sequences were characteristic of the two resistant genotypes H0909 and 4551, specific primers were designed (3.2.19).

The test of the primers (Table 13) was done on the respective purified plasmid containing the insert corresponding to the sequence from which the primers were developed (3.2.4). For all primers the same annealing temperature of 60°C was used. A product with the expected size was obtained in every PCR. The following step was to test the primers on the cDNA of H0909, 4551 and M9 used also for the cDNA-AFLP TP analysis. Not all the primers showed amplification and for those that did not yield a PCR product, a gradient PCR was performed in order to optimise the annealing temperature.

Table 13 PCR primers developed from the sequences obtained from the cloning of cDNA AFLP TP bands as specified in Table 12

Primer name	Correspondence to band	Primer forward	Primer reverse
PAP	CGTHw3U	gga gcc agc gga gat tag tc	ggt cca gca aac tgg aca gag
TRF	CGTHw3 M1	ctg caa gcg cct tct tcg	tcc ctc agc tat gcg ctt g
Kin	CGTHw3 B	tac aag ggc atc gtc aga gac	cca aac aca atg cat acc tct c
STK	CGTDs2 B	gga ggg gat aac gtc aaa agc	ggt tag gtg cca cgt aag tg
PRT1	CGTHw1	ggc cac cat cat tgc tga c	cgt ccc taa gat ttt tcc ttg c
Ef2	CCTHw8 U 1	gct gcg gca ctt gtt acc	gcg atg agt cct gag taa tgc
Unk1	CCTHw8 U 2	gat ccc gtg gtg aag aac atg	cct gag taa tga aac agc atc c
Drht	CCTHw8 U 3	agt gat ccc gat ctt aca aca c	gcg atg agt cct gag taa tcg
Unk2	CCTHw8 U 4	gcg tag tga tcc cgg tac c	gcg tga gtc ctg agt aat gta c
EIPL	CGCDw3 U	atg gtg ctc ctc ctc acg atc c	ctc gat gcg ctg gtt caa ctg
Zcar1	CGCDw M1	ctg cgt agt gat ccg ttt aga c	gtg cgt tcc cca gaa tct tg
Zfing	CGCDw M2 1	ggt agg ttt agg gag aac aca g	gtg cgc aaa gtg aca ctt c
GlyRich	CGCDw M2 2	gtc gtg agg gtg gct atg g	gag cat gag tcc tga gta aca c
HMG1	CGCDw M2 3	gat ccg aac aag ccg aag ag	tgc atg cct gaa tgt tct tg
Zcar2	CGCDw M2 4	tgg atc cgt ttt aga cct gat c	gtg cgt tcc cca gaa tct tg
Unk3	CTGHw B 1	tgc gta gtg atc ctc gtc tc	gag cat gag tcc tga gta agc
14-3-3	CTGHw B 2	ctc aac tcc tcg gag aaa gc	gcg atg agt cct gag taa gc

A PCR product was obtained in all the cDNA tested showing no differences between the three genotypes. In addition, also the genotypes D2212, *Malus sieboldii* and Golden delicious were tested and showed the same results. The primers were further tested on the genomic DNA in order to check if some differences like introns were present. The annealing temperature used was the same as used for the PCR performed on the cDNA. Not all the primers led to a PCR product showing a general poor quality for some of them. Surprisingly, the primer pair Kin showed an interesting behaviour giving amplification at 60°C only with the AP-resistant genotypes H0909, 4551, D2212 and *Malus sieboldii* but not with the susceptible genotypes Golden delicious, M9 and Royal gala. Therefore, a further analysis on this primer pair was performed to better characterise the obtained results. This will be described below in the chapter (4.5.2.1).

4.1.3 Cloning of all selected and isolated cDNA-AFLP fragments from 4551 and H0909

The results obtained from the first hybridisation experiment did not show differences between the three genotypes except in one case with the primers Kin. Several information were derived from the experiment like the sequence similarity of the isolated constitutively expressed genes with genes or proteins present in the database and the individuation of sequences that are still not associated with a putative function. Nevertheless, the presence of more than one single PCR product in every spot was a factor that limited the power of the screening in the hybridisation experiment. In order to obtain more data from the bands isolated in the cDNA-AFLP procedure it was decided to produce a macroarray with a higher density and definition. Therefore, in the following experiment all 464 cDNA-AFLP bands selected of H0909 and 4551 from branch phloem and root phloem were cloned using the USER PCR cloning kit (New England Biolabs, NEB). As described in the protocol (3.2.9.2) new selective primers modified with the corresponding clamp designed from NEB were ordered (3.2.9.2) obtaining four USER-BstYI-C-X and four USER-MseI-X primers. The cDNA-AFLP bands were re-amplified and the products were checked on a 1% agarose gel (3.2.21) in order to control if a PCR product was obtained from all the bands. Subsequently 5 µl of every PCR product obtained from the same primer combination were pooled. From the 15 different primer combinations, 15 different pools were produced. The USER PCR cloning transformation protocol (3.2.9.2) was then performed using 10 µl of the pooled products.

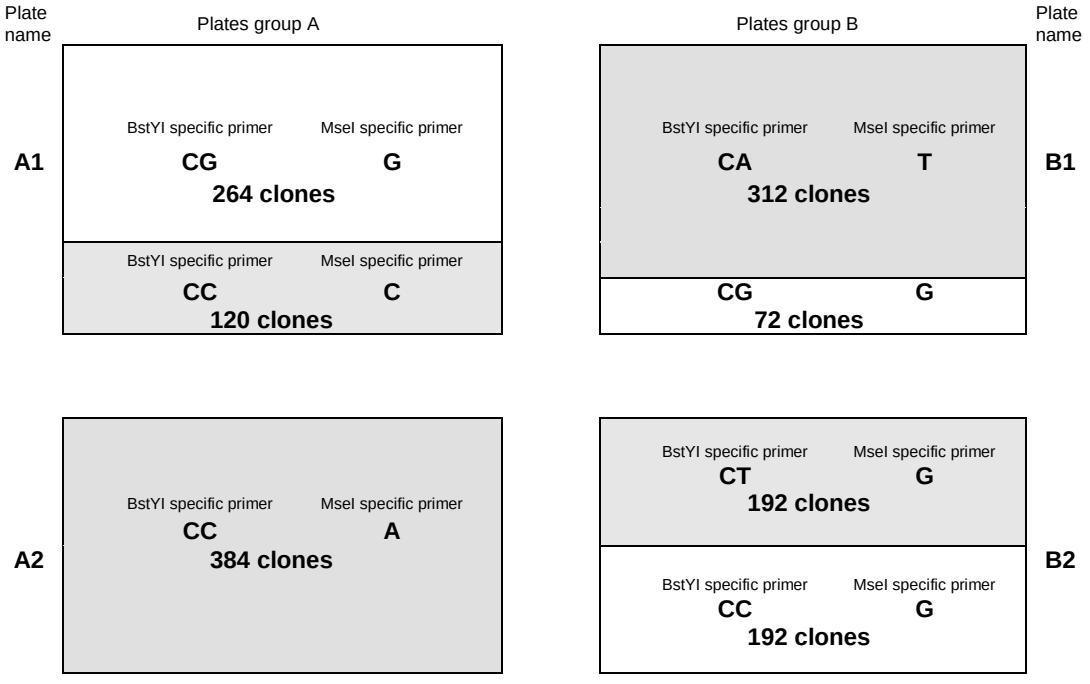
As shown before, a cDNA-AFLP band often contained more than one single PCR product. Therefore, a number of clones ten times the number of the re-amplified bands was picked in order to represent all the cDNA fragments in the cloning (Table 14). From the total 464 cDNA-AFLP bands, 464x10=4640 clones had to be theoretically picked.

At the end a higher number of colonies were effectively picked resulting in a total of 4968 clones that were inoculated in 384 wells plates (3.2.17) and grown overnight at 37°C. The 13 plates obtained containing the clones corresponding to the 15 different pooled combinations were

configured as shown in Figure 12. A copy of every plate was done and then stored at -80°C. The further step was the production of a high density macroarray.

Table 14 Number of cDNA-AFLP bands selected for each primer combination and number of clones effectively picked per combination

Primer combination		Tissue and genotype				Total bands	Total clones	
BstYI-C primer	MseI primer	Branch phloem		Root phloem			Theoretically needed	Effectively picked
		H0909	D4551	H0909	D4551			
A	A	0	0	0	0	0	0	0
	C	2	2	11	11	26	260	264
	G	4	4	8	8	24	240	264
	T	7	7	8	8	30	300	384
C	A	5	5	13	13	36	360	384
	C	2	2	6	6	16	160	168
	G	2	2	6	6	16	160	192
	T	6	6	9	9	30	300	312
G	A	6	6	13	13	38	380	384
	C	4	4	8	8	24	240	264
	G	5	5	11	11	32	320	336
	T	7	7	16	16	46	460	504
T	A	12	12	11	11	46	460	480
	C	4	4	16	16	40	400	408
	G	3	3	6	6	18	180	192
	T	8	8	13	13	42	420	432
TOTAL		**77**	**77**	**155**	**155**	**464**	**4640**	**4968**



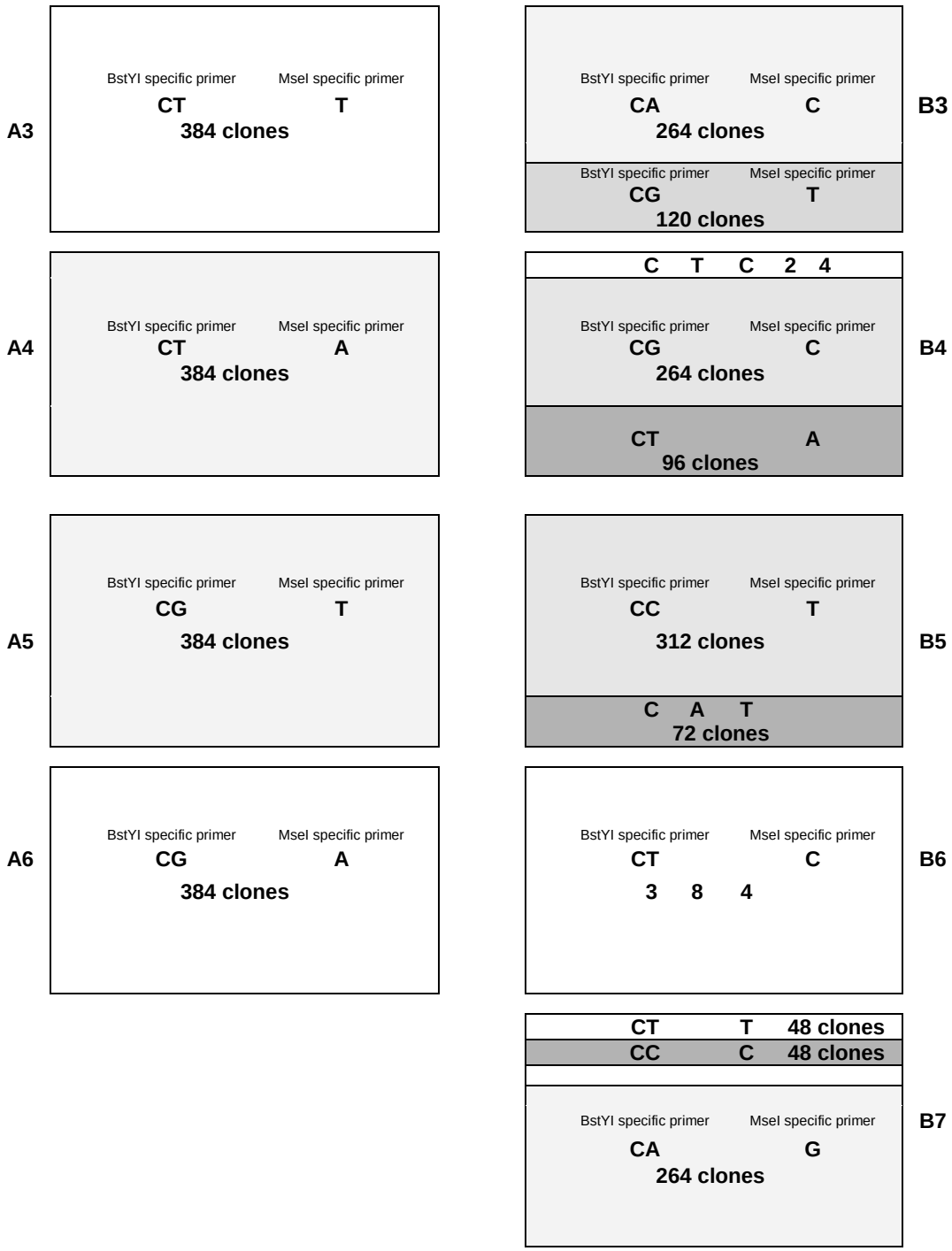


Figure 12 Plate configuration of the obtained clones. Every plate contains 384 wells. The reproduction of the plates shows the disposition of the clones with the selective primer combination of the bands from which they were obtained. The numbers indicate the clones picked for each band type. The different greyscales indicate regions of the plate with the same type of bands. The plates were divided in 2 groups for the membrane spotting: group A from plate 1 to plate 6 and group B from plate 1 to plate 7

4.1.3.1 Production and analysis of a high density macroarray

The inserts contained in the clones were re-amplified with the specific selective primers in 96-well plates performing a colony PCR as described in (3.2.10). Afterwards for every row of the PCR plate two samples were randomly chosen and checked on a 1% agarose gel (3.2.21) in order to control the general quality of the PCR. In all the selected samples, a good amplification of the inserts was obtained. Therefore, the PCR products were transferred from the 96 wells plates to the 384 wells plates using the robot Biomek FX (Beckman Coulter) maintaining the original configuration of the clone plates. The PCR products were then spotted in a high density macroarray as described in the methods (3.2.17.2) using the robot BIOMEK 2000 (Beckman Coulter) for the automatic spotting procedure. The geometry chosen for the filter production had a base grid of 4x4 spots with a 384 floating-pins pin tool (Figure 13). On this type of configuration up to 8 different 384 wells plates can be spotted in replicates. Therefore, two filters (A and B) were needed to spot all the products of the 13 plates.

Filter type A	1	2	3	4
1	A1	A2	A3	A4
2	A5	A6		
3	A4	A3	A2	A1
4			A6	A5

Filter type B	1	2	3	4
1	B1	B2	B3	B4
2	B5	B6	B7	
3	B4	B3	B2	B1
4		B7	B6	B5

Figure 13 Grid schema used for the spotting setup of the high density macroarray. The position 1-1 corresponded to the PCR products of the plate A1 for the filter type A and plate B1 for the filter type B. The positions 1-2, 1-3, 1-4, 2-1, 2-2 were associated with the other plates, respectively. The position 2-3 was occupied only in the filter type B by the plate B7. The replicates were spotted in an opposite order starting with the position 3-4 as corresponding spot for the plate A1 and B1 in the respective filter type.

The filter A was prepared with the products of the plates A1 - A6 for a total of 6x384 spots in replicates (4608 spots). The filter B was spotted with the plates B1 - B7 with a total of 7x384 spots in replicates (minus 24 spots that were empty in the plate B7 for a total of 5328 spots). For both filters, seven copies were done (3.2.17.2).

As in the first hybridisation screening, the probes used to hybridise the new filters were DIG labelled (3.2.18). However, in this experiment combination-specific probes were used in order to have a better resolution of the hybridisation patterns. First, in order to test the quality of the system, a DIG labelled probe was produced from the clone corresponding to the Kin sequence previously isolated. As described before, the Kin sequence was obtained from the band CGTHw3 B. Therefore, the filter A was used because it contained the spots produced from the clones obtained

from the CGT-amplified bands (Figure 12; plate A5). Several spots were individuated on the filter A (Figure 14) and, interestingly, some of them corresponded to products derived from other primer combinations (Figure 14; grid 1, 2 and 4) than the one which produced the Kin band (Figure 14; grid 3, 5 and 6). Six spots were selected (Figure 14; highlighted grids) which showed a comparable intensity in the hybridisation signal in both replicates. In order to gain information about the sensitivity of the procedure, spot couples with weak (Figure 14; grid 3 and 5) and strong (Figure 14; grid 1, 2, 4 and 6) signal (in both spots) were also chosen. The clones associated to these spots were individuated from their corresponding position in the 384 well clone plate. A colony PCR with the specific primers for the Kin sequence was performed on each clone and for all six clones a band with the same size of the Kin clone was obtained (Figure 15). No PCR products were obtained with clones corresponding to hybridisation signals characterised by differences between the replicates. This was an important result showing that spots containing the same fragment but deriving from cDNA-AFLP bands produced from different primer combinations were also individuated. Thus, this approach would allow the screening of fragments that were present in all the genotypes tested but polymorphic at one or more of the selective steps of the cDNA-AFLP procedure (restriction site, first amplification or/and selective amplification).

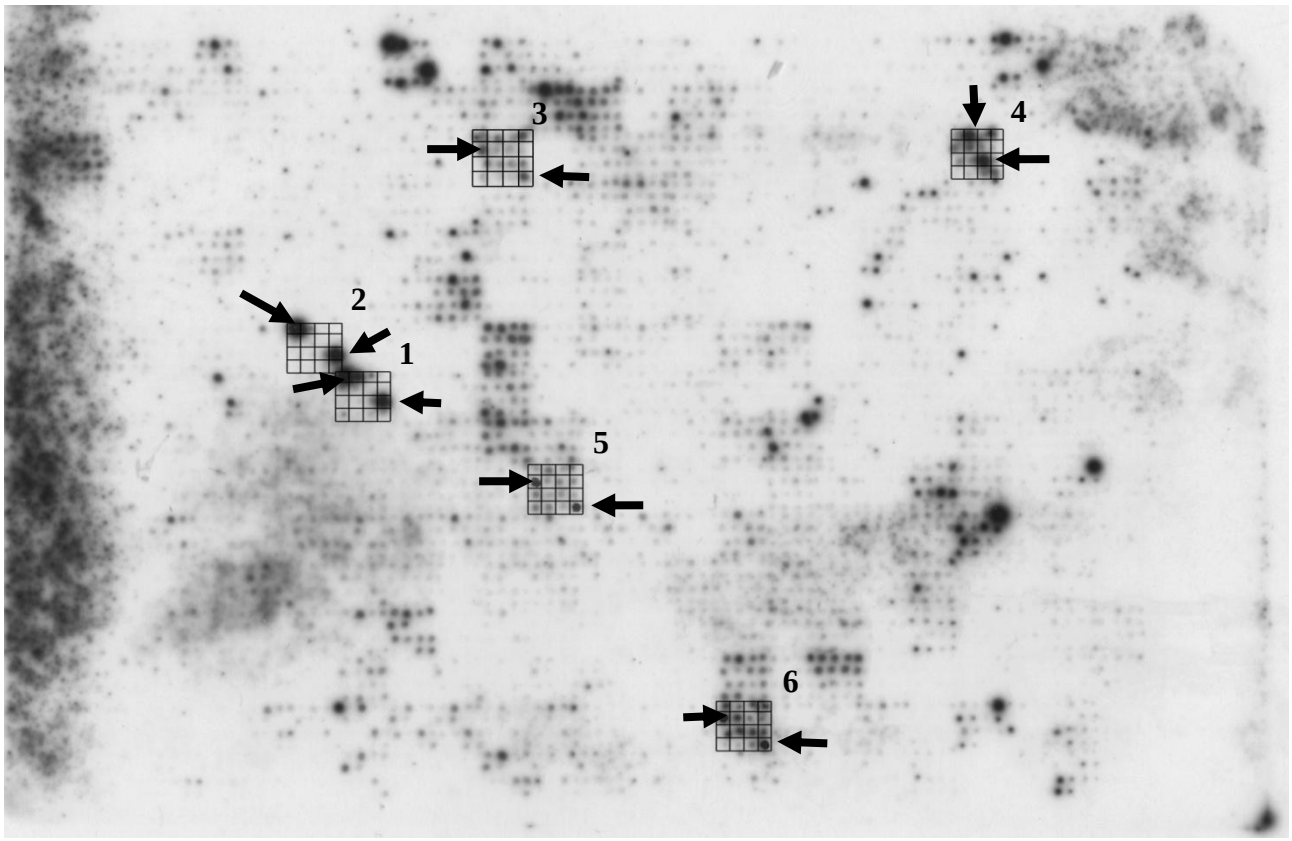


Figure 14 Evaluation of the high density macroarray with a DIG labelled probe of the Kin gene. The grids are representing the 4x4 spotting geometry (4.1.3.1). Few spots that showed the same signal intensity in both replica were chosen for PCR control and are indicated by the arrows. Spot couples with different degrees of signal intensity were chosen to proof the specificity of the hybridisation.

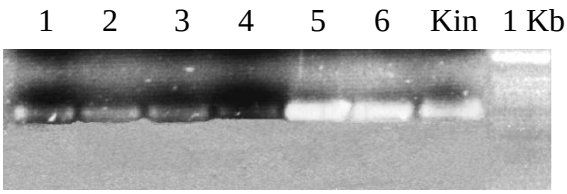


Figure 15 Agarose gel electrophoresis of Kin specific PCR products obtained from clones associated to the spots showing in both replicates comparable hybridization signals with the Kin probe in the macroarray.

Therefore, for each genotype and each tissue one DIG labelled probe was produced using the entire BstC-X/MseI-X set of specific primer combinations. A total of 90 probes was obtained (15 primer combinations x 2 tissues x 3 genotypes = 90 probes) (Diagram 1).

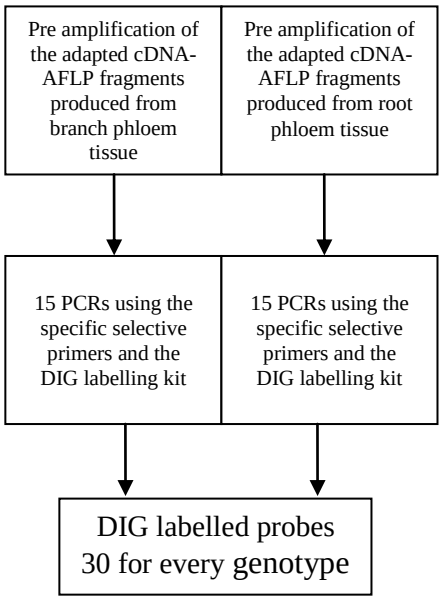


Diagram 1 Steps involved in the production of the DIG labelled probes. This procedure was performed on the cDNA-AFLP adapted fragments of H0909, 4551 and M9.

The hybridisations were performed as described in 3.2.18. Like in the previous screening, only those clones that showed a hybridisation signal with the probes produced from H0909 and 4551 were selected (Figure 16 red and black circles). From the comparison of the hybridisation patterns obtained from the 90 probes, a total of 142 spots was found which had the same signal intensity in the 2 resistant genotypes and no signal in M9. Among these, 110 spots showed a signal with probes produced either from branch or from root. The remaining 32 spots represented hybridisation with both types of probes (Table 15), thus the unique spots in this case were 16. Moreover, further 132 spots having different signal intensities between the H0909 and 4551 were individuated. These are listed in Table 16 but were not further analysed.

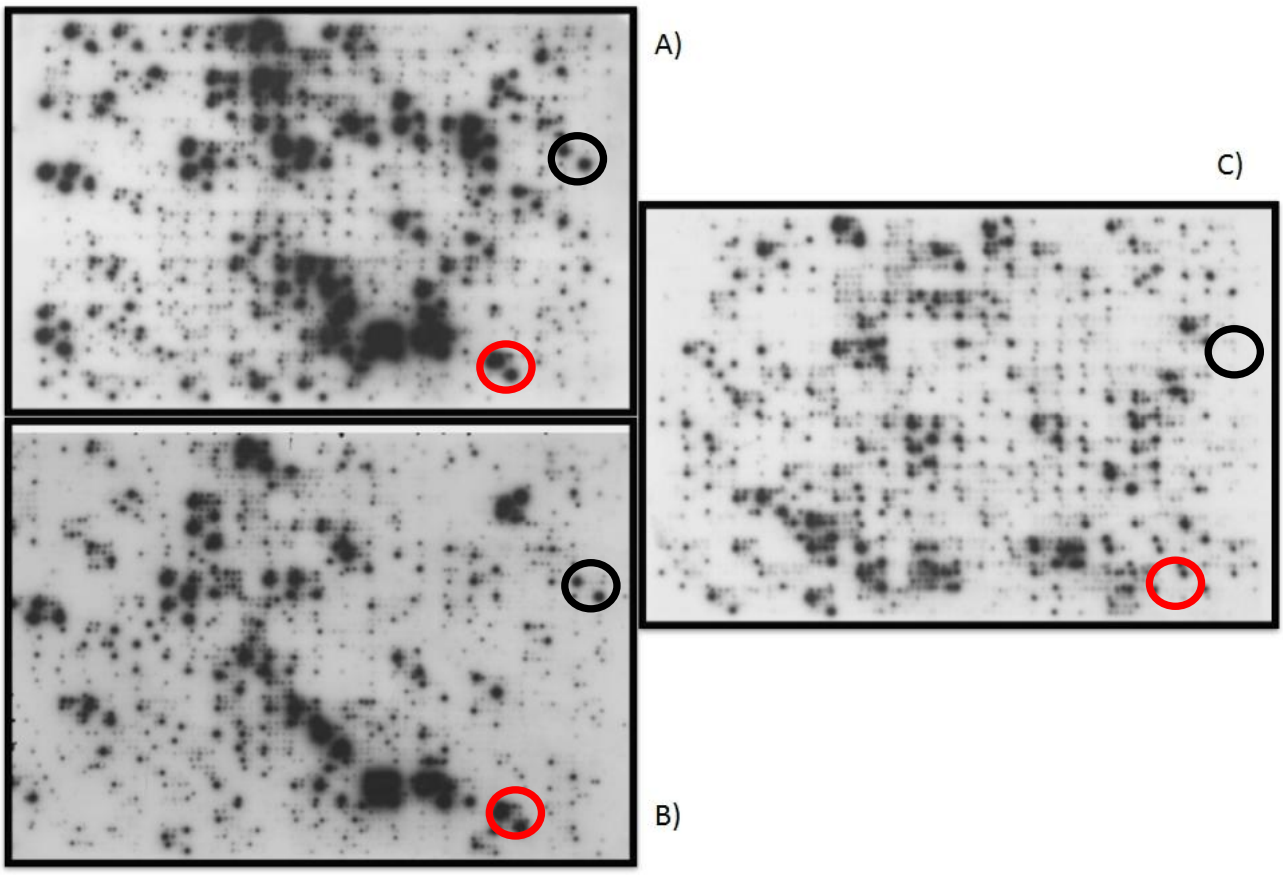


Figure 16 Examples of hybridisations with probes produced with a specific cDNA-AFLP primer combination. In A, B and C a probe produced from cDNA-AFLP fragments obtained from root material of H0909, 4551 and M9, respectively, was used. Differences in the hybridization pattern are highlighted with circles.

Table 15 Number of spots for each probe that showed a signal with the same intensity for both H0909 and 4551. In the case of spots that showed a signal for both the tissues in the same filter position the spot was counted as single (unique). However, in this case the effective number of clones corresponds to the double of the number of spots individuated and this is reported in brackets.

Primer combination		Number of spots that showed a signal with the same intensity for both H0909 and 4551		
BstYI-C primer	MseI	Root phloem	Branch phloem	Branch and root phloem
A	C	2	2	0
	G	1	0	0
	T	1	0	0
C	A	0	0	0
	C	0	1	0
	G	27	6	2(4)
	T	0	0	0
G	A	0	0	0
	C	8	10	13(26)
	G	0	2	0
	T	7	1	0

T	A	8	4	1(2)
	C	0	6	0
	G	4	18	0
	T	2	0	0
TOTAL		60	50	16(32)
Total		110		32

Table 16 Number of spots that showed for each probe a signal in both the resistant genotypes but with different intensities. As described in Table 15 the number in brackets represents the effective clones associated to the reported spots

Primer combination		Number of spots that showed a signal with different intensity in H0909 and 4551		
BstYI-C primer	MseI	Root phloem	Branch phloem	Branch and root phloem
A	C	5	6	0
	G	5	0	0
	T	4	0	0
C	A	1	0	0
	C	0	1	0
	G	1	2	0
	T	1	0	0
G	A	0	0	0
	C	9	6	2(4)
	G	0	2	0
	T	8	5	0
T	A	31	9	2(4)
	C	3	6	0
	G	3	13	0
	T	1	2	0
Total		72	52	4(8)
Total		124		8

4.1.3.2 Sequence analysis of cDNA AFLP bands individuated by high density macroarray

The clones corresponding to the 126 unique spots (110+16) were individuated on the respective 384 well plate. The inserts were re-amplified and their size checked on a 1% agarose gel. Further, the inserts were sequenced (3.2.15) and the products were loaded on a genetic sequencer ABI prism 3100 for the capillary electrophoresis analysis.

A total of 118 sequencing reactions produced a chromatogram with a sufficient quality for further analysis. After the cleaning process in which the primer and the vector sequences were eliminated a clustering analysis was performed on the insert sequence in order to group them by homology. In total, 23 clusters were obtained and the corresponding contigs were blasted on the EST database present at NCBI. In addition, the translated sequences were blasted on the Expasy database to control the similarities at the amino acid sequence level (Table 17).

Table 17 Sequences obtained from the hybridisation screening. NR=no result

Cluster /Contig	Length	Organism NCBI	Organism Expasy	Putative function NCBI	Putative function Expasy	Identities NCBI	Identities Expasy	E-value NCBI	E-value Expasy
1	364	*M. sieboldii*	NR	Elongation factor 1-alpha	NR	321/323 (99%)	NR	3e-174	NR
3	295	Royal Gala	*A. thaliana*	putative high affinity nitrate transporter	F13K23.20 protein	288/296 (97%)	Identities:60%(28/46) ; positives:80% (37/46)	2e-143	6e-14
4	235	*Malus x domestica*	*A. thaliana*	similarity to *A. thaliana* ubiquitin-like protein	Ubiquitin-like protein 5	190/194 (97%)	Identities: 27/30 (90%); positives: 29/30 (96%)	4e-92	1e-16
5	297	Royal Gala	*Medicago truncatula*	Unknown	Ribulose bisphosphate carboxylase amall chain precursor	255/256 (99%)	Identities: 62/83 (74%); positives: 67/83 (80%)	1e-128	4e-33
7	397	*M. sieboldii*	*Pyrus pyrifolia* var. Culta	GAPDH cytosolic	GAPDH like protein	395/399 (98%)	Identities: 115/133 (86%); positives: 116/133 (98%)	0.0	2e-59
10	266	*Malus x domestica*	*Malus x domestica*	Soluble inorganic pyrophosphatase	Soluble inorganic pyrophosphatase	221/225 (98%)	Identities: 37/42 (88%); positives: 37/42 (88%)	1e-110	8e-18
11	332	*Malus x domestica*	*Cucumis sativus*	HMG-1	HMG-1	470/478 (98%)	Identities: 64/78 (82%); positives:74/78 (94%)	0.0	2e-32
17	165	Royal Gala	*A. thaliana*	14-3-3 Protein GF14 kappa	14-3-3 Protein GF14 kappa	172/175 (98%)	Identities: 51/54 (94%); positives:54/54 (100%)	3e-83	3e-28
20	387	*Malus x domestica*	*Cucumis sativus*	HMG-1	HMG-1	261/268 (97%)	Identities: 72/79 (78%); positives:85/92 (92%)	2e-131	2e-37
23	198	*Malus x domestica*	NR	possible apospory-associated; ankyrin repeat-containing protein2	NR	197/201 (98%)	NR	1e-91	NR
6	215	*Malus x domestica*	*A. thaliana*	similar to F8K7.23	Hypothetical protein F8K7.23	179/179 (100%)	Identities: 50/58 (86%); positives: 57/58 (98%)	3e-95	7e-29
9	380	*Malus x domestica*	*A. thaliana*	T4P3.9 protein	T4P3.9 protein	251/265 (94%)	Identities: 44/79 (55%); positives: 59/79 (74%)	8e-113	2e-17
19	199	*Malus x domestica*	NR	T2N18.13 protein	NR	157/158 (99%)	NR	3e-80	NR
22	356	Royal gala	NR	similar to T24D18.26 Protein	NR	347/352 (98%)	NR	0.0	NR
8	249	*Malus x domestica*	NR	Unknown	NR	114/118 (96%)	NR	1e-46	NR
12	305	M9	NR	Unknown	NR	300/305 (98%)	NR	7e-156	NR
13	307	*Malus x domestica*	*Oryza sativa*	Unknown	Unknown	271/275 (98%)	Identities: 44/66 (66%); positives:53/66 (80%)	8e-143	5e-21
16	315	*Malus x domestica*	NR	Unknown	NR	138/140 (98%)	NR	2e-64	NR
21	122	*Malus x domestica*	NR	Unknown	NR	107/113 (94%)	NR	2e-39	NR
2	317	NR	NR	NR	NR	NR	NR	NR	NR
14	101	NR	NR	NR	NR	NR	NR	NR	NR
15	142	NR	NR	NR	NR	NR	NR	NR	NR
18	141	NR	NR	NR	NR	NR	NR	NR	NR

After the blast analysis, four contigs did not show similarity with any sequence present in the database. For the other 19 contigs a similarity with a nucleotide or a protein sequence was found. For five of them it was not possible to associate a putative function; four showed similarity with putative proteins while the remaining 10 were similar to nucleotide or protein sequences with known function. Interestingly, cluster 1, 2, 5, 6, 8, 11, 12, 16, 17 and 21 were represented by sequences corresponding to spots that hybridised, whitin their respective specific primer combination, with both probes produced from cDNA-AFLP fragments derived from sprout- and root-tissue. Moreover, cluster 7 and 10 were associated to a root derived probe whereas cluster

15, 18 and 19 to a sprout derived probe. Finally, cluster 3, 4, 9, 20 and 22 were single sequences (therefore not true clusters) obtained from root-derived probes while cluster 13, 14 and 23 were single sequences obtained from sprout-derived probes.

4.1.3.3 Development of specific primers for the evaluation of the second set of candidate genes

Specific primers were developed on 11 of the sequences of the 23 contigs (Table 18) (3.2.19). The choice was based on the putative function deduced by the sequence similarity information and on the possibility to obtain primers with the parameters used in the algorithm for the primer development (3.2.19). Therefore, sequences that had homology to known genes like glyceraldehyde-3-phosphate dehydrogenase (cluster 7) or ribulose-1.5-bisphosphate carboxylase oxygenase (cluster 5) were not analysed while for others (cluster 6, 9, 14, 18, 20) no compatible primers were found. The developed primers were then tested on the plasmids containing their respective insert. All PCR reactions showed a product using 58°C as annealing temperature. The further step was the analysis on the cDNA of H0909, 4551 and M9. In addition, the cDNA of *Malus sieboldii* and D2212 was tested. All the primers showed a product in each genotype thus no differences were seen. The test was repeated also on the genomic DNA of the same genotypes. Here some primer pairs did not work but in those that showed amplification no differences were individuated except for the C8 primers. Surprisingly, a different behaviour between H0909 and *Malus sieboldii* compared to M9, 4551 and D2212 was detectable. Therefore, further investigations were performed on this primer pair and the results are described in the 4.5.2.2.
Moreover, all the sequences were further analysed for the development of SNP markers which are reported in 4.5.1.

Table 18 Primers developed from the 23 contigs derived from the sequence analysis of cDNA AFLP fragments individuated in the high density macroarray

Primer name	Forward	Reverse
Cont1	CGCAGGCGAAATCTTCTATC	GAAGAACCCAGAGCACCTTG
Cont2	CGATTCTCCGCTCGTTTG	GTTCCCTCCGCCATATCC
Cont3	GTCACTGGGAGTGGTATCAGG	TGAGGGAGACTGGGAGAGTG
Cont8	GCACCTCCTCTTCTGACAGC	GTTTCAGCTCCCTCACCTTG
Cont12	CCAAGGCAGACAAGAAGAGG	TCCCGGGTAAACAAACTGTC
Cont13	GGGCTTTCAGCTTCCTATCC	ACTTCCCGACGTGTTCTTTG
Cont15	GGTCACGGTACTGTTTCTACAAG	TGTGATCAAAACATCCTCGTC
Cont16	CGTGCCAACACACAATTACAG	GTTTCAGCTCCCTCACCTTG
Cont19	GCTGCCACATCCTGGTTTC	ATCCAATGCCCAGGAAGG
Cont21	AGTGTCTGGTCGGACCTC	TAATACAAAGGGAACGACACAC
Cont22	CCTCTGCTTCCATTAGACAAGG	TCCACCAACTTCCCACTCTC
Cont23	GTCGCTCTCCAGAACATGG	CCGATCCAATCGTCCAAG

4.2 Analysis of induced genes in AP-resistant and susceptible plants after infection with '*Ca*. P. mali'

The study of the gene expression needs the establishment of a system in which the parameters that can influence the analysis are under control. In the case of the pathogen-host interaction this is even more complex because of the necessity to produce standardised infected material. In the specific study of the phytoplasma-plant interaction in a woody, perennial plant species this can be a problem. Therefore, the availability of a system in which the infection of the plants and their maintenance can be standardised is a crucial point in the gene expression study of this type of interaction. As shown in previous experiments (Jarausch *et al*., 1996), '*Ca.* P. mali'-infected apple plants can be maintained *in vitro* and, thus, in a controlled system in which parameters like temperature, light and growth can be standardised for the experimental conditions. Furthermore, healthy shoot cultures can be inoculated with '*Ca.* P. mali' under standardised conditions by *in vitro* grafting (Jarausch *et al.*, 1999a). Therefore, in order to prepare the material that could be used in the gene expression analysis of the '*Ca.* P.mali'-plant interaction an *in vitro* system was developed based on the plants that were used in this study allowing the control of a broad range of parameters that can influence the gene expression.

4.2.1 Production of homogenous *in vitro* shoot cultures of AP-resistant and susceptible *Malus* genotypes

In the year 2002, *in vitro* cultures of the genotypes resistant and susceptible to AP (Table 2) were produced by Dr. Jarausch at the AlPlanta Institute. Young sprouts of plants present in the field at the JKI Dossenheim (Germany) were used to initiate *in vitro* cultures as described in Ciccotti *et al.* (2008b). For this work, these cultures were then further propagated and maintained in the growth chamber constituting a standardised set of *in vitro* plants ready for the infection trials (3.1.1).

'*Ca.* P. mali'-infected apple *in vitro* plants were also established from naturally infected trees (referenced in Bisognin *et al.,* 2008a). In the present work, the '*Ca.* P. mali'-infected culture produced by Dr. Jarausch at the AlPlanta Institute was used as inoculum source for the infection trials (3.1.1). This culture was infected with '*Ca.* P. mali' subtype AP (Jarausch *et al.*, 2000) whereas in experiments done in parallel at IASMA (Italy) in the frame of the SMAP project the '*Ca.* P. mali' subtype AT2 was used (3.1.1).

4.2.2 Production of '*Ca*. P. mali'-infected plants by *in vitro* graft-inoculation

The use of *in vitro* culture for the production of the starting material has the great advantage of a high standardisation of the growth conditions through the controlled temperature, light and procedures for the maintenance. In the year 2003, the different genotypes were inoculated with

'*Ca.* P. mali' in order to generate a standardised population of infected *in vitro* plants. The infections were done via *in vitro* grafting as described in Jarausch *et al.* (1999a) (3.2.1.1). One month after the grafting the strength of the junction was tested separating the rootstock from the tip. As the phytoplasma transmission depends on the formation of a phloem-phloem connection between graft tip and rootstock, only grafts with strong junction were considered successful. The tips were then conserved at -20°C. The rootstocks were transferred to a new culture tube and further cultivated. After another two months the plants were propagated in order to obtain more material and tested for the presence of the phytoplasma through PCR (3.2.4). In the case of a negative PCR result, the tips of those graftings that showed a strong junction were tested in order to check if they were infected. This control was not done for the tips of a weak union because in this case the transmission rate is usually very low (Jarausch *et al.*, 1999a). In order to acquire information on the transmissibility of the pathogen to the different genotypes, a series of grafts was performed. Table 19 summarises the results obtained from the inoculation trials done through *in vitro* grafting for the preparation of standardised infected material. Inside of the SMAP project these grafts were only a minor set of the total number of replica that were produced as main part of another study. These data have been published together in Bisognin *et al.,* (2008a). As reported in Table 19, no infected plants were obtained from the grafts of *Malus sieboldii* and H0909. A high rate of mortality, low rate of strong grafts and a low rate of phytoplasmas transmission were observed in these resistant genotypes. Although the mortality rate and the rate of strong grafts were comparable to those observed in *Malus sieboldii* and H0909, for the D2212 genotype infected plants were obtained showing in this case a higher transmission rate.

Table 19 ***In vitro*** **graftings carried out on healthy cultures using the '*Ca.* P. mali'-infected apple-cultures as inoculum**

Genotype - phytoplasma subtype	No. of grafts	No. of strong grafts (SG)	% of SG	No. of PCR + plants (SG)	Transmission rates (SG)	No. of PCR+ plants (SG)	Transmission rates (SG)	No. of PCR+ plants (SG)	Transmission rates (SG)
		1 month post inoculation		3 months post inoculation		6 months post inoculation		12 months post inoculation	
M. Sieboldii - AP	13	3	23%	0	0%	0	0%	-	-
M. Sieboldii - AP	30	8	27%	0	0%	1	13%	-	-
M. Sieboldii - AP	28	15	54%	3	20%	2	13%	0	0%
M. Sieboldii - AP	30	12	40%	0	0%	0	0%	-	-
	101	**38**	**38%**	**3**	**8%**	**3**	**8%**	**0**	**0%**
D2212 - AP	14	8	57%	7	88%	4	50%	3	38%
D2212 - AP	34	6	18%	4	67%	0	0%	-	-
D2212 - AP	9	4	44%	1	25%	1	25%	1	25%
D2212 - AP	43	20	47%	6	30%	6	30%	5	25%

	100	**38**	**38%**	**18**	**47%**	**11**	**29%**	**9**	**24%**
H0909 - AP	15	3	20%	1	33%	0	0%	-	-
H0909 - AP	9	3	33%	2	67%	0	0%	-	-
H0909 - AP	12	4	33%	0	0%	0	0%	-	-
H0909 - AP	29	18	62%	1	6%	1	6%	0	0%
H0909 - AP	30	13	43%	0	0%	0	0%	-	-
	95	**41**	**43%**	**4**	**10%**	**1**	**2%**	**0**	**0%**
GD - AP	23	9	39%	8	89%	8	89%	8	89%
GD - AP	7	6	86%	3	50%	3	50%	3	50%
	30	**15**	**50%**	**11**	**73%**	**11**	**73%**	**11**	**73%**
M 9 - AP	12	4	33%	2	50%	0	0%	-	-
M 9 - AP	11	6	55%	1	17%	1	17%	1	17%
M 9 - AP	24	17	71%	12	71%	8	47%	8	47%
	47	**27**	**57%**	**15**	**56%**	**9**	**33%**	**9**	**33%**

However, in the series of *in vitro* grafting trials done at IASMA (Italy) for the related research activity inside the SMAP project, H0909 and *Malus sieboldii* infected plants were obtained. In this case instead of the AP strain the '*Ca.* P. mali' AT2 strain was used for the infection trials. Therefore, '*Ca.* P. mali' AT2 strain-infected *Malus sieboldii* and H0909 were obtained from the IASMA laboratory in which the same experimental conditions as for our cultures were used to produce the infected material. A set of these infected *in vitro* plants were transferred in our lab and maintained for a year in the same conditions as the other genotypes used in the experiment. In addition, AT2-infected Golden Delicious plants were prepared in order to have a susceptible genotype as a control. All the grafts were started at the same date and after one year in which they were maintained as described in 3.2.1 the infected plants were collected for the cDNA-AFLP analysis.

The healthy plants derived from the same culture lines which were used in the infection trials. These cultures were subcultured at the same day as the graft-inoculated plants and maintained in the same conditions. The transfer in tubes with fresh media was performed following the subculture steps used for the infected plants.

The *in vitro* cultures of the 4551 genotype were not growing homogenously at the beginning of the infection trials and therefore were not included in the analysis.

4.2.3 RNA extraction from healthy and '*Ca*. P. mali'-infected *in vitro* plants

The infected and healthy *in vitro* plants were collected one year after the graft-inoculation. For every genotype (Table 20) 8-10 shoots comparable in size and physiological state were collected, pooled in a falcon tube and immediately frozen in liquid nitrogen. The total RNA was then extracted as described in 3.2.2.

Table 20 Genotypes used for the cDNA AFLP experiment and association of the '*Ca.* P. mali' strains with the genotypes infected through *in vitro* grafting

Genotype	'*Ca*. P. mali' strain
Malus sieboldii	AT2
H0909	AT2
D2212	AP
Golden delicious	AP, AT2
M9	AP

4.2.4 cDNA-AFLP analysis of healthy and '*Ca*. P. mali'-infected *in vitro* plants

For every genotype 100 µg of total RNA extracted from the healthy and the infected plants (Table 20) were used for the poly(A)$^+$ isolation (3.2.12). About 1 µg was then used for the cDNA synthesis as described in 3.2.13. The cDNA was purified and 500 ng were used for every genotype in the cDNA-AFLP procedure as described in 3.2.14. All the 32 possible selective amplifications were performed and the ^{33}P labelled products were loaded on a 6% polyacrylamide gel and then visualised on an autoradiographic film.(Figure 17)

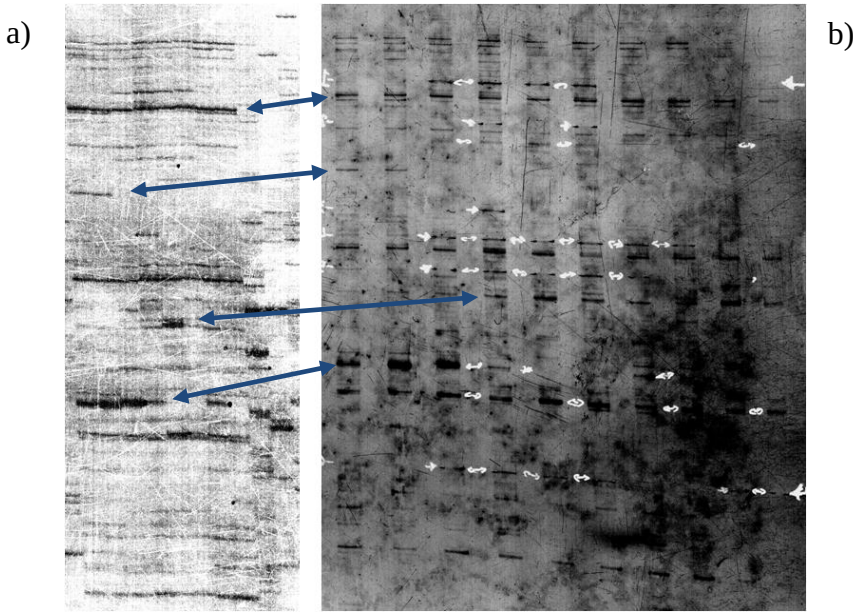


Figure 17 Example of cDNA-AFLP patterns. Autoradiographic films were used for the band visualization. In A the "short" run is shown while in B a "long" run of the same selective PCR with better band resolution is reported. The arrows indicate the corresponding patterns that can be recognised between the two independent runs.

For every primer combination a preliminary analysis was performed to individuate the more interesting differences. Furthermore, two gels with different migration times in the electrophoresis were done in order to obtain a good separation of small framents ("short run" Figure 17a) as well

as of larger fragments (“long run” Figure 17b). This facilitated the selection of interesting bands in a wide range of different fragment sizes. The selective PCR was repeated also to discriminate possible “ghost” bands (bands that are generated unspecifically during the specific amplification of the cDNA-AFLP fragments but that are normally not reproducible. Differentially expressed bands were individuated on the respective autoradiographic films and then excised from the gel (3.2.21).

4.3 Selection of the differentially expressed cDNA-AFLP bands

From the cDNA-AFLP analysis several bands showed differential expression. When a difference between healthy and infected state of a genotype was individuated the same band was then compared in the other genotypes defining a pattern associated with the specific band annotated. About 350 patterns were found in which one or more differences were combined between healthy and infected state of the genotypes (Figure 18; arrows d) corresponding to several hundred of possible differentially expressed bands. In order to reduce the huge work that would be necessary for the analysis of all of them, only those patterns that showed overexpression in the infected resistant genotypes H0909 and *M. sieboldii* (regardless if the other genotypes showed differential expression) were chosen. Subsequently, for each pattern all the bands associated to an increased signal in the infected state were excised. In this way 26 different patterns (Table 21) were individuated and 63 bands (Table 21; red squares) were excised from the blot (Figure 18; arrows a, b, c).

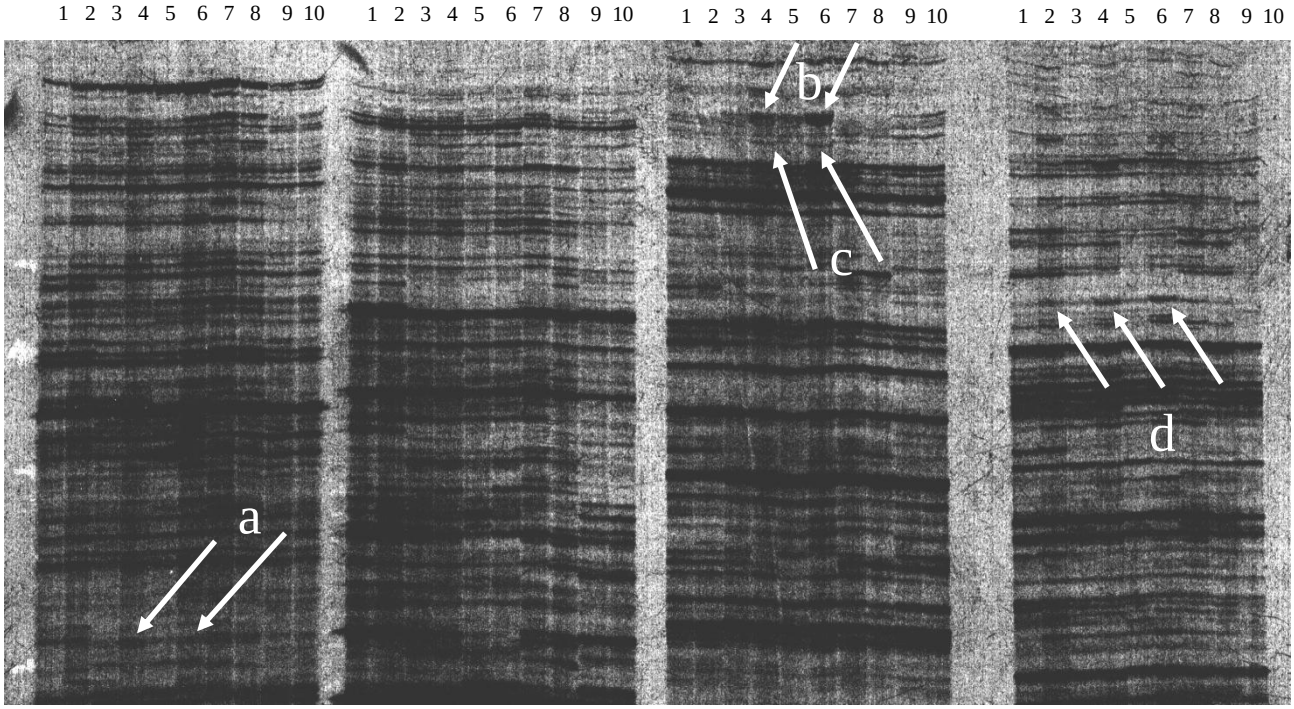


Figure 18 Autoradiogram of a cDNA AFLP experiment, the arrows indicate differentially expressed bands. 1= GDh; 2=GDi; 3=H0909h; 4=H0909i; 5=MSh; 6=MSi; 7=M9h; 8=M9i; 9=D2212h; 10=D2212i. H= healthy; i= infected. a, b, c: differentially expressed bands between heathy and infected resistant genotypes; d: example of pattern of differentially expressed bands

Table 21 Expression patterns associated to the bands selected for the isolation from the blot. Red squares = band that showed a stronger signal; green squares = band that showed a weaker signal; white squares with a minus = no signal, L = faint signal

Primer combination	GD	GDi	H0909	H0909i	MS	MSi	M9	M9i	2212	2212i
CGG 1										
CGG 3	-	-	-				-	-	-	
CGG 6										
CGT 1	L	L					L	L	L	-
CGT 5	-	-						-	-	-
CGT 6	-	-					-	-		
CGT 7								-	-	-
CCA 1									-	-
CCC 1	-	-					-	-	-	-
CCC 2	-							-	-	-
CCG 1	-	-						-		
CCG 2	-	-	-				-	-	L	L
CTC 1									L	L
CTT 2	-	-					-		-	
CAA 1	-	-					-	-	-	-
CAA 2										
TTC 6	-	-	-				-			
TGC 1	-	-	-				-	-	-	-
TCC 1 *	-	-					-	-	-	-
TCC 3 *	-	-					-	-	-	-
TCG 1 *	-	-						L	-	-
TTA 2	-	-					-	-		
TCA 1	-	-					-	-		
TCC 1	-	-	-		-		-	-	-	-
TCC 2	-	-					-	-	-	-
CAT 1 *	-	-					-	-	-	-

* these bands were isolated from a electrophoretic “long-run” ……..

4.3.1 Cloning of the differentially expressed products identified by cDNA-AFLP

All the 63 cDNA fragments were re-amplified with the USER selective primers. The PCR products were checked on a 1% agarose gel (3.2.21) and then separately cloned with the USER PCR cloning kit procedure (3.2.9.2).

4.3.1.1 Selection of clones

For every cloned band 20 colonies were picked and replicated on a new LB agar plate. The clones were amplified through colony PCR (3.2.10) using the general BstYIC0/BstYIT0 – MseI0 primer pair and those clones that showed a product were cultured overnight. After plasmid-miniprep (3.2.11) purification the inserts were sequenced (3.2.15).

4.3.1.2 Sequencing and sequence analysis of differentially expressed genes

The sequence reaction was performed as described in (3.2.15) and the products were analysed from the company GENterprise GmbH. The analysis of similarity at nucleotide level was performed using the NCBI Blast whereas the blast of translated sequences was carried out using the Expasy database. The Table 22 summarises the results obtained from the Blast analyses for the cDNA-AFLP fragments that showed homology with products of known function. The sequences were grouped based on their putative association to cellular mechanisms. Although the major part of the cloned fragments showed similarity to products with unknown function (37%; Figure 19), for all of them homology with a sequence or a translated sequence present in the databases was found. The analysis of the functional groups associated to the differentially expressed genes showed a heterogeneous and broad range of metabolisms related to these sequences (Figure 19). The groups formed by genes associated to similar metabolisms could be further linked in four major functional classes: protein metabolism, energy transport, general stress response and signal transduction. The behaviour and involvement of these genes in the phytoplasma-apple interaction had to be confirmed and proven through further experiments in order to characterise their expression after the infection. This task was performed using the real time PCR technique.

Table 22 Similarity of the sequences obtained from the analysis of the isolated differentially expressed cDNA-AFLP bands. The databases used for the BLAST were NCBI for the nucleotide sequence and Expasy for the translated sequence

Selective primer combination	Homology		Putative function			% of similarity		Identities		E value		Functional class
	NCBI EST	Expasy	NCBI EST	Expasy	Size bp	NCBI EST	Expasy	NCBI EST	Expasy	NCBI EST	Expasy	
CTT	*Malus x domestica*	*Solanum tuberosum*	Ribosomal protein L27A	Ribosomal protein L27a-like protein	323	97%	81%	316/323	62/76 positives 68/76 (89%)	4.00E-164	6.00E-37	Ribosomal protein
CTT	*Malus x domestica*	*A. thaliana*	40S ribosomal protein S18	40S ribosomal protein S18	319	99%	80%	316/319	58/72 positives 59/72 (81%)	2.00E-33	7.00E-172	
CTT	*Prunus dulcis*	*A. thaliana*	UK	60S ribosomal protein L6-2	319	92%	59%	260/281	58/98 positives 70/98 (71%)	2.00E-104	5.00E-24	
CTT	*Prunus persica*	*Medicago truncatula*	Putative 60S ribosomal protein L6	Ribosomal protein L6E	319	88%	60%	276/312	59/98 positives 67/98 (68%)	4.00E-87	2.00E-23	
TCG1	*Malus x domestica*	*Oryza sativa*	40S Ribosomal protein S8	40S Ribosomal protein S8	259	98%	93%	249/254	57/61 positives 57/61 (93%)	5.00E-123	7.00E-34	

TTA2	*Malus x domestica*	NR	60S ribosomal protein L10-3	NR	217	99%	NR	215/217	NR	5.00E-113	NR	
CTT	*Malus x domestica*	*Lycopersicon esculentum*	Arginase	Arginase1	296	98%	96%	292/296	64/66 positives 64/66 (96%)	3.00E-155	2.00E-36	Aminoacid metabolism/transport
CTT	*Malus x domestica*	*Medicago truncatula*	Phenyalanyl-tRNA synthetase beta-subunit	B3/4	317	93%	64%	220/235	50/78 positives 63/78 (80%)	2.00E-92	4.00E-22	
CTT	*Malus x domestica*	*Elaeagnus umbellata* (Autumn olive)	Asparagine synthetase	Asparagine synthetase	319	98%	68%	316/320	62/91 positives 67/91 (73%)	1.00E-164	8.00E-27	
CAA1	*Malus x domestica*	*Brassica juncea* (leaf mustard); *A. thaliana*	Similarity to ISP4 (and glutathione transporter GT1)	Glutathione trasporter GT1; oligopeptide transporter 3 (ATOPT3)	311	99%	83%	297/298	86/103 positives 98/103 (95%); 86/103 positives 97/103 (94%)	2.00E-163	4.00E-52	
TCC2	*Malus x domestica*	*A. thaliana*	Aspartate aminotransferase mitochondrial precursor	Aspartate aminotransferase, chloroplast precursor (transaminase A)	121	100%	80%	83/83	29/36 positives 31/36 (86%)	4.00E-38	2.00E-14	
CCC1	*Malus x domestica*	*Gossypium hirsutum* (cotton)	Lysine-ketoglutarate reductase/saccharo pine dehydrogenase	Bifunctional lysine-ketoglutarate reductase/saccharo pine dehydrogenase	337	98%	77%	287/290	85/110 positives 96/110 (87%)	6.00E-154	6.00E-41	
CAA2	*Malus x domestica*	*A. thaliana*	AKIN gamma CBS domain-containing protein	AKIN gamma; SNF1-related protein kinase regulatory subunit gamma1	144	100%	93%	144/144	44/47 positives 44/47 (93%)	2.00E-74	9.00E-24	Kinase function
CAA2	*Malus x domestica*	*Medicago truncatula; Nicotiana tabacum*	UK; Shaggy related protein kinase theta (ASK-Theta) (serine/threonine specific)	Glycogen synthase kinase3 homolog MSK-3; Shaggy related protein kinase NtK-1	144	100%; 100%	100%; 100%	63/63; 60/60	20/20; 20/20	5e-26; 3e-24	4e-10; 4e-10	
CCG1	*Malus x domestica*	*Oryza sativa; A. thaliana*	LRR protein family, extensine-like protein precursor	OSJNBb 011K07.13 protein; similarity to receptor prot kinase	200	98%; 92%	63%; 66%	162/165; 183/198	41/65 positives 51/65 (78%); 43/65 positives 50/65 (76%)	8e-75; 1e-70	2E-21; 3e-19	
CCG2	*Malus x domestica*	*A. thaliana*	UK	Leucine-rich repeat receptor protein kinase EXS precursor	185	100%	83%	185/185	51/61 positives 53/61 (86%)	1.00E-98	3.00E-27	
CCA1	*Malus x domestica*	*A. thaliana*	UK	Receptor like protein kinase	498	99%	74%	449/450	123/165 positives 144/165 (86%)	0	2.00E-69	
CTT	*Malus x domestica*	*A. thaliana*	UK	Palmitoyltransferase TIP1 Ankyrin repeat-containing S palmitoyltransferase	316	95%	82%	164/172	32/39 positives 37/39 (94%)	3.00E-69	5.00E-19	Growth regulator
CTT	*Glycine max*	*A. thaliana*	GH3 like protein	Indole-3-acetic acid-amido sinthetase GH3 6 auxine response GH3-like protein	310	83%	85%	183/220	88/103 positives 92/103 (89%)	5.00E-31	4.00E-48	
TTC6	*Malus x domestica*	NR	Brassinosteroid-regulated protein BRU1	NR	95	98%	NR	81/82	NR	7.00E-33	NR	
CTT	*Malus x domestica*	*A. thaliana*	UK	SAL1 phosphatase FIERY1 protein	317	99%	79%	210/211	49/62 positives 55/62 (88%)	4.00E-99	2.00E-24	Stress response
TCC*1	*Malus x domestica*	NR	Galactinal synthase	NR	292	90%	NR	175/193	NR	4.00E-59	NR	
TCC*3	*Malus x domestica*	*Capsicum chinense* (Bonnet pepper)	UK	Glutathione-S-transferase GST1	120	97%	67%	97/99	23/34 positives 30/34 (88%)	7.00E-43	3.00E-12	
TCA1	*Malus x domestica*	NR	Catalase	NR	166	95%	NR	159/167	NR	7.00E-50	NR	
TCC2	*Malus x domestica*	*Vitis vinifera*	Putative wound-induced protein	Wound-induced protein-like (fragment)	163	99%	48%	155/156	18/37 positives 23/37 (62%)	4.00E-79	7.00E-05	
CGG6	*Malus x domestica*	*Nicotiana tabacum*	Putative chitinase	Chitinase class V / lysozyme	285	98%	58%	280/285	49/84 positives 60/84 (71%)	3.00E-146	2.00E-22	
CGT5	*Malus x domestica*	*Glycine max*	Chitinase class I	Chitinase class I	289	90%	85%	262/289	82/96 positives 87/96 (90%)	4.00E-95	1.00E-44	
CTC1	*Malus x domestica*	*Oryza sativa*	UK	Formate dehydrogenase mitocondrial precursor	302	100%	91%	301/301	91/100 positives 97/100 (97%)	7.00E-168	4.00E-47	

CTT	*Malus x domestica*	*Malus x domestica*	UK	Putative DNAJ protein	318	99%	98%	310/311	83/84 positives 84/84 (100%)	3.00E-171	8.00E-44	Protein folding
TGC1	*Prunus persica*	*A. thaliana*	Similar to heat shock protein cpHsc70-1 (*A. thaliana*)	Hsp70-like protein	156	96%	88%	149/155	46/52 positives 49/52 (94%)	1.00E-66	5.00E-24	
CCG2	*Prunus persica*	*Medicago truncatula*	Similar to Hsp70	HSP70 kDa	186	89%	86%	134/147	53/61 positives 58/61 (95%)	7.00E-38	4.00E-30	
CCG2	*Malus x domestica*	*Vicia faba*	Immunophilin	Peptidyl-prolyl isomerase FKBP12 (immunophilin)	186	97%	85%	182/186	40/47 positives 43/47 (91%)	9.00E-90	1.00E-22	
CAA1	*Malus x domestica*	*Oryza sativa*	Putative vacuolar protein sorting associated	Vacuolar protein sorting associated protein VPS16-like(intracellular protein transport)	326	99%	58%	209/211	45/77 positives 52/74 (67%)	2.00E-109	7.00E-15	Protein transport
CTT	*Malus x domestica*	*Ipomoea batatas* (sweet potato)	UK	Papaine-like cysteine proteinase isoform	322	95%	83%	249/261	62/74 positives 68/74 (91%)	5.00E-114	1.00E-33	Protein degradation/turnover
TTC6	*Malus x domestica*	*A. thaliana*	UK	Peptidylprolyl isomerase	88	100%	72%	48/48	18/25 positives 20/25 (80%)	3.00E-17	2.00E-06	
TCC*1	*Malus x domestica*	*A. thaliana*	PRT1 protein	Ubiquitin protein ligase PRT1 (proteolysis 1 protein)	290	100%	67%	150/150	47/70 positives 57/70 (81%)	1.00E-77	2.00E-25	
TTA2	*Malus x domestica*	*Triticum aestivum*	UK	Ubiquitin	218	100%	100%	218/218	72/72	2.00E-118	6.00E-42	
TTA2	*Malus x domestica*	*Triticum aestivum*	UK	Ubiquitin	219	99%	98%	218/219	60/61 positives 61/61 (100%)	4.00E-114	2.00E-34	
TCC1	*Malus x domestica*	*A. thaliana*	PRT1 protein	Ubiquitin protein ligase PRT1 (proteolysis 1 protein)	296	100%	68%	150/150	48/70 positives 57/70 (81%)	1.00E-77	7.00E-26	
CCC2	*Malus x domestica*	*Oryza sativa*	UK	Putative aminopeptidase C	147	100%	85%	142/142	40/47 positives 43/47 (91%)	3.00E-73	9.00E-26	
CCC2	*Malus x domestica*	*Oryza sativa*	F1707.15 protein (UPL2)	E3 ubiquitin protein ligase UPL1, putative expressed	216	97%	91%	210/216	66/72 positives 69/72 (95%)	8.00E-103	3.00E-40	
CGG1	qi 48417146 qb CN944333.1 Royal Gala	Q8GX91 *A. thaliana*	UK	pepsin/ASP-protease	517	95	73	146/153	107/146 positives 123/146 (83%)	5.00E-62	2.00E-56	
CCA1	*Malus x domestica*	*A. thaliana*	Aminopeptidase-like protein	Aminopeptidase-like protein	500	97%	72%	435/447	120/166 positives 144/166 (86%)	0	1.00E-65	
TCC*3	*Prunus persica*	*A. thaliana*	cDNA similar to WD-40 repeat family protein/beige related	F1003.12 protein	120	96%	85%	116/120	34/40 positives 36/40 (90%)	1.00E-50	2.00E-19	Protein-protein interaction
CGG6	*Malus x domestica*	*Oryza sativa*	Expressed protein	DENN (AEX-3) domain-containing protein-like	285	100%	82%	251/251	78/94 positives 88/94 (93%)	6E-138	7.00E-40	
TCG1	*Malus x domestica*	NR	Keech repeat containing F-box protein family	NR	82	98%	NR	81/82	NR	3.00E-35	NR	
TCA1	*Malus x domestica*	*Malus x domestica*	Sorbitol dehydrogenase	Sorbitol dehydrogenase	221	99%	100%	219/221	68/68	6.00E-113	3.00E-44	Sugar modification
TTA2	*Malus x domestica*	*Cyamopsis tetragonoloba* (Cluster bean)	UDP-galactose 4-epimerase	UDP-glucose 4-epimerase (GEPI 42)	217	98%	78%	195/198	37/47 positives 41/47 (87%)	1.00E-101	9.00E-21	
TTA2	*Malus x domestica*	*Fragaria ananassa*	UK	UDP-glucose glucosyltransferase (GT7)	216	91%	63%	152/167	46/72 positives 55/72 (76%)	4.00E-52	3.00E-24	
CGT6	*Malus x domestica*	*Cyamopsis tetragonoloba* (Guar)	Similar to GAE1_PEA UDP-galactose 4-epimerase	UDP-glucose 4-epimerase GEPI42	221	99%	78%	115/116	37/47 positives 40/47 (85%)	4.00E-55	8.00E-16	
TGC1	*Malus x domestica*	*Pisum sativum* (Garden pea)	UK	UDP-D-glucuronate carboxylyase	157	99%	56%	156/157	28/50 positives 31/50 (62%)	1.00E-79	2.00E-09	

TGC1	*Malus x domestica*	*Medicago truncatula*	Hypothetical 25.7 KD protein	Sugar transporter superfamily	157	94%	80%	98/104	17/21 positives 18/21 (85%)	4.00E-36	1.00E-06	Sugar transporter
TTC6	NR	*Malus baccata*	NR	Malus baccata putative TIR-NBS type R protein 4 (R4) gene	91	NR	70%	NR	21/29 positives 22/29 (75%)	NR	1.00E-08	Signal transduction
TCC*1	*Malus x domestica*	NR	Calmodulin related protein	NR	340	97%	NR	125/128	NR	7.00E-55	NR	
TCG1	*Malus x domestica*	*A. thaliana*	Copine like protein copine BONZAI1	BONZAI1 copine like protein	372	95%	89%	336/352	83/93 positives 88/93 (94%)	7.00E-160	3.00E-38	
CTT	*Malus x domestica*	*Vitis aestivalis* (Grape)	Transcription factor WRKY4	Putative WRKY transcription factor	317	94%	83%	288/306	85/102 positives 91/102 (89%)	2.00E-126	6.00E-44	Transcription factors
CTT	*Malus x domestica*	*Medicago truncatula*	UK	Zinc finger DHHC type	318	99%	75%	171/172	72/96 positives 81/96 (84%)	2.00E-88	2.00E-34	
TTC6	*Malus x domestica*	*A. thaliana*	Putative DNA binding protein	Putative Myb family transcription factor (MYB transcription factor)	92	100%	90%	92/92	27/30 positives 28/30 (93%)	1.00E-43	1.00E-18	
TCC*3	*Malus x domestica*	NR	Zinc finger like protein	NR	122	97%	NR	119/122	NR	8.00E-52	NR	
TCG1	*Malus x domestica*	*Oryza sativa*	Hypothetical protein p85RF	Putative zinc finger protein	85	97%	68%	83/85	22/32 positives 23/32 (71%)	1.00E-34	8.00E-11	
TCC1	*Malus x domestica*	*Medicago truncatula*	CCCH-type zinc finger protein-related (*A. thaliana*)	Zinc finger CCCH-type	259	98%	51%	254/259	22/43 positives 27/43 (62%)	8.00E-131	6.00E-06	
TTA2	*Malus x domestica*	NR	Nucleotide binding protein	NR	219	99%	NR	210/212	NR	5.00E-110	NR	
CCC2	*Malus x domestica*	*Medicago truncatula*	UK	SNF2-related (transcription activator) HMG-I and HMG-Y DANN binding Bromodomain ATP-requiring DANN elicase RecQ	368	98%	54%	278/281	67/122 positives 81/122 (66%)	4.00E-146	7.00E-29	
CGT1	qi 51095612 qb CO865462.1 *Malus x domestica*	Q9FZ14 potato	MYB-related protein	Tuber-specific and sucrose-responsive element binding factor	446	97	55	391/402	45/81 positives 60/81 (73%)	0	8.00E-17	
CTT	*Malus x domestica*	*Medicago truncatula*	Putative cytochrome P450	E-class P450 groupI	319	85%	56%	239/281	51/91 positives 73/91 (80%)	2.00E-55	9.00E-26	Electron transport/Redox reactions
CAA2	*Malus x domestica*	NR	Iron sulfur subunit of succinate dehydrogenase precursor	NR	144	94%	NR	139/137	NR	2.00E-53	NR	
TCC*1	*Malus x domestica*	*Medicago sativa* (alfalfa)	Squalene monooxygenase	Squalene epoxydase (SE1)	291	99%	86%	290/291	83/96 positives 90/96 (93%)	2.00E-159	5.00E-46	
TCC*1	Malus x domestica	NR	Geranylgeranyl hydrogenase	NR	295	98%	NR	182/185	NR	1.00E-62	NR	
TCC*3	*Malus x domestica*	*A. thaliana*	Glycolate oxidase	Probable peroxisomal (S)-2-hydroxy-acid oxidase 2 (glycolate oxidase 2) (GOX2); short chain alpha-hydroxy acid oxidase 2	120	97%	85%	117/120	34/40 positives 37/40 (92%)	5.00E-23	1.00E-20	
TCC2	*Malus x domestica*	*A. thaliana*	Glycolate oxidase	Glycolate oxidase-like protein	120	97%	91%	117/120	22/24 positives 22/24 (91%)	5.00E-53	8.00E-11	
TCC1	*Malus x domestica*	*Medicago sativa* (alfalfa)	Putative squalene monooxygenase 2	Squalene epoxydase (SE1)	291	99%	85%	289/291	82/96 positives 89/96 (92%)	5.00E-157	2.00E-45	
CCG1	*Malus x domestica*	*A. thaliana*	Cytochrome b6-f complex chain petM precursor	Cytochrome b6-f complex subunit precursor (petM)	201	100%	93%	164/164	30/32 positives 31/32 (96%)	4.00E-86	7.00E-12	
CTT	*Malus x domestica*	*A. thaliana*	UK	Probable histone H2A variant 2	318	99%	89%	304/305	53/59 positives 57/59 (96%)	1.00E-167	2.00E-22	Cell cycle

TTA2	*Malus x domestica*	*Oryza sativa*	Cyclin-E binding protein 1	BNR repeat domain protein-like	215	99%	62%	164/165	34/54 positives 41/64 (75%)	2E-84	3.00E-16	
CGT7	*Triphysaria versicolor*	*Sesbania rostrata*	H+Transport ATPase -	P-type H+-ATPase (heavy metal transmembr.transport)	160	90%	96%	144/149	51/53 positives 52/53 (97%)	7.00E-43	1.00E-28	Ion uptake/metal detoxyfication
TCC*3	*Malus x domestica*	*Prunus persica*	Vacuolar H+ pyrophosphatase	Vacuolar proton-inorganic pyrophosphatase (HVP1)	122	98%	100%	111/113	19/19	8.00E-49	8.00E-07	
TCC2	*Malus x domestica*	NR	Metallothionein-like protein type2	NR	121	98%	NR	77/78	NR	1.00E-32	NR	
CCA1	*Malus x domestica*	*Malus x domestica*	Metallothionein-like protein	Metallothionein-like protein	502	98%	100	388/394	54/54	0	5.00E-27	
CCC2	*Malus x domestica*	*Pyrus pyrifolia* (Japanese pear)	Metallothionein-like protein	Metallothionein-like protein	365	100%	77%	210/210	34/44 positives 37/44 (84%)	2.00E-113	2.00E-16	
CTT	*Sus scrofa*	*Pseudomonas fluorescens*	UK	Lipoprotein putative precursor	320	99%	80%	318/320	71/88 positives 75/88 (85%)	3.00E-174	5.00E-34	Lipoprotein
CAA1	*Malus x domestica*	NR	GAPDH	NR	162	98%	NR	159/162	NR	4.00E-78	NR	Enzyme GAPDH
CAA2	*Malus x domestica*	NR	GAPDH	NR	163	98%	NR	160/162	NR	3.00E-80	NR	
TTC6	*M.sieboldii*	*Oryza sativa*	EF1alpha	EF1alpha	77	100%	100%	69/69	23/23 positives 23/23 (100%)	6.00E-30	1.00E-13	Translation machinery
TCC1	*Malus x domestica*	*Lycopersicum esculentum*	EF2	EF2	212	97%	100%	205/211	62/62	2E-97	6.00E-36	
CTT	*Prunus persica*	*A. thaliana*	Similar to pir T49035 acid phosphatase	Purple acid phosphatase (PAP22)	316	87%	82%	262/300	85/103 positives 97/103 (94%)	2.00E-76	8.00E-49	Phosphatase
CTT	*Malus x domestica*	*Medicago truncatula*	UK	Las1-like	319	97%	47%	160/164	22/46 positives 30/46 (65%)	2.00E-76	2.00E-07	Plant morpjogenesis
TCC*1	*Malus x domestica*	*Nicotiana benthamiana*	UK	NRG1	291	96%	62%	282/291	27/43 positives 31/43 (72%)	6.00E-138	1.00E-11	N protein, resistance response
TCC*3	*Populus trichocarpa*	*Medicago truncatula*	UK	Regulator of chromosome condensation / beta-lactamase-inhibitor prot III	122	86%	77%	72/83	17/22 positives 19/22 (86%)	8.00E-12	5.00E-08	Nuclear transport
TCC*3	*Malus x domestica*	*Orysa sativa*	UK	Molybdenum cofactor sulfurase protein-like	123	98%	63%	114/116	23/36 positives 24/36 (66%)	1.00E-50	3.00E-06	Sulfur cycle
TCC2	*Malus x domestica*	*Oryza sativa*	UK	Molybdenum cofactor sulfurase protein-like	123	99%	63%	115/116	23/36 positives 24/36 (66%)	5E-53	3.00E-06	
TCG1	*Malus x domestica*	*Populus tremula*	Pumilio like protein; mRNA post-transcriptional fate regulation	Pumilio domain-containing protein PPD1	298	98%	82%	295/298	82/99 positives 86/99 (86%)	9.00E-159	5.00E-42	RNA binding
TCC2	*Malus x domestica*	*Medicago truncatula*	Putative DNA polymerase	Aminotransferase class II	110	98%	68%	107/109	24/35 positives 29/35 (82%)	7.00E-49	9.00E-13	Transferase
TTA2	*Malus x domestica*	*Populus tremula*	Plasma membrane major intrinsic protein 2	Putative plasma membrane intrinsic protein (pip 2.1)	218	97%	55%	177/182	19/34 positives 22/34 (64%)	1.00E-82	2.00E-04	Membrane protein
CCC1	*Malus x domestica*	*A. thaliana*	Hypothetical protein	COBRA-like protein 9 precursor	392	98%	57%	383/390	26/45 positives 36/45 (80)	0.00E+00	2.00E-08	Cell development
CCC1	*Prunus persica*	NR	Acetyltransferase-like protein	NR	390	88%	NR	80/90		8.00E-18	NR	Molecular function

CCC1	*Malus x domestica*	*Medicago truncatula*	UK	MTD1	396	97%	50%	174/178	4/91 positives 61/91 (67%)	3E-82	6.00E-15	
CCC2	*Malus x domestica*	*Medicago truncatula*	Fibrillin	PAP fibrillin	366	98%	66%	323/328	48/72 positives 51/72 (70%)	1.00E-170	3.00E-18	Membrane interaction
CCC2	*Malus x domestica*	*Medicago truncatula*	Putative lectin	Concavaline A-like lectine/glucanase	348	98%	46%	332/337	35/76 positives 48/76 (63%)	8.00E-175	7.00E-15	
CCG2	*Malus x domestica*	*Alnus glutinosa*	THI 4-ALNGL Thiazole byiosinthetic enzyme, chloroplast precursor (AG6)	Thiazole biosynthetic enzyme chloroplast precursor (AG6) (THI4)	166	99%	67%	165/166	27/40 positives 28/40 (70%)	4.00E-85	9.00E-11	Secundary metabolism
CGG1	qi 51564881 qb CV085532.1 *Malus x domestica*	Q5TIN3 *Gossypium raimondii* (cotton)	Hypothetical 53.1 KD protein	Beta-6-xylosyltransferase; TIGR:UK	518	99%	96%	380/381	140/161 positives 151/161 (92)	0.00E+00	4.00E-82	Cell wall modification
CGG3	qi 53858717 qb CV525043.1 *Malus x domestica*	Q9SY97 *A. thaliana*	Chlorophyll A-B binding protein 8 precursor	PSI type III chlorophyll a/b-binding protein	389	98%	89%	382/389	79/88 positive 83/88 (93%)	0.00E+00	3.00E-42	Chlorophill a/b binding protein

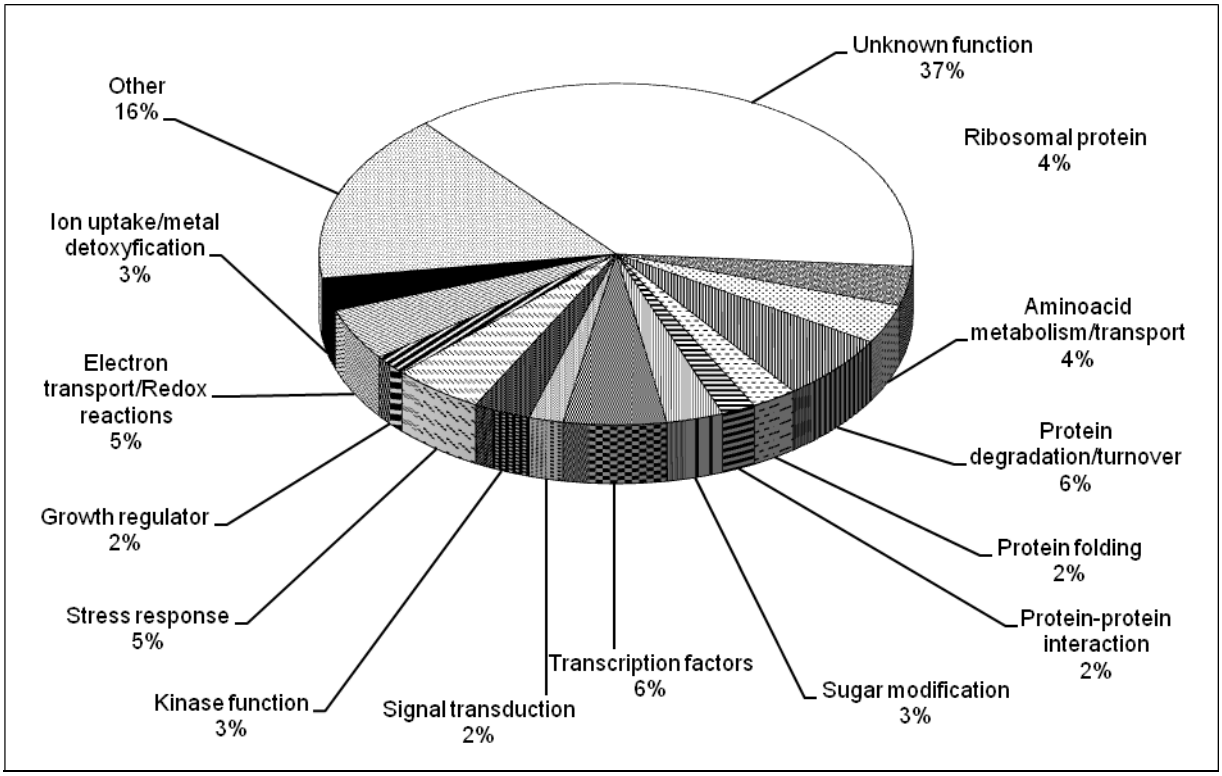


Figure 19 Schematic grouping of the sequences of the differentially expressed genes based on their putative function

4.4 Study of the differential gene expression by real-time PCR analysis

Several cDNA fragments were individuated as differentially expressed in the cDNA-AFLP analysis on infected and healthy plants. In order to validate the differential expression of the genes corresponding to these fragments, a real-time PCR analysis was carried out to investigate their expression levels after the phytoplasma infection. As shown in Table 22, different sequences were obtained for a given cloned cDNA-AFLP fragment. As it was not possible to know which sequence was associated to the differentially expressed product, the real-time PCR analysis was first focused on those genes represented by sequences of inserts found in several clones and present in clones obtained from different genotypes. Specific primers were then designed on these sequences.

Although, as previously reported, the *in vitro* system allows a high standardisation of the growth and maintenance parameters, the gene expression can be influenced not only by the presence of the phytoplasma but also by the physiological state of the plant at the timepoint of sampling. Therefore, in order to confirm those genes that were differentially expressed due to the phytoplasma infection an independent experiment was carried out on a new set of infected and healthy *in vitro* plants. RNA extraction and cDNA synthesis were performed as described for the preparation of the starting cDNA in the cDNA-AFLP analysis (3.2.2; 3.2.12; 3.2.13; 3.2.14).

4.4.1 Primer design for the real-time analysis

The real-time analysis was initially carried out on a first set of expressed sequences. These were selected between those obtained from the same type of cDNA-AFLP band and that were homologous in different genotypes. 23 sequences were individuated (Table 23) on which specific primers were developed using the web form and the parameters described in 3.2.19.

Table 23 cDNA-AFLP fragments selected for the real-time qPCR analysis

Band code	cDNA-AFLP primer combination from which the band was identified		Putative function (based on nucleotidic sequence)	Putative function (based on aminoacidic sequence)
	BstYI primer	MseI primer		
3	C-T	MseI-C	Unknown	Unknown
5	C-G	MseI-T	Chitinase class I	Chitinase class I
6	C-G	MseI-G	Chlorophyll A-B binding protein 8 precursor	Chlorophyll a/b binding protein type III
7	C-C	MseI-C	Unknown	COBRA-like protein 9 precursor

9	C-C	MseI-A	Metallothionein-like protein	Metallothionein-like protein
11	C-C	MseI-A	Aminopeptidase-like protein	Aminopeptidase-like protein
14	C-G	MseI-T	Unknown	Unknown
15	C-C	MseI-G	LRR protein family	Similarity to receptor protein kinase
21	C-C	MseI-G	Unknown	Unknown
22	C-G	MseI-G	Unknown	pepsin/ASP-protease
23	C-G	MseI-T	H+Transport ATPase	P-type H+-ATPase (heavy metal transmembrane transport)
24	C-A	MseI-A	AKIN gamma CBS domain-containing protein	AKIN gamma CBS domain-containing protein
25	T-C	MseI-C	Squalene monooxygenase	Squalene monooxygenase
26	C-T	MseI-T	Unknown	Hypothetical protein At5g58510
27	T-C	MseI-G	Copine like protein copine BONZAI1	Copine like protein copine BONZAI1
28	C-A	MseI-A	Unknown	Unknown
29	T-G	MseI-C	Hsp70-like protein	Hsp70-like protein
30	T-G	MseI-C	NR	NR
31	T-C	MseI-A	Unknown	Unknown
32	T-T	MseI-A	Unknown	Unknown
33	T-C	MseI-C	Glycolate oxidase	Glycolate oxidase
35	T-C	MseI-C	Putative squalene monooxygenase 2	Putative squalene monoxygenase

4.4.1.1 Proof of the presence of the analysed expressed product in the different genotypes

Specific primers were developed on the 5' and 3' extremity of the cDNA-AFLP fragment chosen for the real-time PCR study (Table 24). The primers were tested (3.2.5) on the cDNA synthesised from *Malus sieboldii*, H0909 and M9 as described in 3.2.13. The PCR products were then directly sequenced as described in 3.2.15. In all the genotypes a product corresponding to the original sequence was found except for the primer pair 30 for which a PCR product was obtained only in H0909 and *Malus sieboldii*.

Table 24 Primer used to test the presence in the different genotypes of the transcripts corresponding to the selected cDNA-AFLP differentially expressed bands

Primer name	Forward	Reverse
Band 3	GGTATTGCGGTCGGGATG	CCTTCATCGGAATAAGAAGGTG
Band 5	AGATCATTTCCGATTGCCTTG	GGGTTTGGCACAACTGGAG
Band 6	CCCATTAGGGCAAGTTCAAAG	AAGGAGCGGACAGACAGTTG
Band 7	GAACTGGCCTCCGATCTTG	CACTCCCAAGAATGCTACCC
Band 9	GTAGGACGTGCAGCTCAGAAG	AACAGAGCTTGCGTTTATCTGC
Band 11	TCCTTCATCCACTTCCATGC	TGCAGAGAGGCTTGGTTGG
Band 14	AAGTACAACACGCTGCGTACC	TCTTGCATGGCTATGGAGTG
Band 15	ACGATCTTGGCTCCTCACC	AACTCGCGTCGAATACGG
Band 21	CAGAGCCATCGTCAACAGAAC	CCTTGAACCGTGAAAGTGG
Band 22	GTTACCGCTCCCGAAATC	CACCTAGGCCAAGCTTTTG
Band 24	TACGTTGTGGACGATGATGG	TTCTGAGGCAGCGGTAGAAC
Band 25	GGGGAACCTGAGGAGCAAC	AAATCTCCGCCGCAATCTC
Band 26	CTGAACCCCTCCTGGTTAG	CAGAAACTTCACATGCTTGC
Band 27	CAAGAAACCAAAGATGCCCTTG	CTGCGGTCATCCGAAGTCTC
Band 28	GGGGTGCTTTTGTAGCTTATG	ATATCAAAGGGCAAACCAGTTG
Band 29	CTTGACAGGCTTAGAACACCAG	CGTAACATTTGGTTCCTTCC
Band 30	TCATTGAAACCACAAGACTGG	AATTCCAACGACCCCTTG
Band 31	CGGTGTGAATCAAGATAATCAG	ATCCGATAGCACTCTCTATGAG
Band 32	ATCAAAGAAAGGCACTAAAAGG	TACTTGCTTACTGGAGTTGGAG
Band 33	GGAAAGATGGATAAGACCAATG	TTGGTAATTTGGTGATTGTCTG
Band 35	GGGGAACCTGAGGAGCAAC	AAATCTCCGCCGCAATCTC

4.4.1.2 Development of gene-specific primers for the real-time PCR analysis

For real-time qPCR analysis, gene-specific primers were selected for each gene. Because the detection method chosen for the gene expression analysis was SYBR Green the development of good quality oligonucleotides was crucial for the production of reliable results. The conditions that had to be fulfilled were more strictly as for the common primer (3.2.19). Important parameters were:

- size of the amplicon: has to be between 120 and 220 bp
- annealing temperature: was set in a range between 58 and 63°C with an optimal value of 60°C
- primer dimers: the possibility of an inter- and intra-primer annealing was set very low

In order to reduce biases related to the size of the amplicon (due to differences in the fluorescence signal intensity and in the Taq amplification efficiency) only primers able to generate a product between 120 and 220 bp were selected. The annealing temperature plays an important

role because while at high temperature the possibility of dimers formation is reduced the chances to find a putative primer pair with the oligo design algorithm are also lowered. Finally, the parameter controlling the number of possible intra- and inter-primer base pairing was maintained low giving a very stringent selection of the possible primer pair. However, in few cases the primer quality resulted non optimal and due to the short length of the corresponding cDNA-AFLP fragment it was not possible to develop a further primer pair.

To test the suitability of the primers (Table 25) for the qPCR, first a gradient PCR was performed using for each primer pair the corresponding cloned cDNA fragment as template. The primers that showed a good amplification and a unique product at 60°C were further validated. A PCR was then carried out using as template a dilution series (3.2.20.3) of the respective plasmid using 60°C as annealing temperature. Only the primers showing a unique PCR product (Figure 20) in all the diluted samples without the formation of primer dimers were chosen and used in the real-time qPCR.

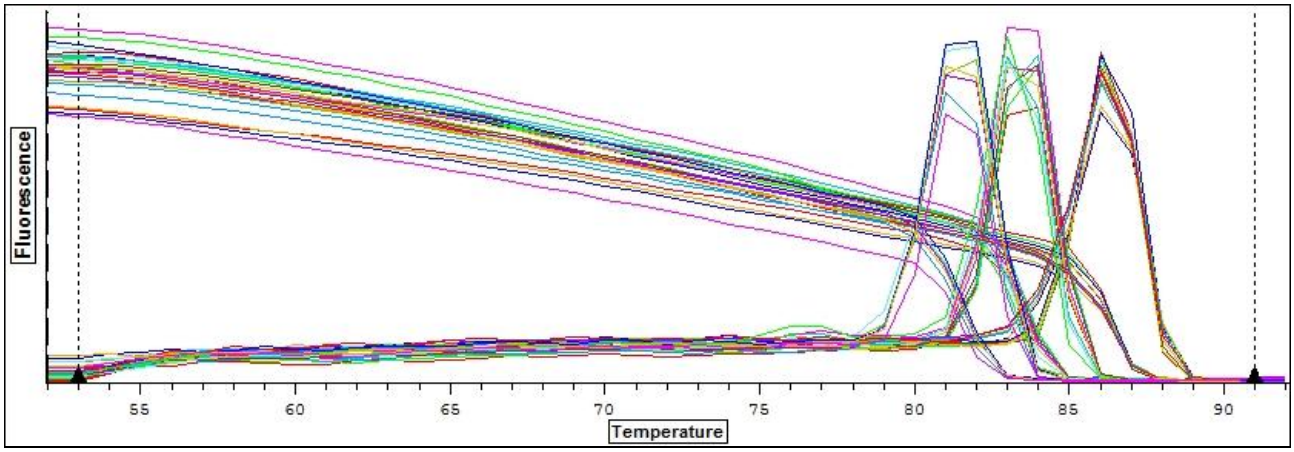


Figure 20 Melting curve analysis of the PCR products generated in the real-time qPCR reaction. Each peak corresponds to a product generated from a specific primer pair in several samples. The presence of a unique peak associated to each curve indicates a high specificity and good quality of the primer pair

Table 25 Primers used for the real-time qPCR analysis

Gene	Forward	Reverse
Gene 3	CAAACGAAGTCACAAGTTCTCC	CGCAATTTTGAGATGACAGAAG
Gene 5	ATATTGACCGGATGGTGTGC	GCTGCTTTCTTGGCTCAAAC
Gene 6	AGTCTCGGGCGGGATTAG	CTCAGACCCAGAAGGCACAG
Gene 7	GAACTGGCCTCCGATCTTG	GACAGGAAGTCGCCGTTG
Gene 9	ACACGACCTTCGTAAATTTGG	CCTGCAACTGCAAGTGAGAG
Gene 11	TCCTTCATCCACTTCCATGC	TAGCGTCTTGTCCCGATCC
Gene 14	CCCAGAAGAACTCATCATTTAGC	CTTGCCGATAGGGTCAGC
Gene 15	CCTCACCTGCCCATTACTTC	AACTCGCGTCGAATACGG
Gene 21	TGGGAACACTTCGATTAGCC	GGCTGTTATACAAACCCTTGAAC
Gene 22	AGCGATGGATTTGTGTTTCC	CCCGATACAGCAGCCTCTC
Gene 24	TGGACGATGATGGGAATCTG	GCTATTCTGAGGCAGCGGTAG
Gene 25	GGAGCAACCACATTTTTCAAG	CTTATGCAAACCACGGACAC
Gene 26	TCCCATCAAAGTTTTCACCAG	TTGCAATATGCATTGAGATGAAG

Gene 27	AATTGTTGGAGTCGGAGGAG	GAAATTGAGCCGGTAGTTCTG
Gene 28	ACATGTTGTGGCTTGTTTCAG	CATGCTAACACATCTCGCTTG
Gene 29	ACCAGTTGAAACCGCCTTG	ACATTTGGTTCCTTCCCTGTC
Gene 30	GGGTGTAGAGATGGGGGTTC	ATTCCAACGACCCCTTGC
Gene 31	TCGGTGTGAATCAAGATAATCAG	CTTGGAAGAGTTGGGATACTTTG
Gene 32	AGGCACTAAAAGGAAGCAAGAG	GAGGGAGCGTGATGTTATTTATG
Gene 33	TGGATAAGACCAATGATTCTGG	AATTGGTAATTTGGTGATTGTCTG
Gene 34	ATGGATAAGACCAATGATTCTGG	AATTGGTAATTTGGTGATTGTCTG
Gene 35	CGGGAACATCTACCAAACAAC	CGCCGCAATCTCTGTACTC

4.4.2 Analysis of the real-time qPCR data

The procedure for the elaboration of the raw data obtained in the qPCR was developed from two different analysis methods previously published by (Peirson *et al*., 2003) (data analysis for real-time PCR; DART-PCR) and (Ramakers *et al.*, 2003) (LinReg PCR). Both methods are based on the direct analysis of the fluorescence plot obtained for each sample in the real-time PCR (Figure 21).

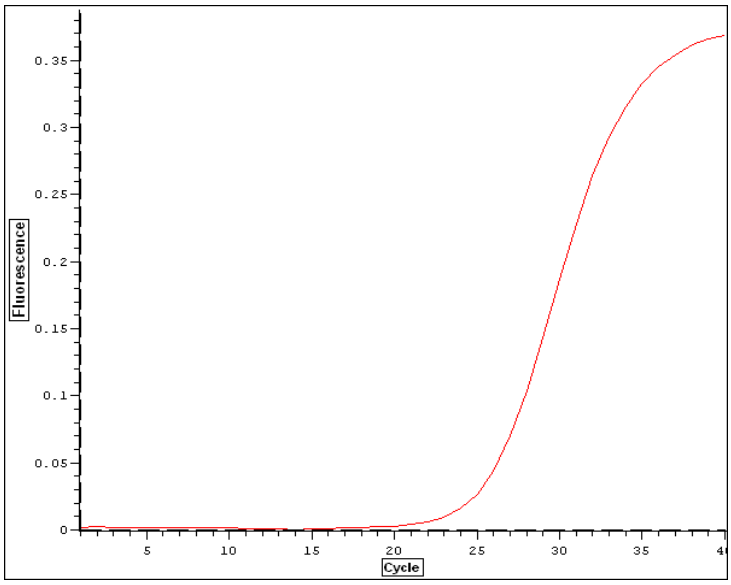


Figure 21 Example of fluorescence plot obtained from the real time PCR. The fluorescence (Y axes) per cycle (X axes) of reaction is reported.

For each reaction the specific efficiency, on which the PCR kinetic depends, can be calculated from the fluorescence plot. The equation that describes the PCR kinetic can be written as

$$N_C = N_o \cdot 2^C \qquad \text{Eq.1}$$

where C is the cycle number, N_C the number of molecules obtained after C cycles and N_o the initial amount of target template. Here the theoretical doubling of the template at each cycle is assumed and therefore the constant 2 is employed in the equation. Commonly this does not represent the real behaviour of the PCR in which the reaction efficiency is below 100%. Therefore, the Eq.1 can be modified in

$$N_C = N_o \cdot E^C \qquad \text{Eq.2}$$

with *E* assuming values between 1 (0% efficiency) and 2 (100% efficiency). DART-PCR and LinReg PCR are both based on the calculation of the real efficiency value *E*. The Eq. 2 can be represented in the logarithmic form (Figure 22):

$$\log N_C = \log N_o + C \cdot \log E \qquad \text{Eq.3}$$

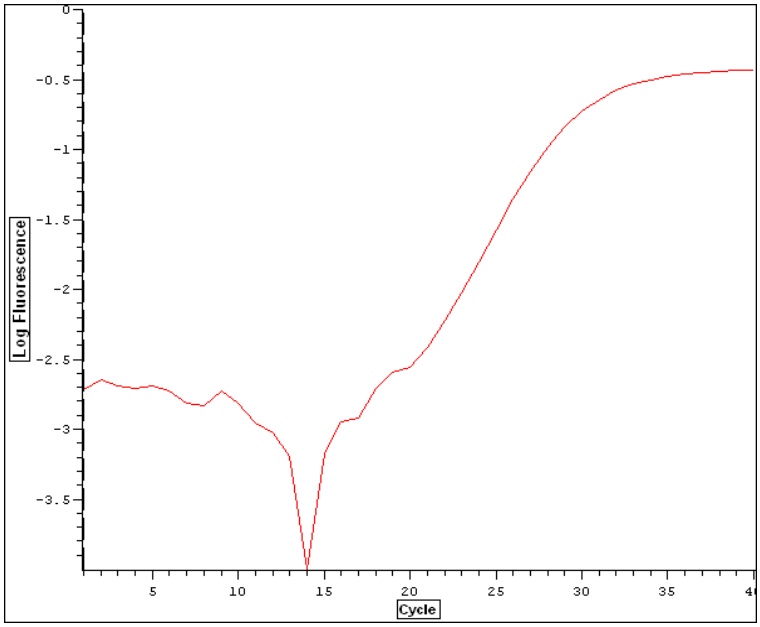


Figure 22 Example of a logarithmic plot of the fluorescence plot obtained in the real-time PCR (Y axes = log values of the fluorescence; X axes = cycle number).

The Eq. 3 can be rearranged as follows

$$C = -\frac{1}{\log E} \cdot \log N_o + \frac{\log N_C}{\log E} \qquad \text{Eq.4}$$

this equation represents a linear equation with slope

$$-\frac{1}{\log E} = m \qquad \text{Eq.5}$$

and the efficiency *E* value can be calculated as

$$E = 10^{-\frac{1}{m}} \qquad \text{Eq.6}$$

The slope *m* can be derived from the logarithmic form of the fluorescence plot. Although in both the DART and the LinReg methods *m* is calculated on the same region of the logarithmic curve the way to individuate the linear range is different. In the DART-PCR method a 10-fold range is defined around the midpoint (obtained with the equation $M = R_{noise} \times \sqrt{(R_{max}/R_{noise})}$ where R represents the fluorescence value and R_{noise} is calculated as the standard deviations of the fluorescence values of the first 10 cycles) of the fluorescence plot calculated in a statistical manner. The slope of the corresponding logarithmic curve is then calculated on these points. The LinReg PCR method uses an algorithm that defines the best range of points on the logarithmic curve of the fluorescence plot in which the linear regression has the best correlation coefficient (R^2). In this case the range of points on which the linear regression is performed can vary depending on the settings but in any case it is always calculated between 3 and 6 data points. In both methods once *m* (and consequently *E*) is defined the initial amount of fluorescence present in the reaction can be calculated as follows:

DART-PCR

In this method, the midpoint (M) of the fluorescence plot is chosen as threshold. Two values are associated to this point: a fluorescence value and the corresponding cycle Ct at which this value is reached in the real-time PCR reaction. Substituting C with Ct in Eq. 2 and the fluorescence intensity *I* to the quantity *N,* Eq. 2 becomes

$$I_{Ct} = I_o \cdot E^{Ct} \qquad \text{Eq.7}$$

I_{Ct} = fluorescence at cycle Ct
I_o = fluorescence at cycle 0
E = efficiency value (between 1 and 2)

The intensity at cycle 0 can be calculated as

$$I_o = I_{Ct} / E^{Ct} \qquad \text{Eq.8}$$

LinReg PCR

In the LinReg PCR method the equation 3 is used to define the intercept at cycle 0 (C = 0) of the linear regression calculated on the linear phase of the log plot.

$$\text{Intercept value} = \log I_o \qquad \text{Eq.9}$$

therefore

$$I_o = 10^{\text{intercept value}} \qquad \text{Eq.10}$$

One strong point of the DART-PCR method is how the efficiency values are statistically analysed after their calculation. Efficiencies calculated for samples amplified with the same primer pair are compared through ANOVA in order to proof if a unique E value can be used in all the samples for the calculation of each I_0. In this way, a reduction in the variability of the data can be obtained. Nevertheless, the LinReg PCR has a much powerful way to define the efficiency individuating the range of cycles in which the logarithmic curve has the best linearity and therefore the maximum reaction efficiency. In the DART-PCR the midpoint is calculated based on the background signal (noise) of the fluorescence curve (indicated as the standard deviations of the fluorescence values) in the first 10 cycles. Although this range can be modified, it will be applied to all the curves giving a wrong result in cases in which the maximum fluorescence signal intensity assumes low values, the noise is particularly high, or the efficiency has relatively low values. However, also the LinReg PCR method produces a wrong individuation of the best range in the case of curves with short range of linearity.

Therefore, in this thesis a method for the analysis of the data obtained from the real time PCR was developed to improve the procedure for their elaboration. An algorithm was defined for the individuation of the range at which the PCR reaches the maximum efficiency on the logarithmic curve of each sample using the second derivative of the plot. As in the DART-PCR for each sample, the efficiency values of the primers used in the gene expression study are analysed through ANOVA in order to proof if a unique value per primer can be adopted for the further calculations. On the range of linearity instead of a specific threshold value for each sample like in the DART-PCR or a unique threshold value for all the samples another modification was introduced. For a specific primer pair a common threshold value was defined in the range of linearity for all the curves obtained from a specific sample. In this way, the variability of the data produced with a primer pair in the sample is reduced. The amount of fluorescence associated to the initial quantity of template present in the reaction is then calculated as in the DART-PCR method using the Eq. 8. The description of the method used in this work is reported in 3.2.20.4 while the modifications applied to the data elaboration procedure were performed as follows:

1) the raw data of the fluorescence values per cycle (F/cycle) obtained from the real-time PCR analysis are exported in an Excel® spreadsheet
2) the data are organised to have the sample replicates in two neighbouring columns and all the data obtained for the same gene (on different genotypes) disposed in adjacent columns
3) the logarithmic base 10 value ($\log_{10}$) of the F/cycle values and the first derivative of the $\log_{10}$ ($\log_{10}$') are calculated
4) the second derivative ($\log_{10}$'') curve is then calculated on the first derivative
5) starting from the end of the $\log_{10}$'' curve (Figure 23D) its tendency is analysed to find the points before and after whose the $\log_{10}$'' curve assumes value=0. The value (V0) on the $\log_{10}$'' curve is used to define a range of V0 +/- 2 values on the $\log_{10}$' curve (Figure 23C). These 5 values

are then analysed to find the group of consecutive points that give the minimum standard deviation (SD). All the 5 values are compared together or in groups of 4 or 3 values of consecutive data. The group of points giving the minimum SD defines a range that has a maximum value (Cmax) and a minimum value (Cmin) of cycle

6) the corresponding range of cycles (Cmin-Cmax) is reported on the original logarithmic curve (log_{10}) (Figure 23B) and the fluorescence values of Cmin and Cmax are annotated
7) the slope of the curve calculated on this range corresponds to the efficiency (E) of the primer pair associated with the curve
8) from the fluorescence value associated with Cmin and Cmax a mean fluorescence value (threshold) is calculated for this range
9) after the three runs for each gene 6 threshold values are obtained and a mean value is calculated (T)
10) the T value is then used to calculate for each sample curve at which cycle the fluorescence reaches this threshold. This cycle corresponds to the Ct value (Figure 23A)
11) the 6 E values obtained for each gene in a healthy or infected genotype are organised in a unique table in which all the efficiency data for the gene obtained from each sample are reported. An analysis of variance (ANOVA) is performed on the group of efficiencies in order to check if a unique value can be used for the analysed primer in all the different genotypes. In the case that the ANOVA analysis gives a negative results the specific efficiency is calculated for the group/s that was/were rejected
12) the E value, the T value and the Ct value are used to calculate for each sample the initial fluorescence intensity (I) using the equation: $I_{initial}=I_T*(E)^{-Ct}$
13) for each run the mean value of the $I_{initial}$ replicates is calculated obtaining 3 values from the three runs. The final mean value with the associated standard deviation can be calculated for each gene and sample
14) these values are then normalised on the internal control respectively using the equation: (mean $I_{initial}$ of gene) / (mean $I_{initial}$ of control)=gene normalised value
15) the normalised data are used to calculate the fold change of the gene with the equation: (gene normalised value infected state) / (gene normalised value healthy state) = gene fold change
16) in order to define if the fold change has a significant value the data group of the three $I_{initial}$ values deriving from the expression level in healthy state and the expression level in infected state are then compared with a F test in order to check if the variances are equal or different
17) in case of equal variances a t-test (alpha=0.05) is performed assuming the equal variance parameter otherwise the different variance parameter is assumed

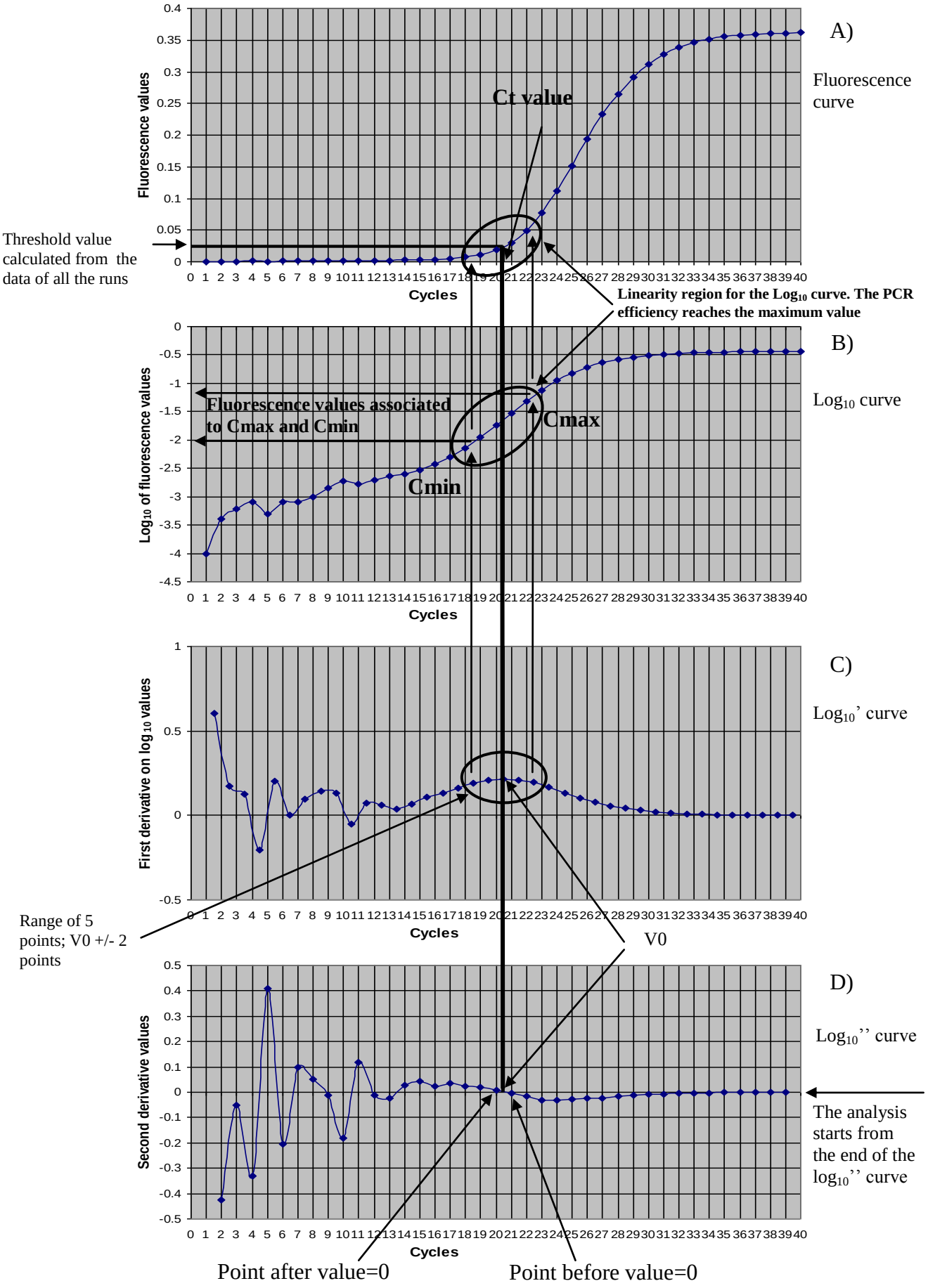


Figure 23 Data processing procedure for the threshold and Ct values identification

4.4.3 Housekeeping gene expression study for the development of an internal control for the real-time qPCR analysis

The data obtained from the gene expression analysis on the candidate genes have to be normalised on an internal control (commonly a housekeeping gene) that is expressed at a constant level in both the healthy and infected state. Unfortunately, there is a general lack of information in the phytoplasma-plant interaction about which gene could be used as suitable reference for the normalisation of the gene expression data. Furthermore, information about suitable housekeeping genes of *Malus* x *domestica* and *Malus sieboldii* was not available at the beginning of this work. Therefore, different commonly used housekeeping genes in other systems were tested:

- Eukaryotic translation initiation elongation factor 4A (eIF4)
- Tubulin alpha (Tua)
- Tubulin beta
- Actin
- Glyceraldeyde triphosphate dehydrogenase (GAPDH)
- Ubiquitin
- Elongation factor-1 α (EF)
- 18S rRNA (18S)

Only for some of the reported housekeeping genes it was possible to obtain information about their sequence in apple. Specific primers for one region of the GAPDH gene (Table 17, contig 7), Tua gene (GenBank Mdo.3499), 18S rRNA gene (GenBank DQ341382), eIF4 gene (GenBank CN936325) and EF (Table 17; contig 1) gene were designed (Table 26) based on sequences available at NCBI or on sequences directly obtained from cloned cDNA-AFLP fragments (4.1.3.2). In another case, primers for a region of an actine gene were also tested but they did not show good results (Dr. Komjanc, IASMA, personal communication). In addition, primers on an ubiquitine gene (Table 17; contig 4) yielded multiple PCR products probably due to the fact that this gene is represented by a family (Callis *et al.*, 1995) and were, thus, unsuitable. Therefore, only the first four genes mentioned above were used in the experiment (Table 26). The primers were tested through RT-PCR on RNA of different genotypes as described in 3.2.20.1 and through PCR on the cDNA obtained from the same RNA (3.2.20.2). A single band with the right size was obtained for GAPDH, Tua and eIF4 in RNA and cDNA of the different genotypes. For the 18S rRNA gene, one product was obtained only in the RNA but not in the cDNA. For the EF gene a product was obtained only in the cDNA but not in the RNA. The PCR products were then cloned as described in 3.2.9.2 and sequenced (3.2.15). The sequence data confirmed the specificity of the primers for the respective gene. The primers were then tested in a real-time qRT-PCR reaction at different RNA dilutions (3.2.20.1) in order to find the optimal reaction conditions. For all the genes, a 1:20 dilution of the same quantity of DNAse treated RNA was a good compromise while for the 18S rRNA gene analysis the best dilution was 1:200. (Figure 24). The housekeeping gene expression analysis was conducted on RNA extracted from three different tissues of Golden

delicious: *in vitro* plants, leaves from *ex vitro* plants and roots from *ex vitro* plants. For every tissue, material from both healthy and infected plants was separately collected and RNA was extracted as described (3.2.2). The qRT-PCR was performed in three replicas with two duplicates per replica. Six values per primer pair were obtained for every sample and processed through the data analysis procedure described above (3.2.20.4).

Table 26 Primers developed for the real-time qPCR analysis of housekeeping genes in the '*Ca*. P. mali'-plant interaction in *Malus*

Gene		Primer sequence	Predicted product size
Eukaryotic translation initiation factor 4A	Forward	ATTCAAGCGGAATGACAAGC	180
	Reverse	TCAGTGACAAAGCCAACAGC	
Tubulin alpha	Forward	CAACACCTTCTTCAGCGAGAC	175
	Reverse	AGCAGCATCTTCCTTTCCAG	
Elongation factor-1 α	Forward	CGCAGGCGAAATCTTCTATC	200
	Reverse	GAAGAACCCAGAGCACCTTG	
Glyceraldeide triphosphate dehydrogenase	Forward	GTTCGTCCACCTCTCCAGTC	190
	Reverse	GCACGTTGTTGGTGTGAACG	
18S rRNA	Forward	AAACGGCTACCACATCCAAG	249
	Reverse	ACCCAACCCAAGGTCCAACTA	

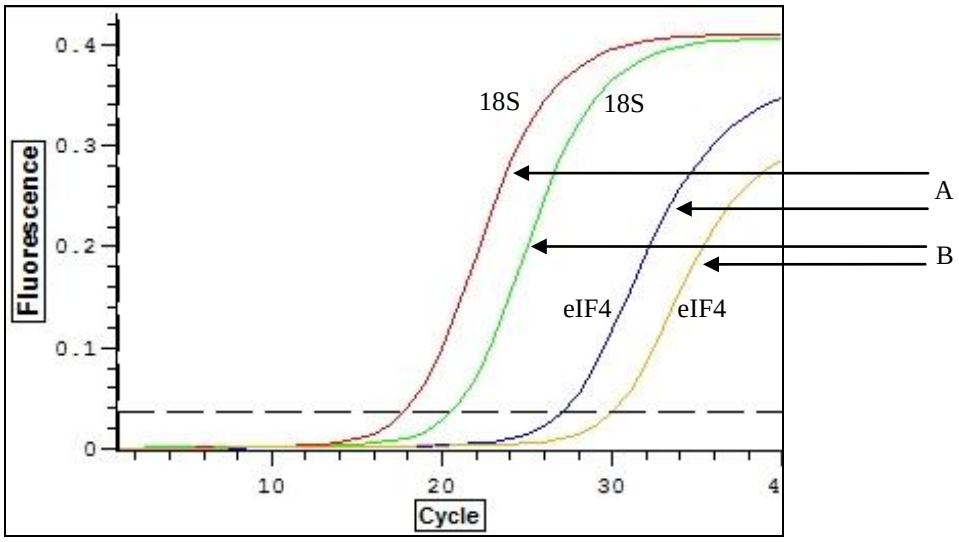


Figure 24 Comparison between the expression level of 18S and eIF4 in the same RNA (here *M. sieboldii*) sample in two different dilutions A= 1:20 B=1:200

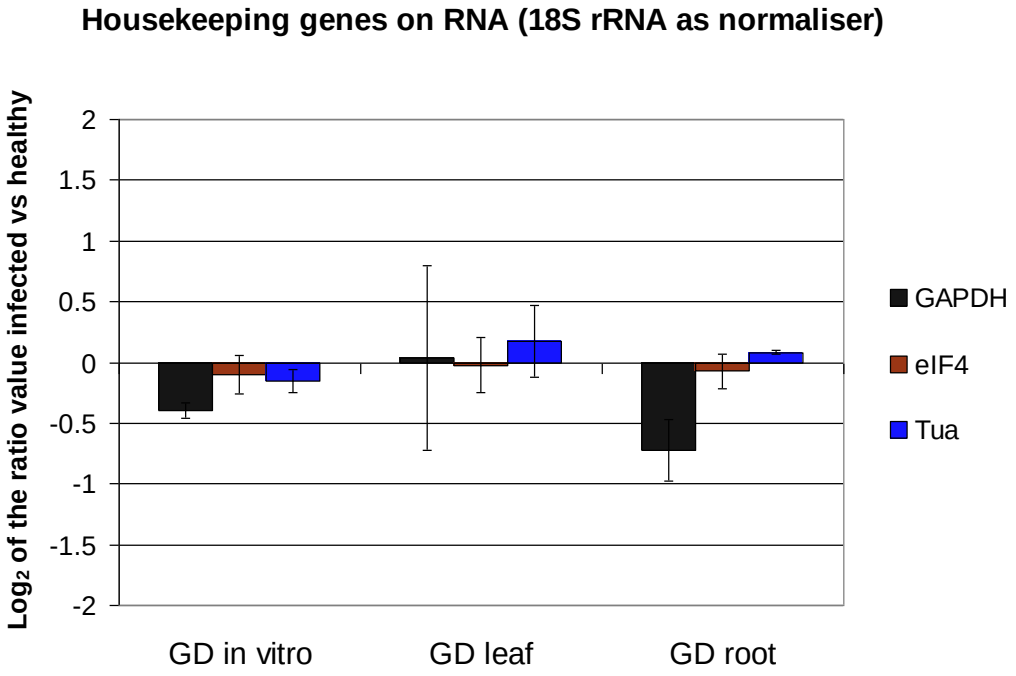


Figure 25 Expression levels of the housekeeping genes in RNA of *in vitro* shoots, *ex vitro* leaf and *ex vitro* root tissues of infected relative to healthy Golden delicious. The data were normalised on the expression level of the 18S rRNA.

The expression level of the genes GAPDH, Tua and eIF4 were normalised on the expression level of the gene 18S rRNA and then the relative expression levels between the infected and healthy state were calculated (Figure 25). The data obtained from the qRT-PCR analysis were statistically compared with a two tailed t-test (alpha=0.05). As shown in Figure 25 and in Table 27 the most stable gene in all the three tissues was the eIF4 gene showing no significant differences between its expression level in RNA obtained from healthy plants compared to the expression level in RNA from infected plants. For the GAPDH and Tua genes, no significant differences were found in their relative expression between healthy and infected plants in RNA of *in vitro* shoots and *ex vitro* leaf tissues. However, a significant difference between infected vs. healthy state was found in the expression level of both genes in RNA extracted from root tissue (Table 27).

Table 27 P values of the T-test performed on the expression levels between healthy and infected state for the genes GAPDH, eIF4 and Tua. In bold are evidenced the values significantly different.

P in a Two tailed T-test with alpha=0.05	*In vitro* plants	Leaves *ex vitro* plants	Roots *ex vitro* plants
GAPDH	1.06E-01	9.51E-01	**4.20E-02**
eIF4	5.28E-01	9.00E-01	4.44E-01
Tua	6.60E-02	3.86E-01	**2.35E-02**

The relative higher expression level of the 18S rRNA gene compared to the other genes could introduce an error due to possible differences in the reaction efficiency. In addition, using ribosomal RNA as a control could reduce the variation in expression underestimating the effective relation between the gene expression levels (Nicot *et al.*, 2005) whereas endogenous housekeeping genes would be more representative for such analyses. Therefore, the normalisation of the two genes GAPDH and Tua was repeated based on the expression levels of the gene eIF4 (Figure 26) as this gene showed the most stable expression in the first experiment with 18S rRNA as normaliser

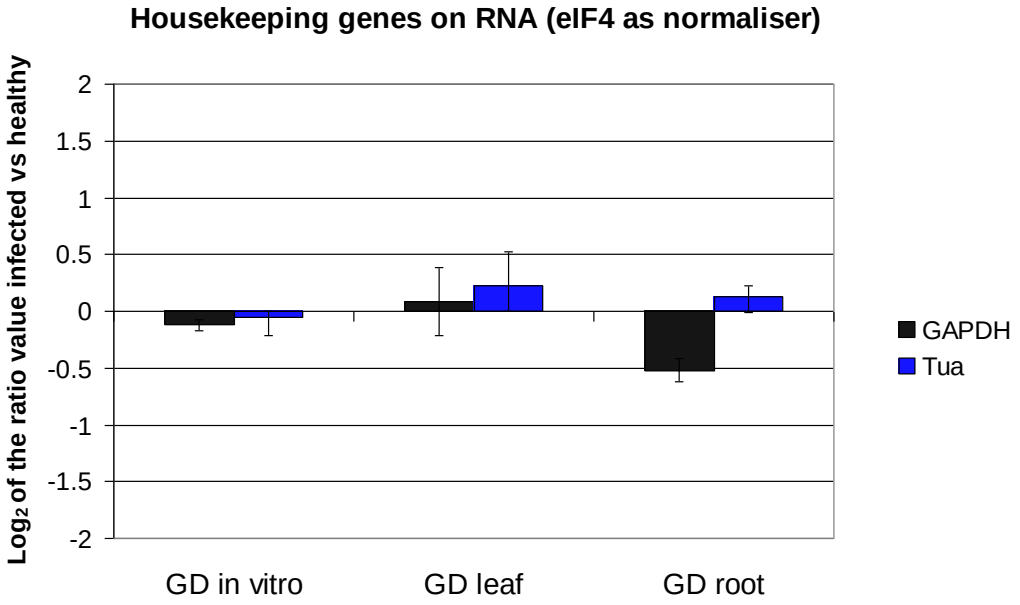


Figure 26 Expression level of GAPDH and Tua normalised on the eIF4 expression level and calculation of the relative expression between infected and healthy tissues.

Table 28 P values of the T-test performed on the expression levels between healthy and infected state for the genes GAPDH and Tua on RNA. In bold are reported the values significantly different

P in a Two tailed T-test with alpha=0.05	*In vitro* plants	Leaves *ex vitro* plants	Roots *ex vitro* plants
GAPDH	6.69E-02	7.35E-01	**1.53E-02**
Tua	7.08E-01	1.62E-01	2.66E-01

As shown in Figure 26 the respective levels are comparable to the results obtained from the analysis with the 18S rRNA gene. Nevertheless, in this case the differences in the expression level between infected and healthy state were significant only for GAPDH in *ex vitro* root tissue and not for Tua as it was for the data normalised on the 18S rRNA level.

In order to check if these results can be reproduced also using cDNA as template, the same analysis was performed on the cDNA synthesized from the same RNA used for the housekeeping genes analysis (Figure 27).

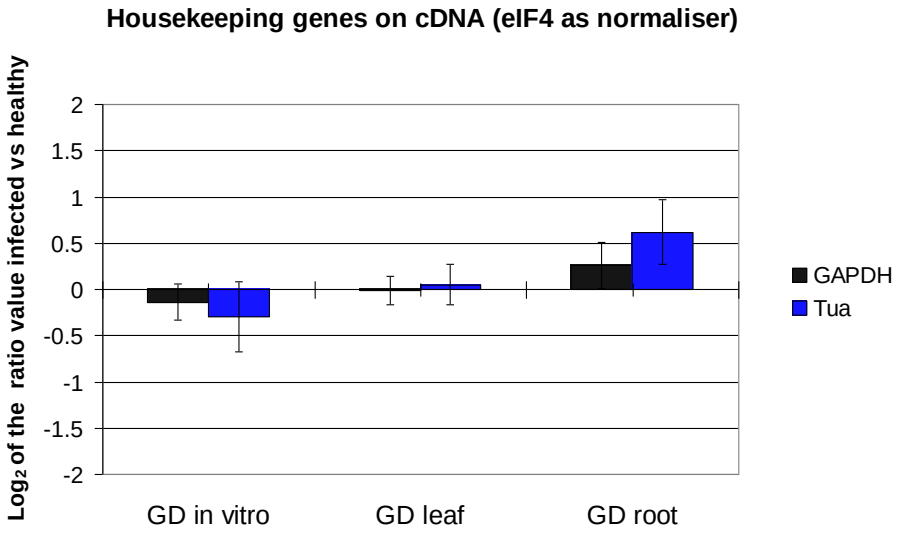


Figure 27 Relative expression level in cDNA synthetised from infected and healthy tissues. The expression level of eIF4 was used for the data normalisation

Table 29 P values of the T-test performed on the expression levels between healthy and infected state for the genes GAPDH and Tua on cDNA. In bold are evidenced the values significantly different.

P in a Two tailed T-test with alpha=0.05	*In vitro* plants	Leaves *ex vitro* plants	Roots *ex vitro* plants
GAPDH	5.21E-01	9.51E-01	1.95E-01
Tua	2.64E-01	7.34E-01	**4.20E-02**

As shown in Figure 27 there is a correspondence between the results obtained from the RNA of *in vitro* shoots and *ex vitro* leaf material. Differently, for the root material the expression levels were not congruent with those obtained from the analysis performed on RNA.

The primer pair developed on the sequence corresponding to the elongation factor (EF) was also tested for its suitability in qRT-PCR. Unfortunately, it was not possible to obtain a clear product directly from RNA because of the presence of primer dimers or unspecific products that hindered the production of reliable data. Nevertheless, using cDNA as template a unique product was obtained. Therefore, a comparison of EF expression level was done in different pools of GD healthy *in vitro* plants vs. GD '*Ca*. P. mali'-infected *in vitro* plants using the eIF4 gene for the data normalisation (Figure 28).

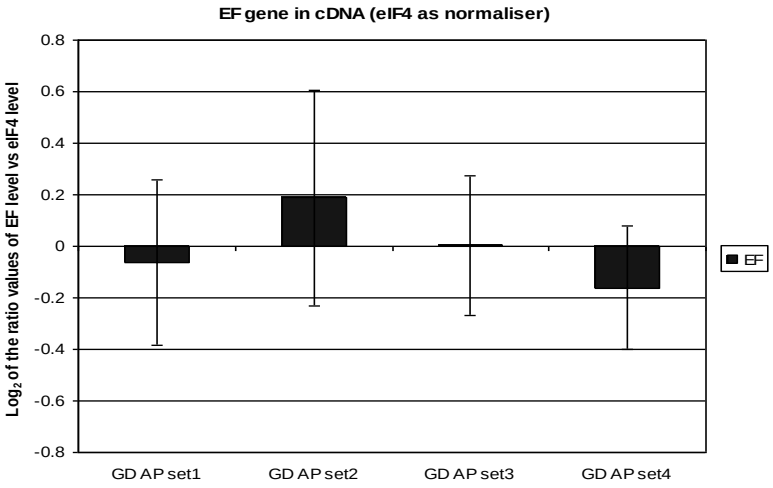


Figure 28 EF expression levels compared in cDNA synthetised from different pools of infected and healthy GD *in vitro* plants.

Table 30 P values of the t-test performed on the expression levels between healthy and infected state for the EF gene on cDNA.

P in a Two tailed t-test with alpha=0.05	GD AP set1	GD AP set2	GD AP set3	GD AP set4
EF	7.50E-01	1.73E-01	9.82E-01	7.88E-02

As shown in Figure 28 the expression levels of the EF gene were comparable between healthy and infected state and no significant differences were found after the t-test performed on the infected/healthy datasets (Table 30).

4.4.4 Real-time analysis on cDNA of *in vitro* plants

The real-time analysis was carried out on the cDNA synthesised from the infected and healthy genotypes as described in 3.2.20.2. For every gene, three repetitions were performed in which the reactions were done in double. The housekeeping genes eIF4 and EF were used as internal control for the data normalisation of the gene expression level. The ratio between infected and healthy status was then calculated and the $\log_2$ value reported for each gene. In addition, the normalised values were directly compared in order to show differences in the expression level between the genotypes. In this case GD healthy was taken as reference and the ratios between normalised expression level of the gene in a genotype and that of GD healthy was calculated.
Not all the primers showed a good performance in the real-time PCR analysis. For the primer 23 the results obtained were too variable and therefore they were not analysed. The primer 26 did not show a good kinetic of the reaction giving primer dimers that were evidenced in the melting curve analysis (Figure 20). A further problem was represented by the quantification of the eIF4 gene in GD AT2-infected *in vitro* plants where the variability between the replicas did not allow the

generation of reliable data. Therefore, these results were not included in the graphs reported for each gene. Nevertheless, with the other primers reliable results were obtained.

Gene 3

As shown in the Figure 27A and 27B for the different genotypes the relative gene expression between infected and healthy state calculated on the two housekeeping genes EF and eIF4 were comparable. The values obtained for the eIF4 expression level in GD AT2 were at the limit of sensitivity of the detection method. Therefore, the normalisation on eIF4 was not possible. On the contrary, the EF expression levels allowed the analysis and the comparison of the relative expression between GD AT2 infected plants and the healthy GD. In all the genotypes, no significant differential expression was seen after the infection (t-test Table 31). However, the comparison of the expression level of gene 3 (normalised on the EF gene) between the different genotypes and states (Figure 27C) showed genotype-specific differences. M9 showed a lower expression level in general than the other genotypes.

Gene 5

For the gene 5, significant differences were found in *M. sieboldii*, H0909 and GD AP (Table 32; values in bold) after the infection. Moreover, also in AT2-infected GD the gene resulted to be significantly differentially expressed (EF data) but the confirmation of this results by the eIF4 data is missing. No significant difference was seen in D2212 and M9 after the infection.
In the comparison of the general expression level *M. sieboldii* showed a higher expression of the gene compared to the other genotypes. Interestingly, the levels in both healthy H0909 and *M. sieboldii* were similar but after the infection in *M. sieboldii* the level increased while in H0909 it decreased.

Gene 6

Gene 6 was differentially expressed in M9, *M. sieboldii* and H0909. Nevertheless, only in M9 and *M. sieboldii* this difference was significant in both the normalised data eIF4 and EF while for H0909 this was the case only in the eIF4 normalised dataset (Table 33). The general expression levels of gene 6 were very similar in all the genotypes except for M9 in which it was found to be between six and eight times lower.

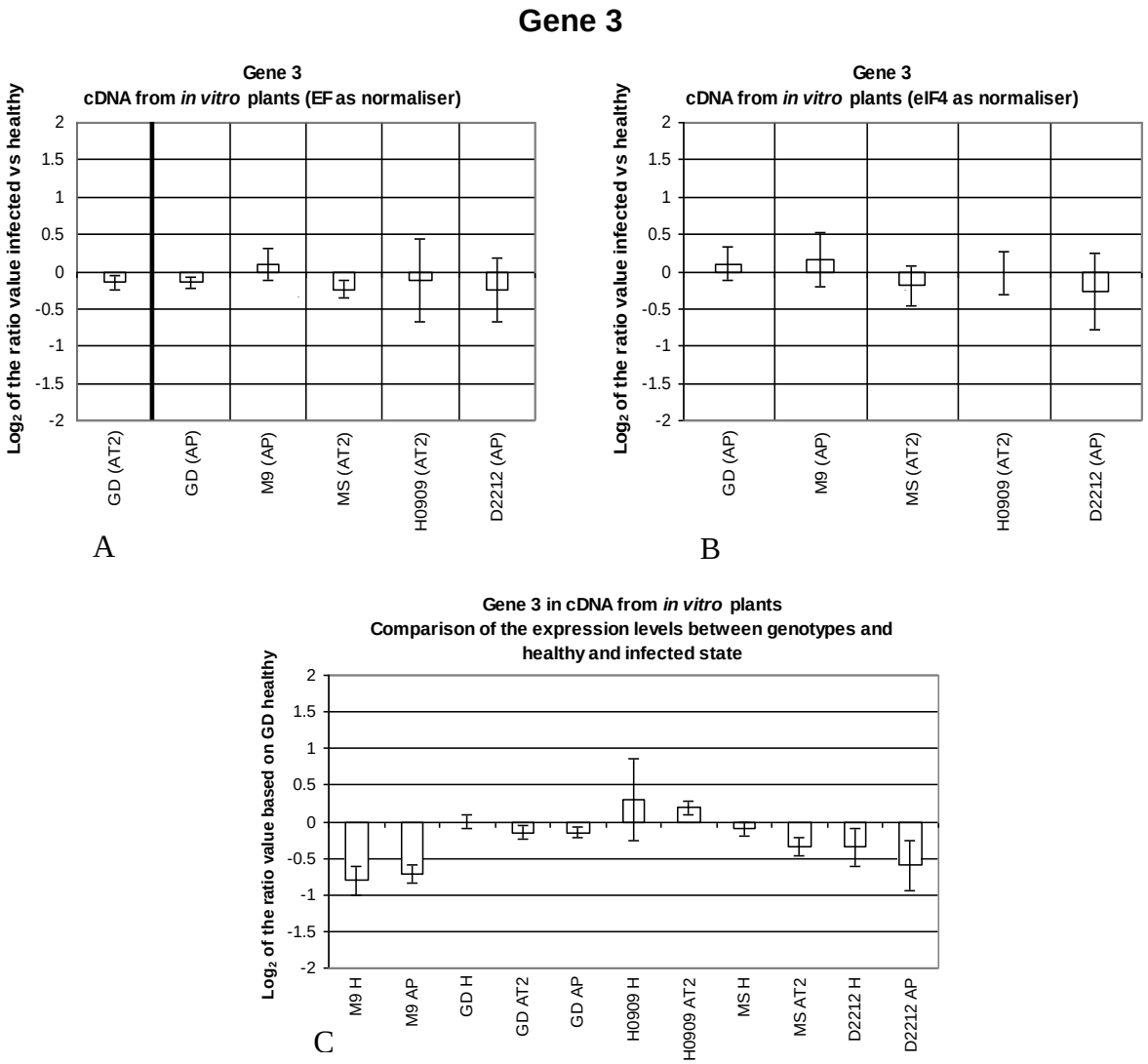


Figure 29 Gene 3 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF)

Table 31 P values for the two-tailed t-test (α=0.05) on the Gene 3 expression levels between infected and healthy state

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
1.44E-01	1.02E-01	4.85E-01	9.80E-02	7.66E-01	4.32E-01	EF
	5.81E-01	5.57E-01	4.33E-01	9.11E-01	4.43E-01	eIF4

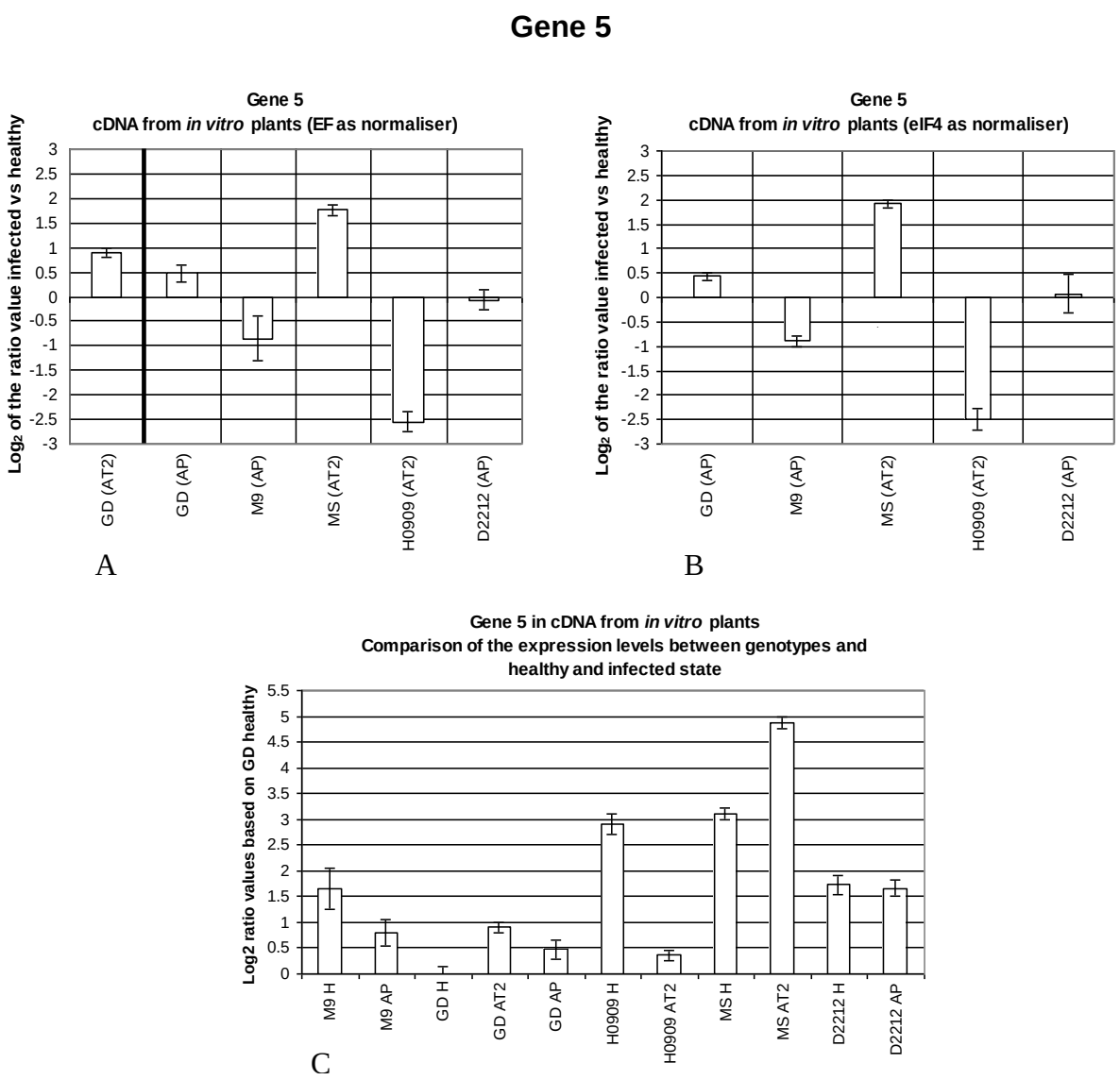


Figure 30 **Gene 5 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF). Blue and green coloured bars represent significant differences in the expression level between infected and healthy state.**

Table 32 **P values for the two-tailed t-test (α=0.05) on the Gene 5 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.**

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
5.85E-04	**2.27E-02**	1.55E-01	**2.57E-03**	**1.16E-02**	6.23E-01	EF
	1.21E-02	7.30E-02	**3.48E-04**	**7.44E-03**	7.50E-01	eIF4

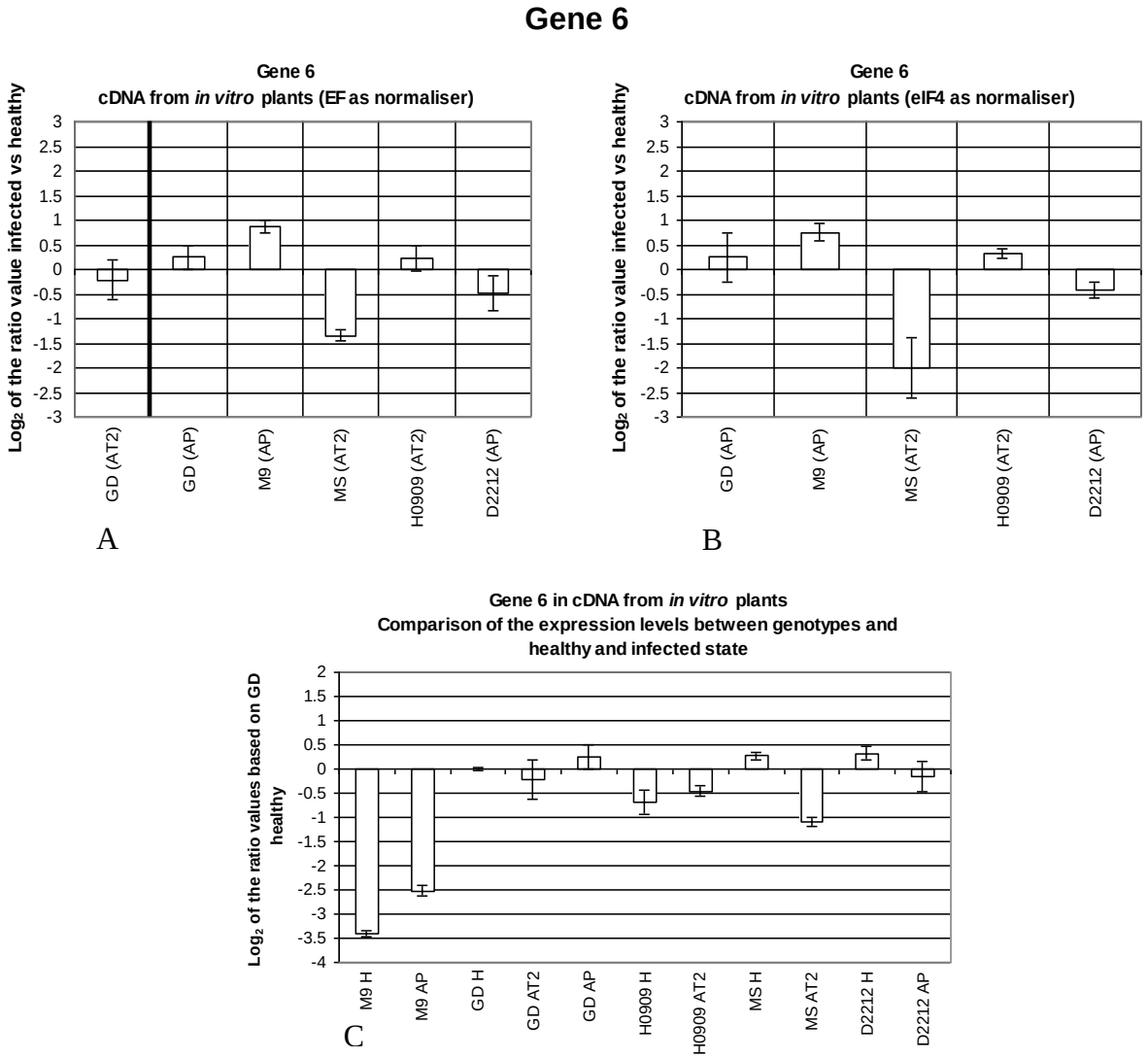


Figure 31 Gene 6 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF). Blue coloured bars represent significant differences in the expression level between infected and healthy state.

Table 33 P values for the two-tailed t-test (α=0.05) on the Gene 6 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
4.87E-01	3.27E-01	**4.49E-03**	**4.14E-03**	3.30E-01	1.60E-01	EF
	5.42E-01	**4.67E-02**	**8.45E-02**	**3.94E-02**	6.33E-02	eIF4

Gene 7

No significant differences were observed for the gene 7 except for the GD AP eIF4 normalised data. The general expression level was comparable in H0909, *M. sieboldii* and M9 although the latter was characterised by high standard deviation values. In GD and D2212 the expression level was lower.

Gene 9

For gene 9 a significant difference in the relative expression after the infection in both the EF and eIF4 normalised data was detected in H0909. In the other genotypes no significant differences were found. Considering the high standard deviation the general expression levels were comparable in all the genotypes though in the apomictic ones the level was tendentially lower (except in H0909 AT2 infected).

Gene 11

The gene 11 was significantly differently expressed only in *M. sieboldii*. The comparison of the general expression levels showed no differences between the genotypes. However, for this gene the high standard deviation did not allow to obtain reliable results.

Gene 14

The expression level of gene 14 showed for GD AP infected an up-regulation after the infection. The high standard deviation in the EF normalised values did not allow any consideration about the relative expression for the other genotypes although in *M. sieboldii* AT2 the difference calculated with eIF4 resulted to be significant. The general expression level showed a tendency with higher expression values in the infected state for the *M*. x *domestica* plants while this trend was opposite for *M, sieboldii* and *M. sieboldii*-derived genotypes. However, the high standard deviation values did not allow further considerations.

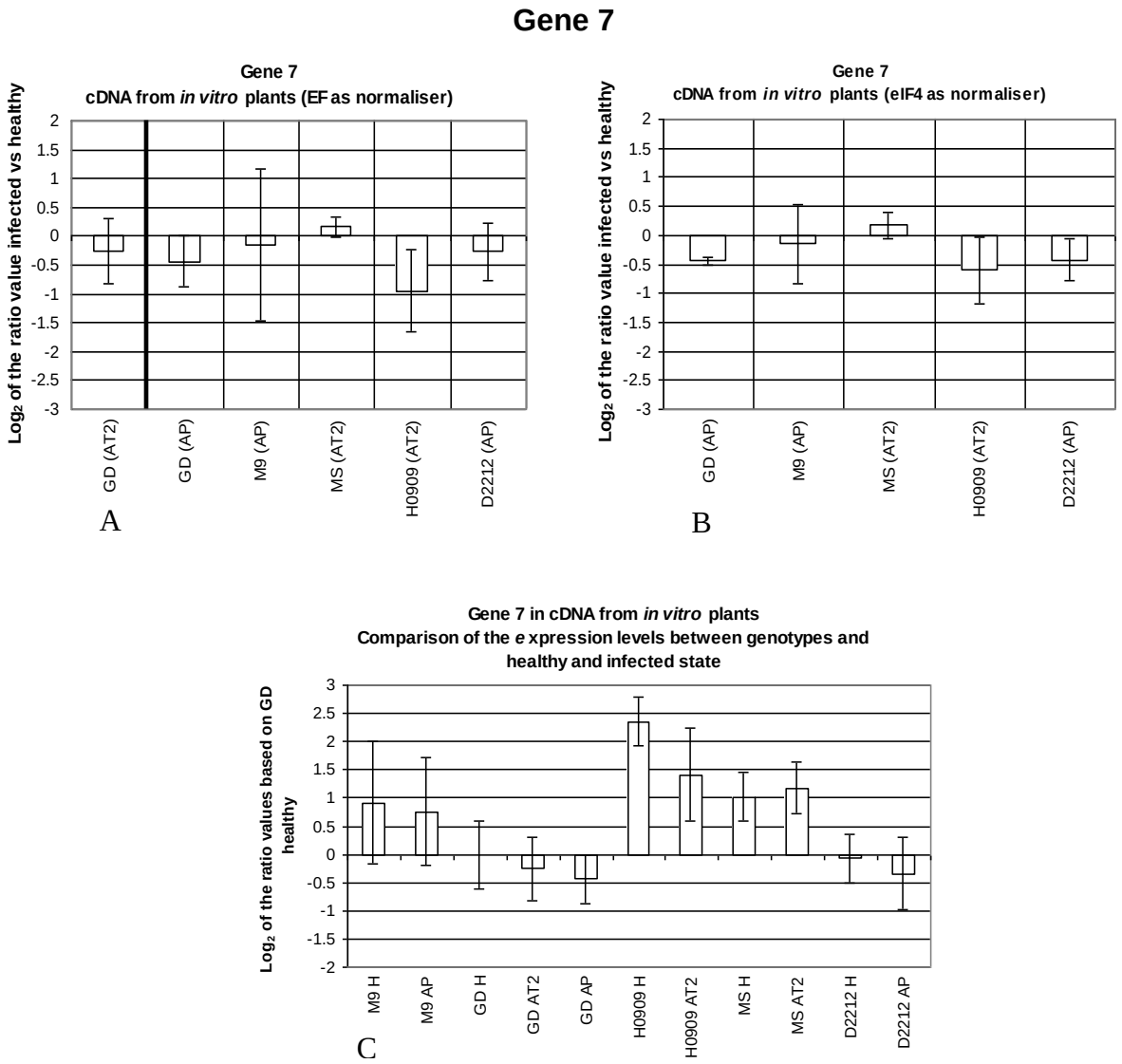


Figure 32 Gene 7 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF).

Table 34 P values for the two-tailed t-test (α=0.05) on the Gene 7 expression levels between infected and healthy state.

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
5.99E-01	3.40E-01	8.84E-01	3.17E-01	1.20E-01	4.79E-01	EF
	8.08E-03	7.80E-01	4.11E-01	2.21E-01	2.49E-01	eIF4

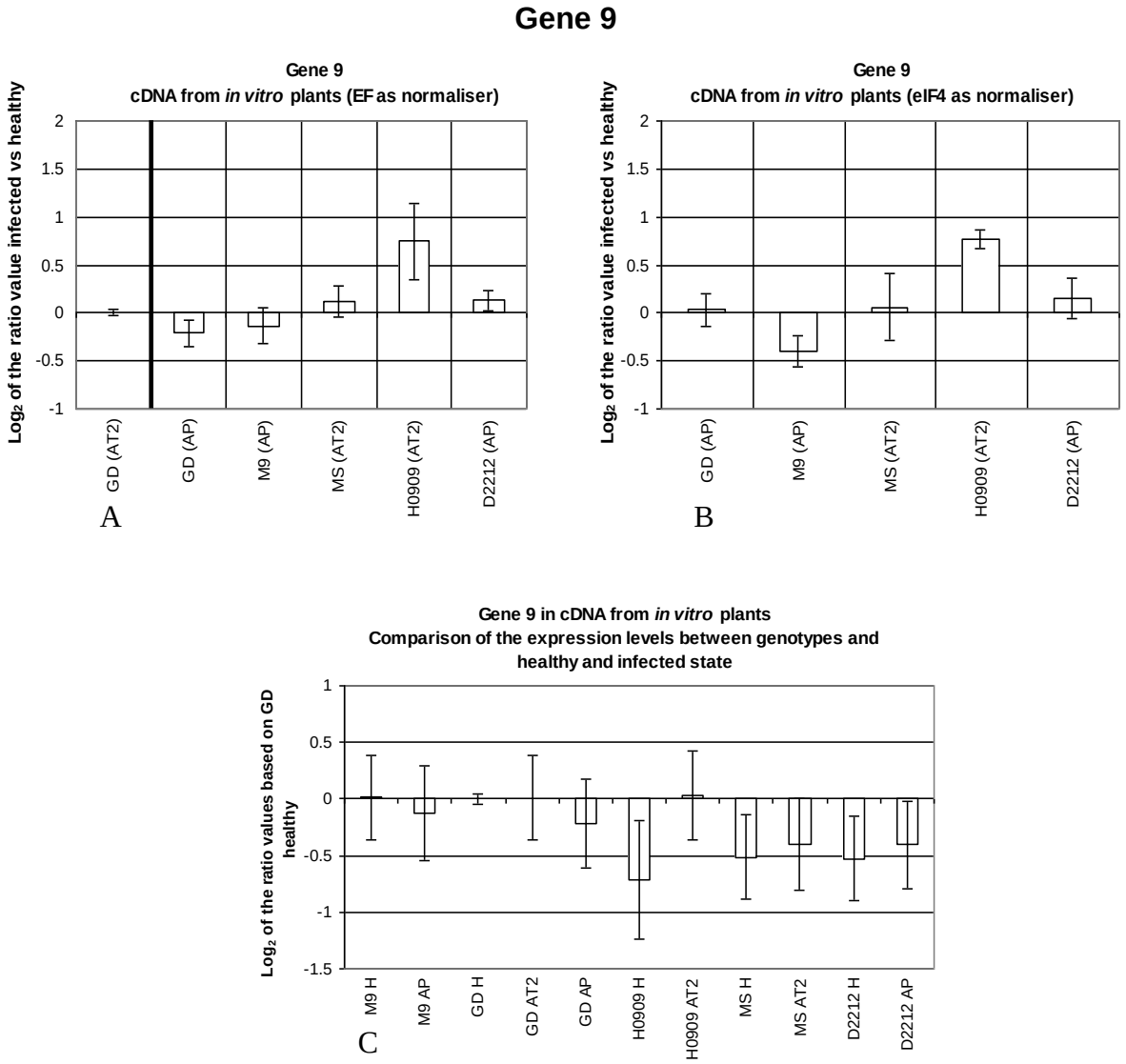


Figure 33 Gene 9 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF). Blue coloured bars represent significant differences in the expression level between infected and healthy state.

Table 35 P values for the two-tailed t-test (α=0.05) on the Gene 9 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
6.93E-01	1.38E-01	4.79E-01	4.44E-01	**3.13E-02**	2.45E-01	EF
	8.01E-01	9.23E-02	8.33E-01	**7.20E-04**	4.34E-01	eIF4

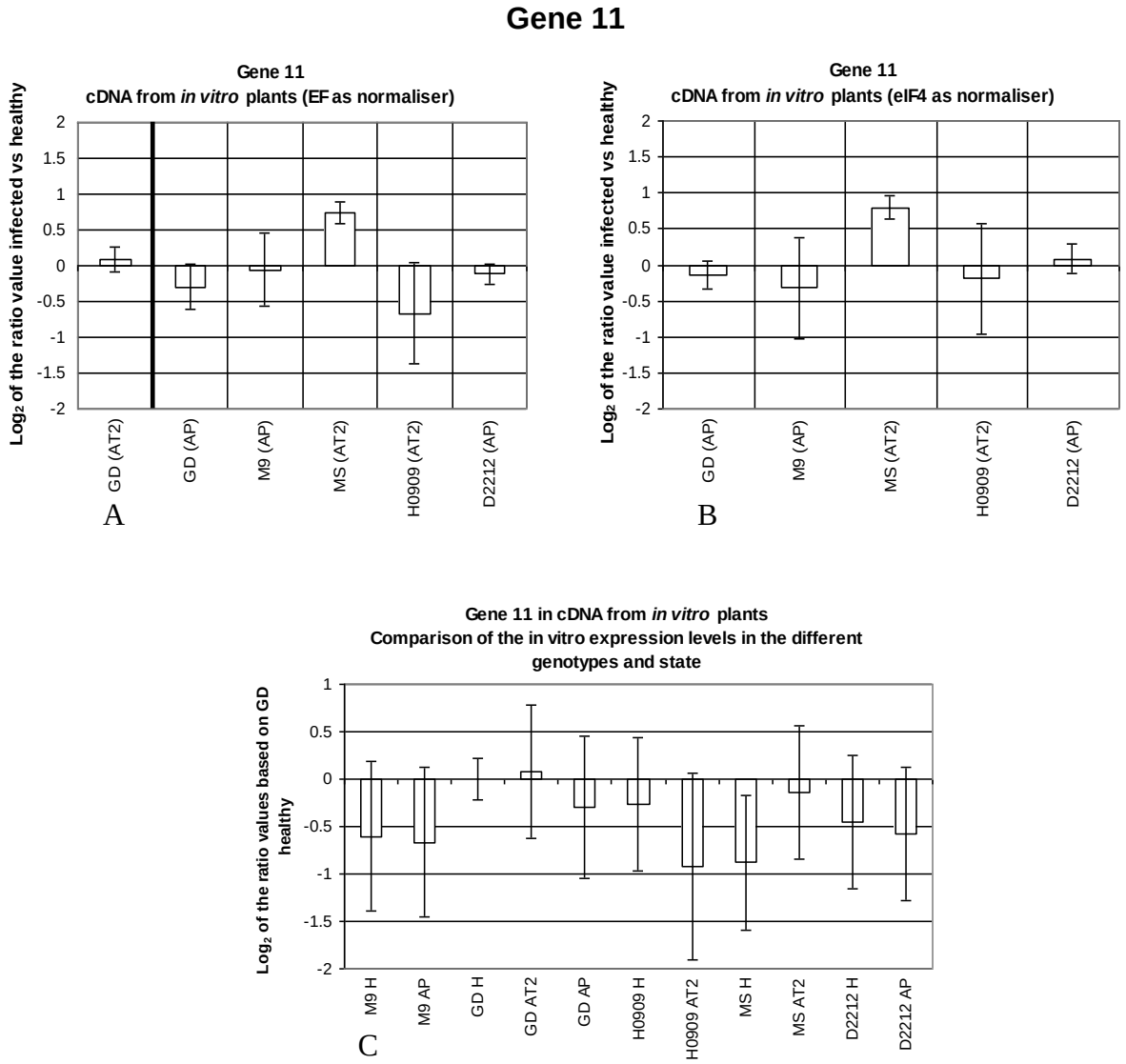


Figure 34 Gene 11 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF). Blue coloured bars represent significant differences in the expression level between infected and healthy state.

Table 36 P values for the two-tailed t-test (α=0.05) on the Gene 11 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
5.67E-01	2.91E-01	8.83E-01	**2.27E-02**	2.33E-01	2.34E-01	EF
	4.18E-01	5.96E-01	**2.97E-03**	7.40E-01	6.36E-01	eIF4

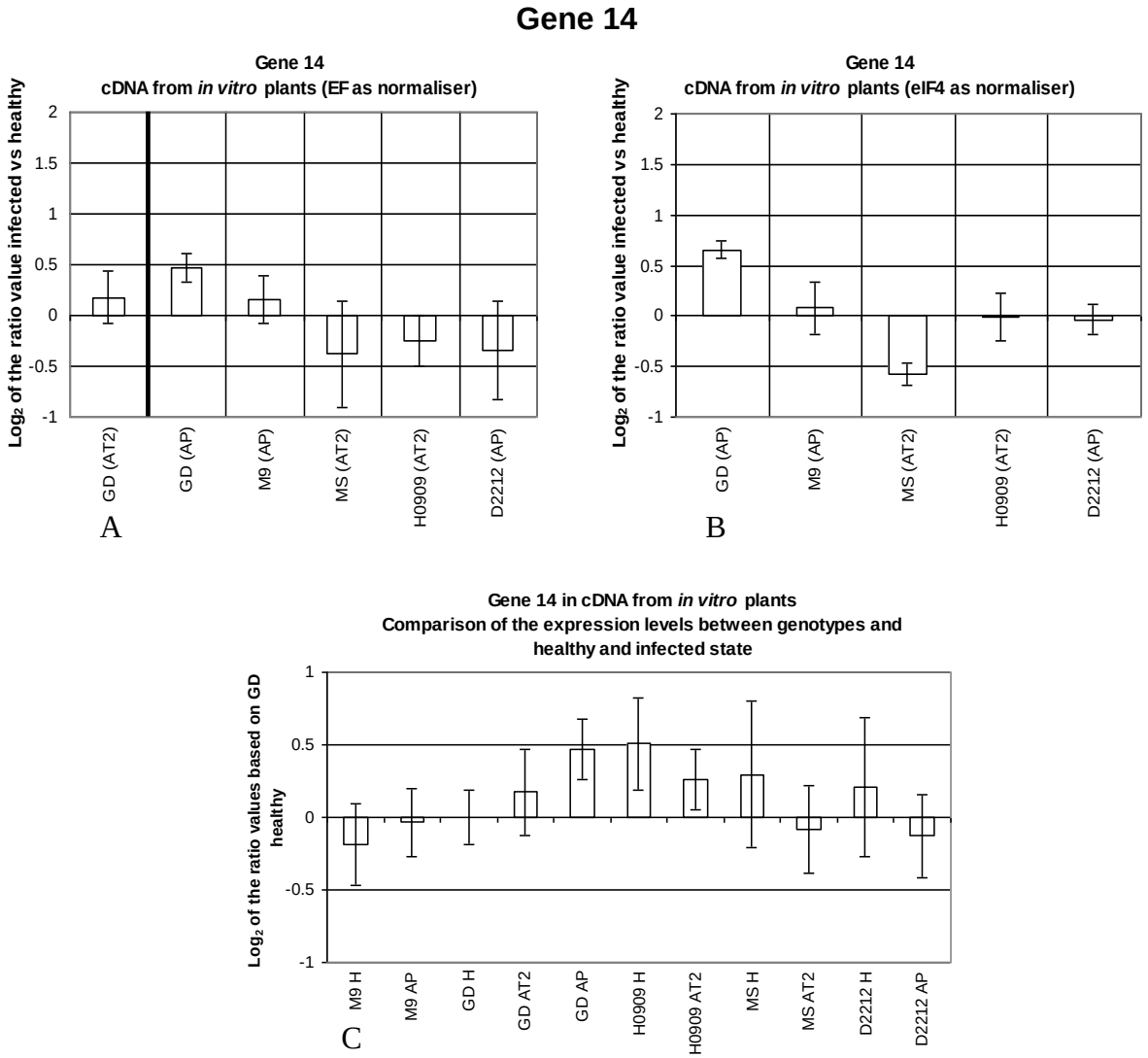


Figure 35 Gene 14 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF). Blue coloured bars represent significant differences in the expression level between infected and healthy state.

Table 37 P values for the two-tailed t-test (α=0.05) on the Gene 14 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
3.13E-01	**1.36E-02**	3.20E-01	3.09E-01	3.29E-01	3.27E-01	EF
	1.22E-02	7.18E-01	**2.05E-02**	9.24E-01	7.77E-01	eIF4

Gene 15

The relative expression of gene 15 was significantly different in D2212 in both the normalised datasets. The gene resulted to be down-regulated after the infection. The same was observed for H0909 but in this case the difference between infected and healthy state was significant only with the data normalised on eIF4. The expression level comparison showed similar level in M9 and *M. sieboldii* healthy and infected, in GD infected (both AT2 and AP), H0909 and D2212 healthy. The level decreased in the case of H0909 and D2212 infected while it was lower in GD healthy.

Gene 21

The gene 21 resulted to be up regulated in H0909 with a statistically significant difference between the expression level in infected and healthy state in both EF and eIF4 normalised data. Significant differences were observed also for *M. sieboldii* and D2212 but only in the eIF4-normalised dataset. In the comparison of the general expression level it is interesting to observe that while the value for H0909 healthy is comparable to the other apomictics, after infection the level increased and became higher than the reference GD healthy.

Gene 24

The relative expression level of gene 24 between infected and healthy state showed a down regulation in H0909 in both the data set normalised on EF and eIF4 but only the latter difference was significant. The expression in GD AP infected and D2212 resulted significantly decreased after the EF normalisation but not in the normalisation on eIF4. About the general expression level it can be observed that it was tendentially higher in GD healthy and GD AT2 (remaining almost unchanged) compared to the other genotypes. In M9, GD AP, *M. sieboldii* in both healthy and infected, in D2212 healthy and in H0909 healthy the level was comparable while in H0909 AT2 and D2212 AP the expression was tendentially lower (indication of a down-regulation). However, these considerations were supported from the statistical analysis either in the EF- or in the eIF4-normalised dataset but not in both.

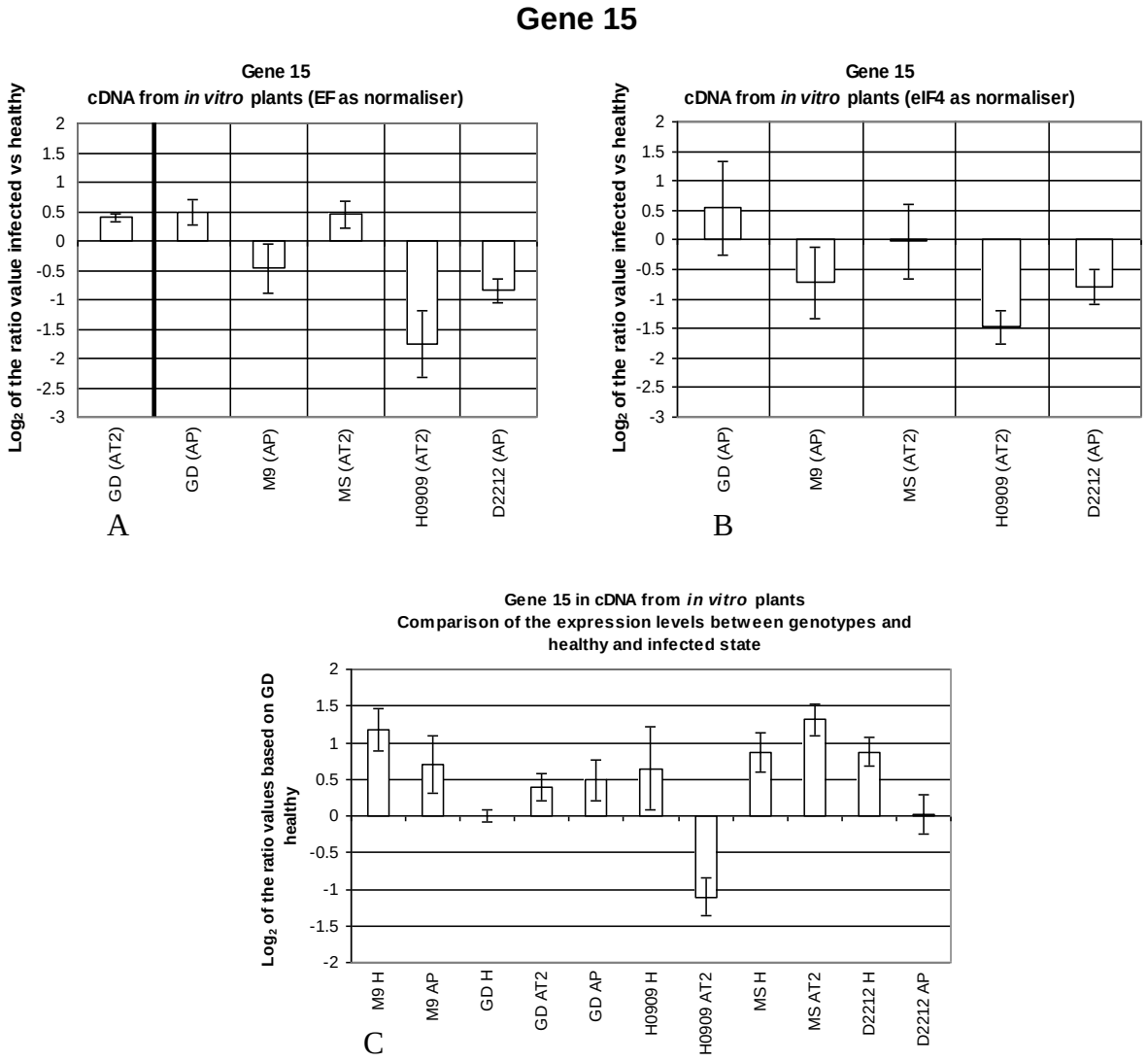


Figure 36 Gene 15 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF). Blue and green coloured bars represent significant differences in the expression level between infected and healthy state.

Table 38 P values for the two-tailed t-test (α=0.05) on the Gene 15 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
1.15E-02	1.17E-01	2.36E-01	5.31E-02	1.18E-01	**1.02E-03**	EF
	2.89E-01	2.37E-01	9.47E-01	**1.60E-02**	**4.10E-02**	eIF4

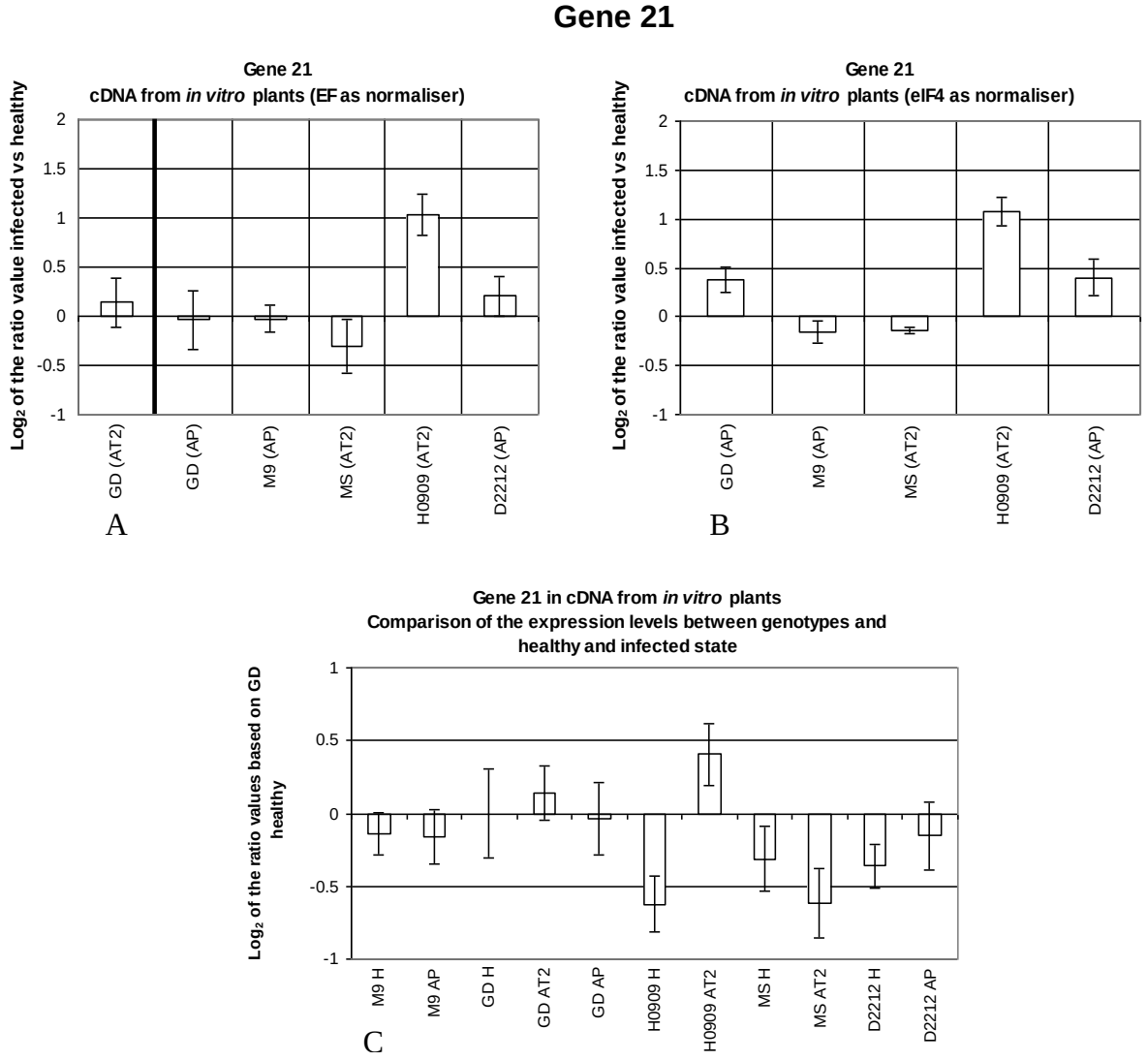


Figure 37 Gene 21 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF). Blue coloured bars represent significant differences in the expression level between infected and healthy state.

Table 39 P values for the two-tailed t-test (α=0.05) on the Gene 21 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
3.73E-01	8.65E-01	8.22E-01	2.55E-01	**2.83E-02**	3.11E-01	EF
	1.27E-01	1.72E-01	**2.15E-02**	**9.56E-05**	**3.06E-02**	eIF4

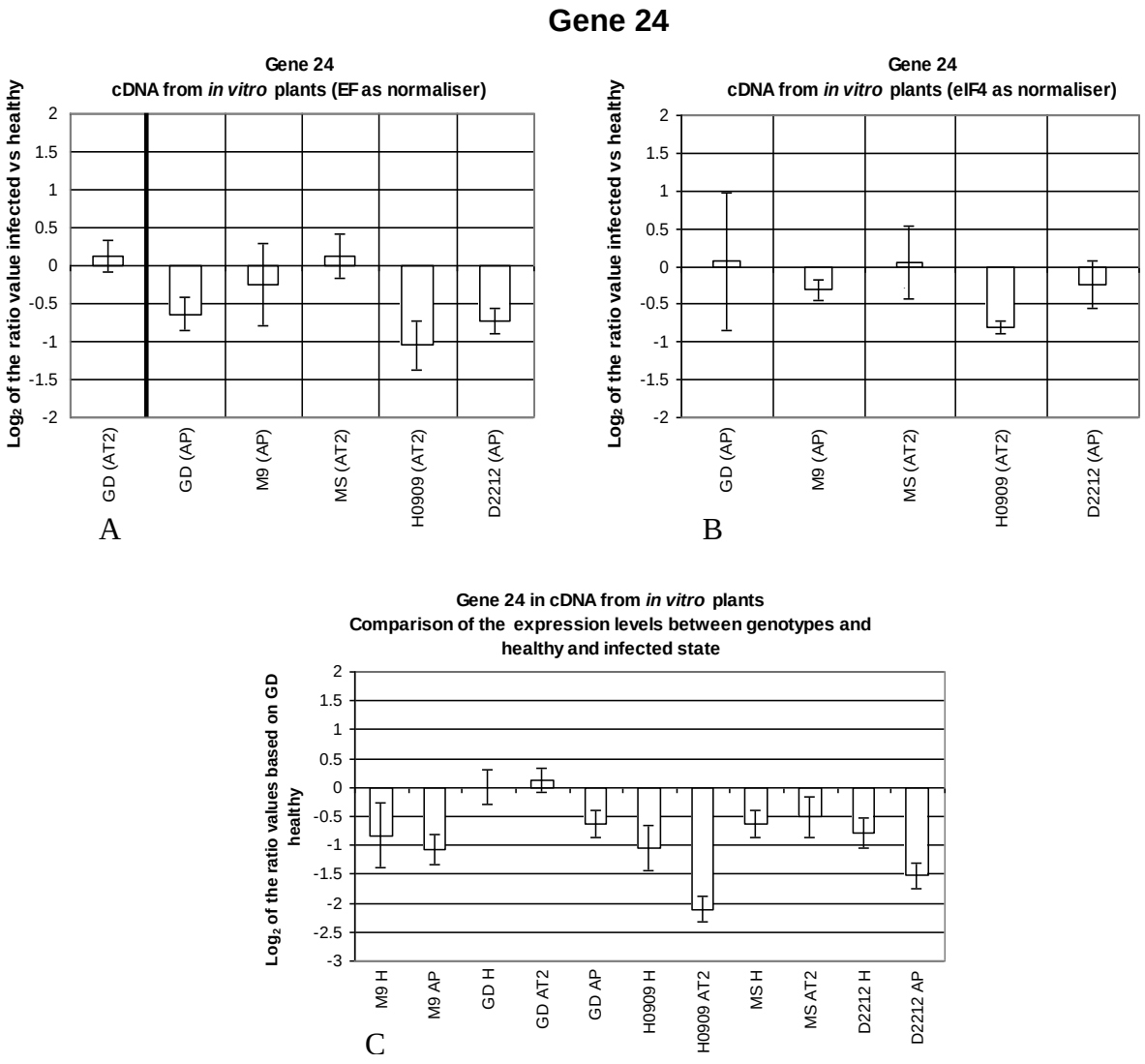


Figure 38 Gene 24 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF).

Table 40 P values for the two-tailed t-test (α=0.05) on the Gene 24 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
3.48E-01	**1.48E-02**	5.04E-01	5.17E-01	8.46E-02	**1.58E-02**	EF
	9.10E-01	**1.28E-02**	8.83E-01	**8.33E-03**	2.15E-01	eIF4

Gene 25

The values obtained from both the EF and eIF4 normalised data sets for gene 25 resulted to be significantly differentially expressed in *M.* and H0909 after the infection. Nevertheless, the relative expression in *M. sieboldii* showed up regulation of the gene whereas in H0909 the gene resulted to be down regulated. Moreover, the comparison of the expression level in the different genotypes showed in M9 a general lower value comparable to the value in H0909 AT2 infected. The levels in H0909 healthy, *M. sieboldii* healthy, D2212 healthy and infected showed similar ranges while GD had the highest level in both states healthy and infected (remaining unchanged) but similar to that observed in *M. sieboldii* AT2 infected.

Gene 27

For the gene 27 significant differences were found only in the eIF4 normalised data of M9 and in the EF normalised data of D2212. However, in almost all the genotypes the ratio values were very close to 1 indicating no differential expression after the infection. In the comparison of the general expression levels in GD the gene 27 resulted to be relatively highly expressed in contrast to what was observed for the other genotypes.

Gene 28

The relative expression level of the gene 28 after the infection resulted to be significantly different for both the EF- and eIF4-normalised data in GD AP-infected, *M. sieboldii* and D2212. In H0909, the difference was significant only in the EF-normalised dataset and so was the expression in GD AT2 though in this case the eIF4 data are missing. In GD, *M. sieboldii* and D2212 the gene 28 showed over expression in the infected state supported also by the statistical analysis. In contrast, although the t-test P value was significant only for the EF-normalised data in H0909, the gene showed under expression. From the comparison between the level in the different genotypes and states the gene 28 showed similar values in M9 (healthy and infected) and GD infected (with both AP and AT2). Interestingly, in the apomictic genotypes the level was lower (especially for H0909 infected and MS healthy) or similar to that in GD healthy.

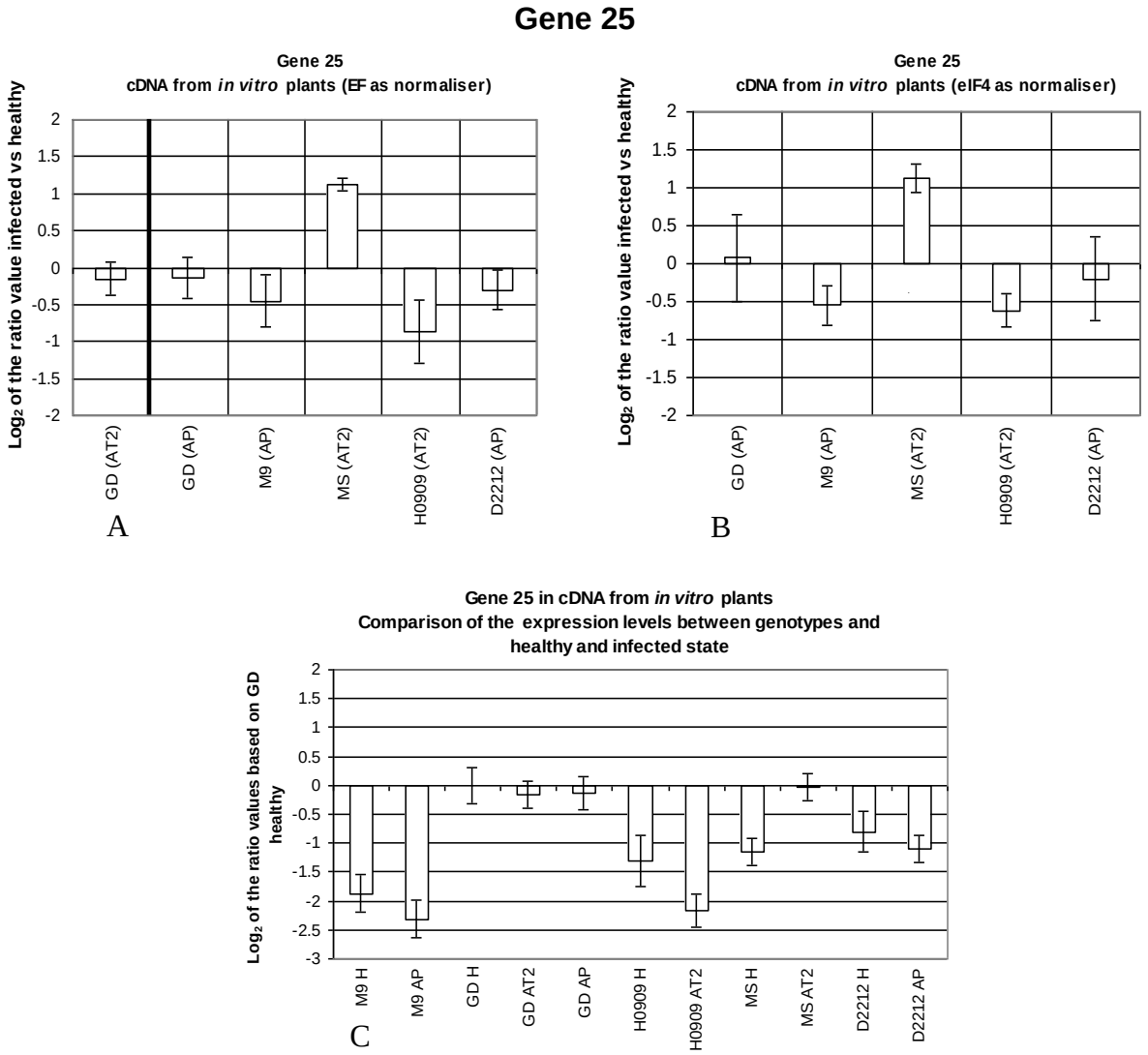


Figure 39 Gene 25 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF). Blue coloured bars represent significant differences in the expression level between infected and healthy state.

Table 41 P values for the two-tailed t-test (α=0.05) on the Gene 25 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
4.71E-01	5.65E-01	1.26E-01	**4.05E-03**	**4.83E-02**	2.77E-01	EF
	8.39E-01	9.32E-02	**7.59E-03**	**1.30E-02**	5.38E-01	eIF4

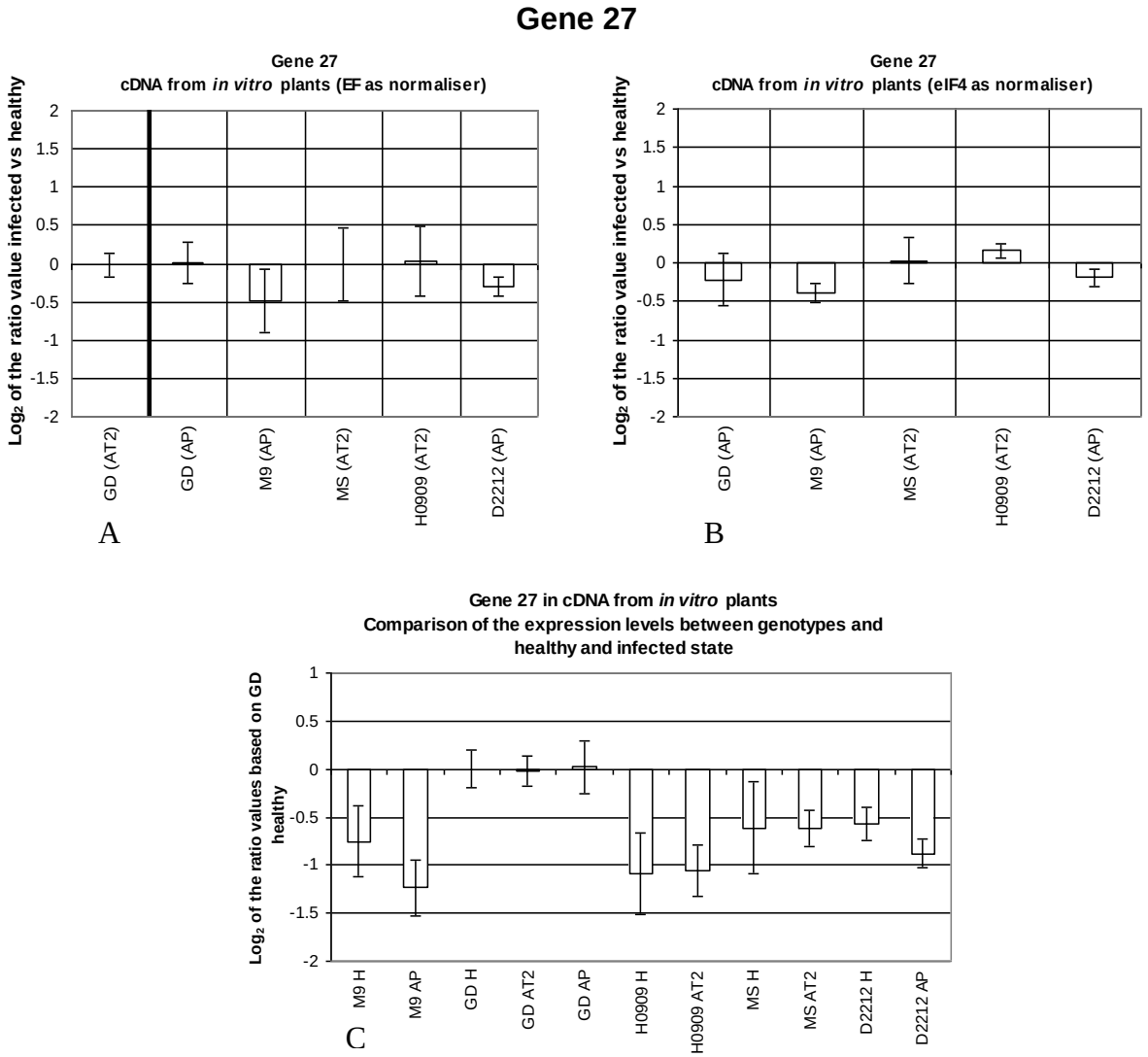


Figure 40 **Gene 27 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF).**

Table 42 **P values for the two-tailed t-test (α=0.05) on the Gene 27 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.**

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
8.53E-01	9.08E-01	1.60E-01	9.80E-01	9.20E-01	**1.36E-02**	EF
	3.38E-01	**7.48E-03**	8.94E-01	1.37E-01	6.31E-02	eIF4

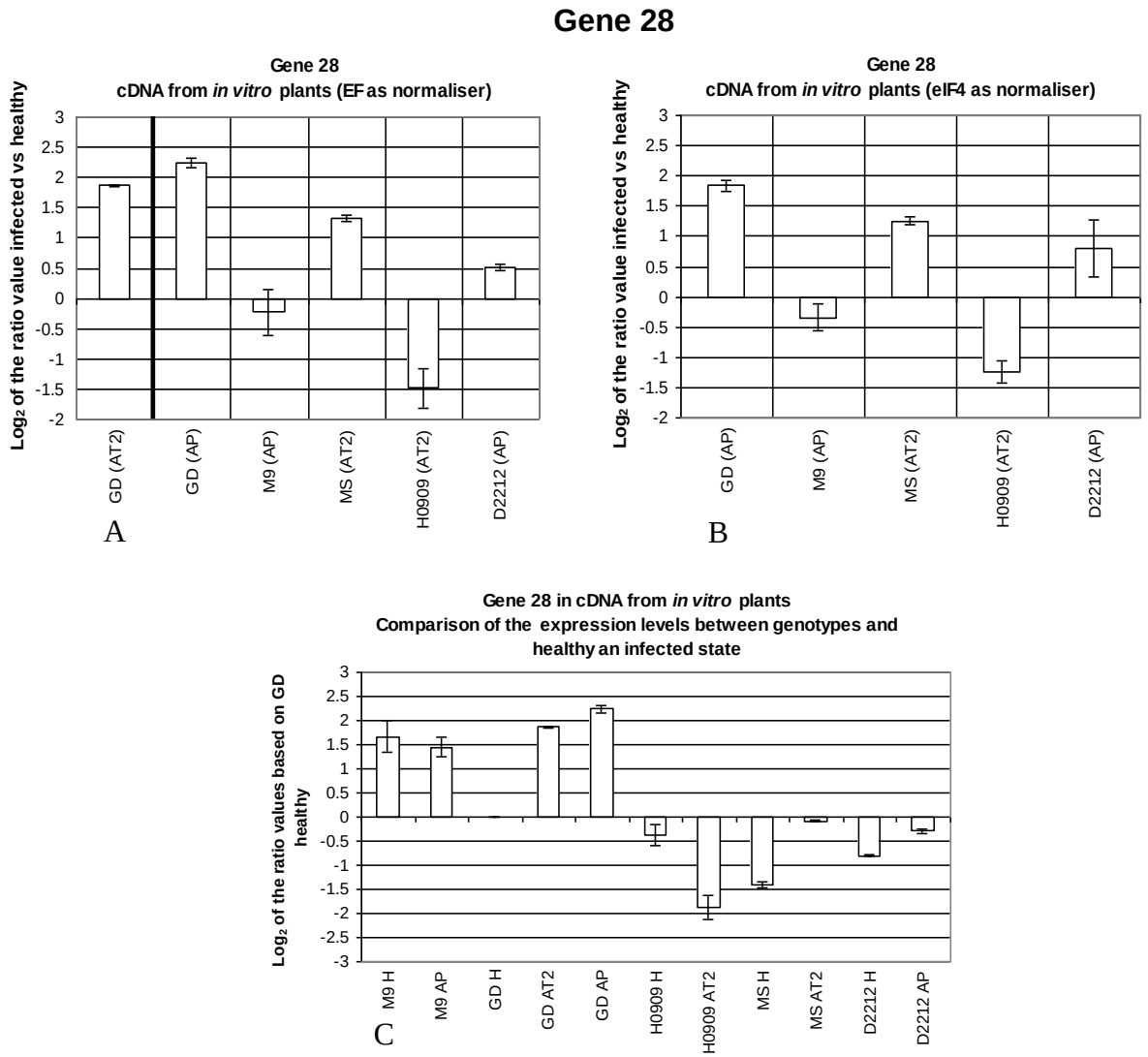


Figure 41 Gene 28 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF). Blue and green coloured bars represent significant differences in the expression level between infected and healthy state.

Table 43 P values for the two-tailed t-test (α=0.05) on the Gene 28 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
8.00E-06	**2.78E-02**	3.82E-01	**6.62E-04**	**5.87E-03**	**4.48E-03**	EF
	2.65E-03	1.77E-01	**5.60E-04**	1.03E-01	**5.97E-02**	eIF4

Gene 29

The gene 29 did not show significant differences in expression except in the EF normalised GD AP and GD AT2 datasets. The comparison of the expression level between the genotypes showed similar values in GD healthy and infected but slightly higher values in *M. sieboldii* and D2212 healthy and infected. The H0909 and M9 showed the lowest expression level in both the states.

Gene 30

The expression level of the gene 30 showed a significant difference only in *M. sieboldii* normalised on eIF4. For the genotypes M9 and GD no results were obtained because the primers showed no amplification. This result was further investigated and reported below in the chapter (4.5.2.3) on the marker development section. The general level was considerably higher in *M. sieboldii* and D2212 compared to that shown by H0909.

Gene 31

Significant differential expression was found in the genotypes GD AP infected normalised on EF and M9 normalised on eIF4. For H0909 the difference between the expression in infected and healthy plants was significant in both EF- and eIF4-normalised data. However, while in GD an over expression was observed in H0909 the gene resulted to be down regulated. In M9 the ratio was near to 1 although it resulted significantly different in the eIF4 normalised data. From the comparison of the general expression level similar values for GD in both infected states and *M. sieboldii* and D2212 in both the healthy and infected states could be observed. In addition, while H0909 healthy showed similar expression level to GD healthy in H0909 infected the level was comparable to that in M9 healthy and infected.

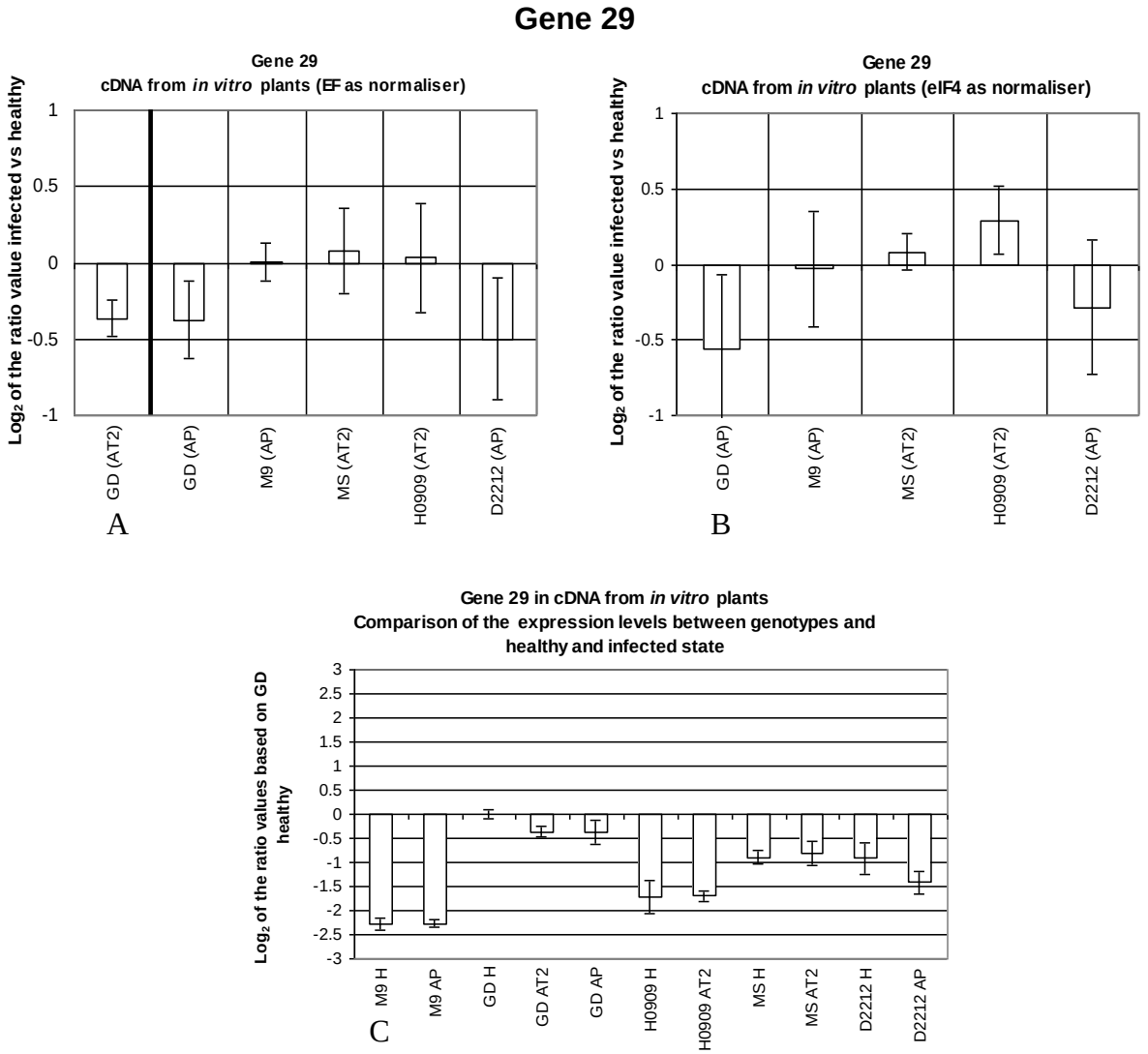


Figure 42 Gene 29 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF). Green coloured bar represent significant differences in the expression level between infected and healthy state.

Table 44 P values for the two-tailed t-test (α=0.05) on the Gene 29 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
5.35E-03	**4.59E-02**	9.49E-01	6.54E-01	8.79E-01	1.17E-01	EF
	2.85E-01	9.16E-01	4.37E-01	1.03E-01	3.48E-01	eIF4

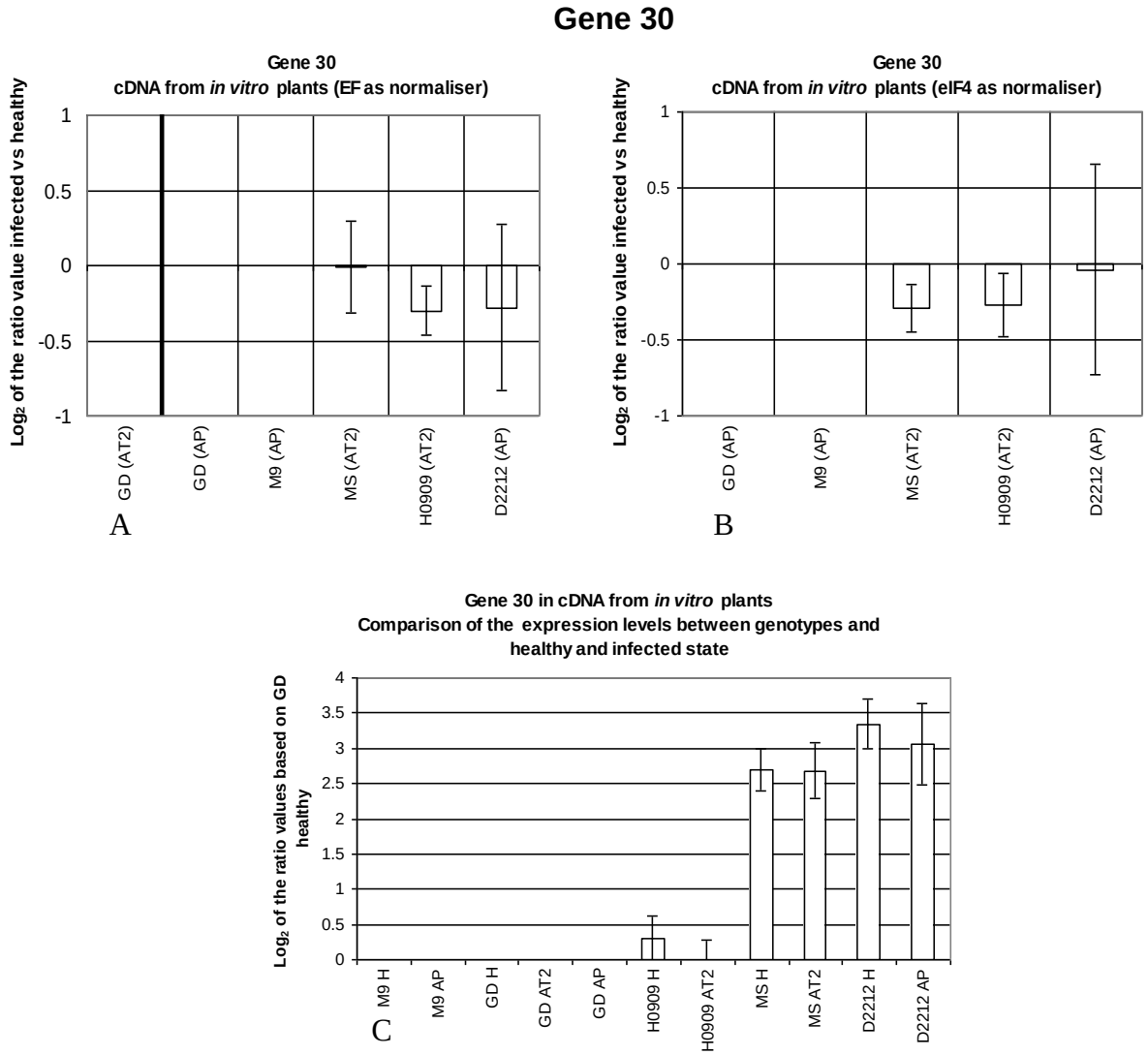


Figure 43 Gene 30 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF).

Table 45 P values for the two-tailed t-test (α=0.05) on the Gene 30 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
			9.65E-01	9.97E-02	4.03E-01	EF
			3.47E-02	2.03E-01	9.30E-01	eIF4

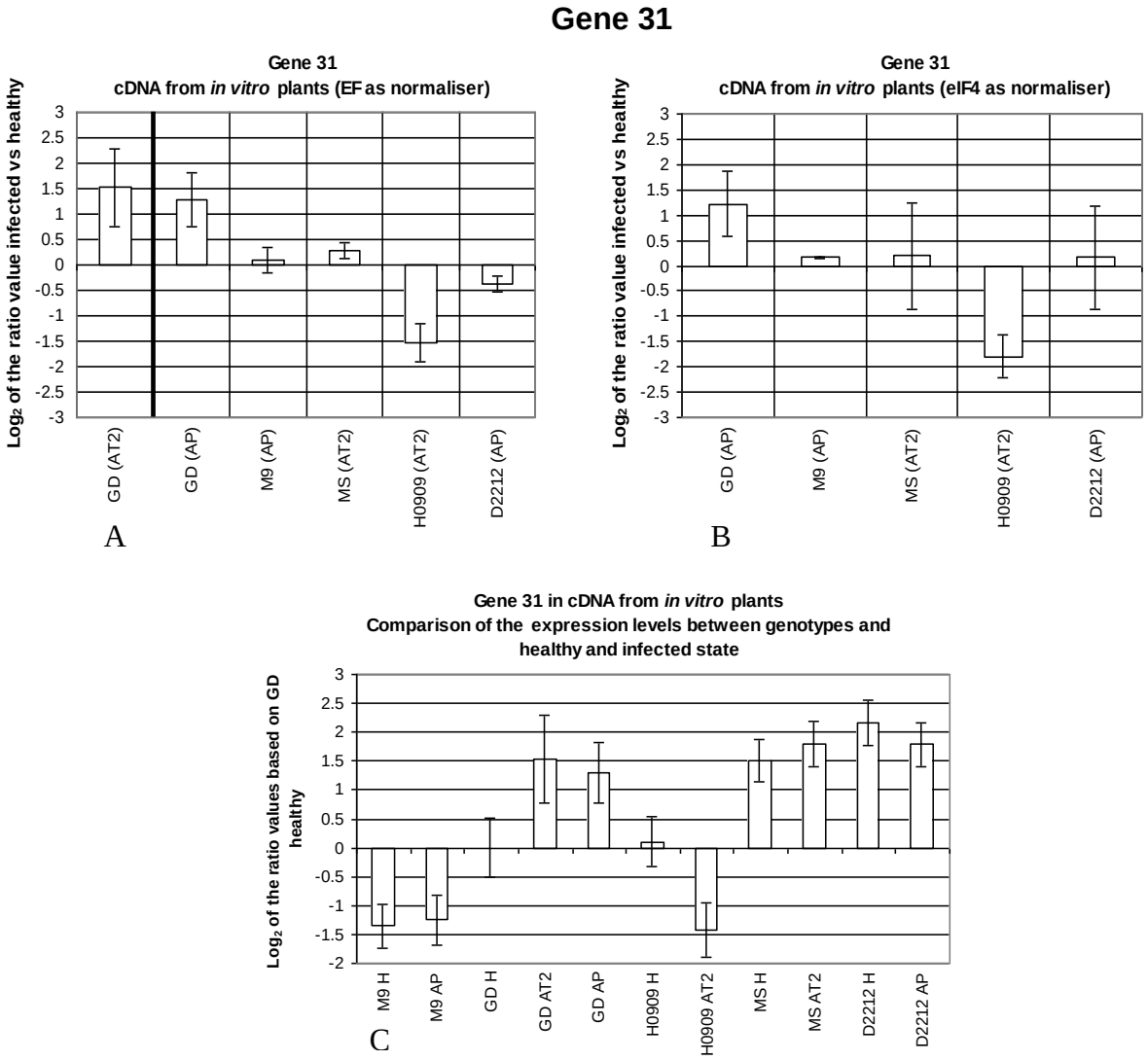


Figure 44 Gene 31 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF). Blue coloured bars represent significant differences in the expression level between infected and healthy state.

Table 46 P values for the two-tailed t-test (α=0.05) on the Gene 31 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
7.50E-02	**2.35E-02**	6.34E-01	1.39E-01	**3.55E-02**	8.93E-02	EF
	1.48E-01	**5.48E-03**	7.65E-01	**1.02E-02**	8.06E-01	eIF4

Gene 32

The relative expression values of the gene 32 resulted significant in both normalised datasets for GD AP, *M. sieboldii* and H0909. Moreover, in GD AT2-infected the difference was significant for the EF normalised data. In the comparison between the genotypes the highest expression level was shown in M9 healthy, H0909 healthy and *M. sieboldii* infected. Comparable levels were shown in M9 infected, GD AT2-infected, H0909 infected, *M. sieboldii* healthy and D2212 healthy and infected. Interestingly, although in all the genotypes the level was higher than in GD healthy (the reference) the contrary was observed in GD AP-infected were the level was about 2 times lower.

Gene 33

No significant differential expression was seen for the gene 33 except for the eIF4-normalised data of M9 resulting down regulated. Although the data produced for the gene 33 were affected by high standard deviation it could be observed that tendentially the expression level compared to the reference GD healthy was relatively higher in M9 and lower in H0909 infected.

Gene 35

The gene 35 resulted significantly differentially expressed in *M. sieboldii* and H0909 in both the normalised data sets. However, while the gene showed to be up regulated in *M. sieboldii* in H0909 it resulted to be down regulated. In addition, it was significantly down regulated in the EF-normalised dataset of GD AP and in the eIF4-normalised dataset of M9 AP. The expression levels compared in all the genotypes and states showed similar levels between M9 and H0909 in infected and healthy state, respectively. *M. sieboldii* and D2212 in both infected and healthy state and GD AP-infected showed a slightly higher expression level compared to M9 and H0909. Hovewer, in all the genotypes the level was lower that in GD healthy except in GD AT2-infected were it remained almost unchanged.

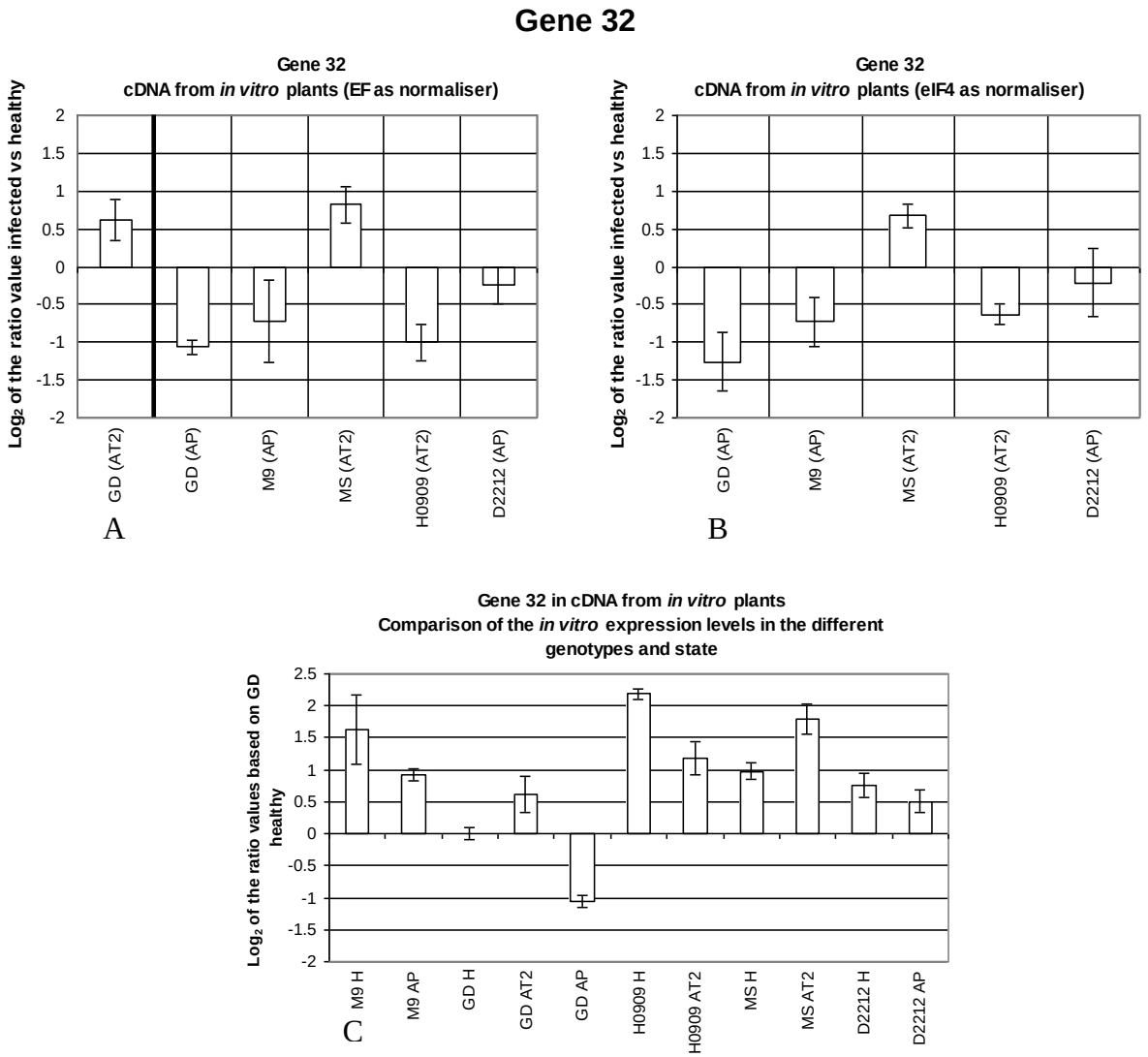


Figure 45 Gene 32 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF). Blue and green coloured bars represent significant differences in the expression level between infected and healthy state.

Table 47 P values for the two-tailed t-test (α=0.05) on the Gene 32 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
3.23E-02	**7.70E-05**	2.73E-01	**8.92E-03**	**1.44E-02**	1.46E-01	EF
	7.37E-03	6.39E-02	**2.61E-02**	**2.93E-02**	4.57E-01	eIF4

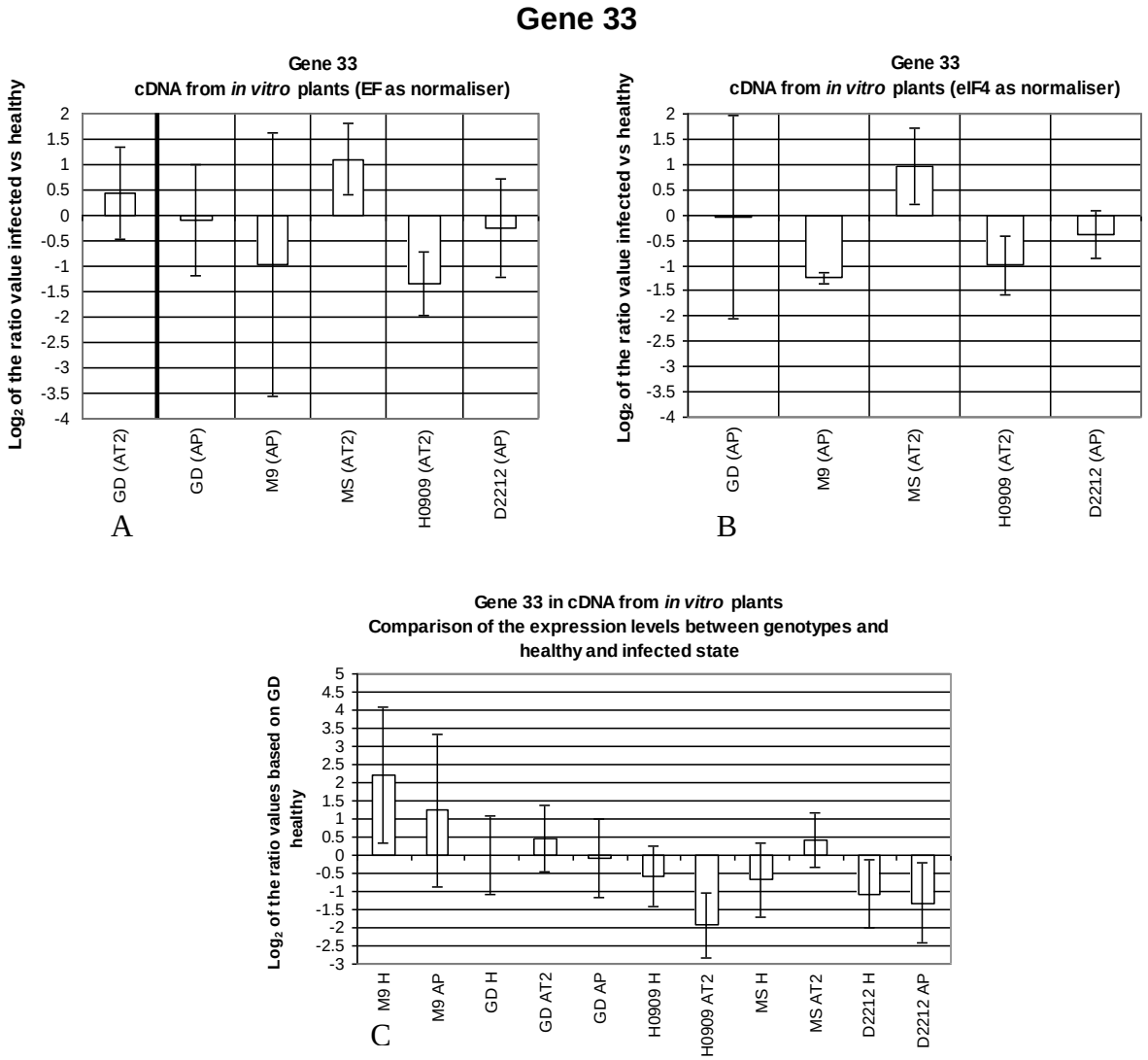


Figure 46 Gene 33 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF).

Table 48 P values for the two-tailed t-test (α=0.05) on the Gene 33 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
5.53E-01	9.11E-01	5.87E-01	1.84E-01	1.07E-01	7.51E-01	EF
	9.76E-01	**9.45E-03**	2.56E-01	1.81E-01	3.71E-01	eIF4

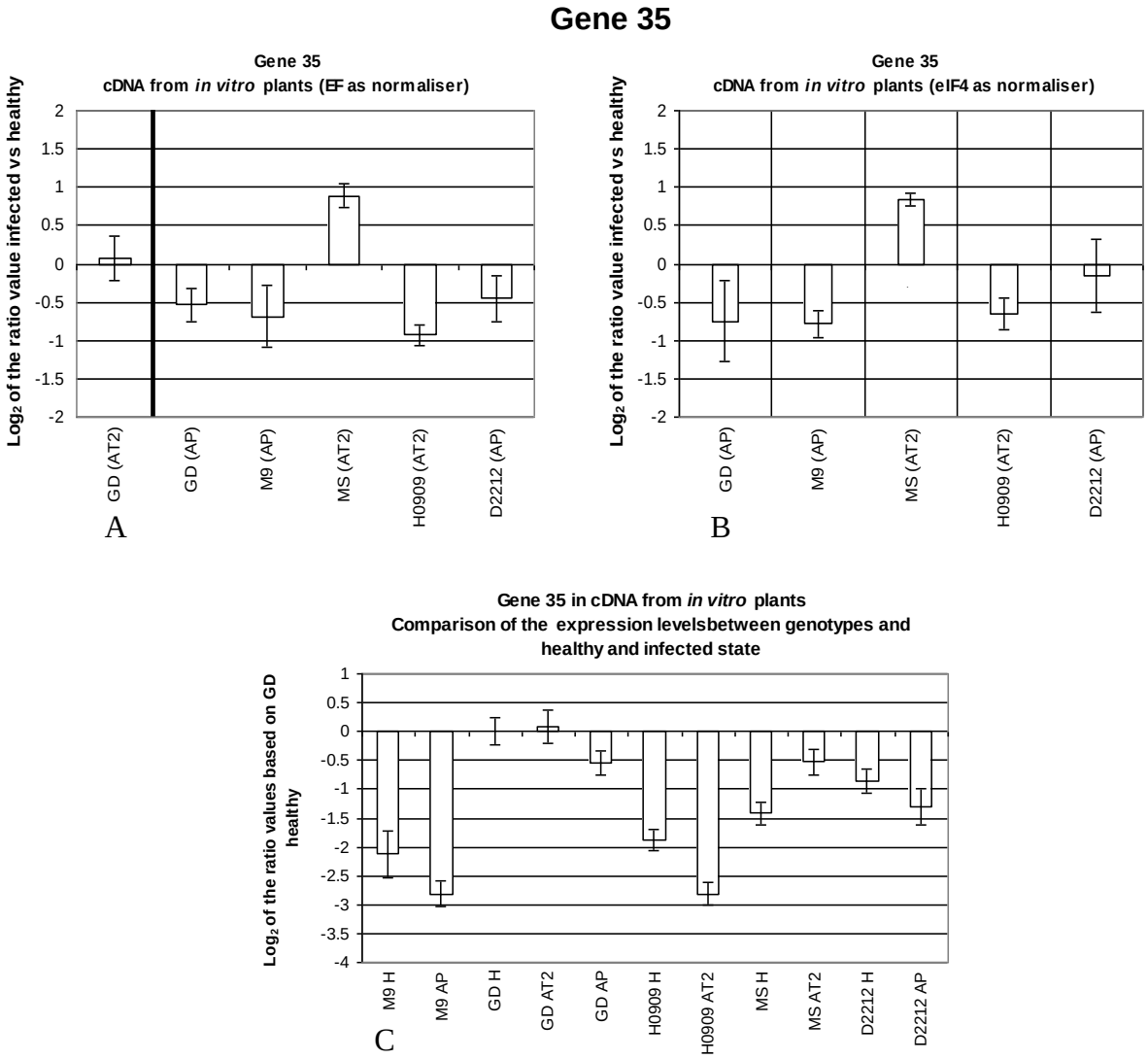


Figure 47 Gene 35 expression levels; A= normalised on EF; B= normalised on eIF4; C= general expression level (normalisation on EF). Blue coloured bars represent significant differences in the expression level between infected and healthy state.

Table 49 P values for the two-tailed t-test (α=0.05) on the Gene 35 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP	
6.53E-01	**1.47E-02**	7.24E-02	**2.11E-02**	**1.27E-02**	1.53E-01	EF
	1.87E-01	**2.04E-03**	**4.69E-04**	**1.52E-02**	5.78E-01	eIF4

In the Table 50 are resumed the results of the real-time analysis on cDNA of *in vitro* plants.

Table 50 Summary of the gene expression analysis by qPCR on cDNA of *in vitro* plants. Significantly differentially expressed genes after normalisation with eIF4 or EF are indicated by arrows for upregulation (▲) or downregulation (▼)

Gene	Normaliser	GD AT2	GD AP	M9 AP	MS AT2	H0909 AT2	D2212 AP
3	eIF4		-	-	-	-	-
	EF	-	-	-	-	-	-
5	eIF4		▲	-	▲	▼	-
	EF	▲	▲	-	▲	▼	-
6	eIF4		-	▲	▼	▲	-
	EF	-	-	▲	▼	-	-
7	eIF4		▼	-	-	-	-
	EF	-	-	-	-	-	-
9	eIF4		-	-	-	▲	-
	EF	-	-	-	-	▲	-
11	eIF4		-	-	▲	-	-
	EF	-	-	-	▲	-	-
14	eIF4		▲	-	▼	-	-
	EF	-	▲	-	-	-	-
15	eIF4		-	-	-	▼	▼
	EF	▲	-	-	-	-	▼
21	eIF4		-	-	▼	▲	▲
	EF	-	-	-	-	▲	-
22	eIF4						
	EF						
24	eIF4		-	▼	-	▼	-
	EF	-	▼	-	-	-	▼
25	eIF4		-	-	▲	▼	-
	EF	-	-	-	▲	▼	-
27	eIF4		-	▼	-	-	-
	EF	-	-	-	-	-	▲
28	eIF4		▲	-	▲	-	▲
	EF	▲	▲	-	▲	▼	▲
29	eIF4		-	-	-	-	-
	EF	▼	▼	-	-	-	-
30	eIF4				▼	-	-
	EF				-	-	-
31	eIF4		-	▲	-	▼	-
	EF	-	▲	-	-	▼	-
32	eIF4		▼	-	▲	▼	-
	EF	▲	▼	-	▲	▼	-
33	eIF4		-	▼	-	-	-
	EF	-	-	-	-	-	-
35	eIF4		-	▼	▲	▼	-
	EF	-	▼	-	▲	▼	-

4.4.5 Real-time PCR analysis on RNA of *ex vitro* plants

The high standardisation of the *in vitro* plant material allowed to minimise the effects of uncontrolled parameters in the studied system. Nevertheless, the expression profile analysis in plants grown *in vitro* could be not fully representative of the real behaviour of the genes in conditions at which the plants normally grow in open field. In particular, no root organs were tested in the *in vitro* system as the plants were cultured without roots. Therefore, the real-time analysis

was performed on healthy and infected *ex vitro* plants maintained in the greenhouse but grown in conditions comparable to those present in the orchards. For each genotype, the preparation of the *ex vitro* plants was carried out starting from 20 *in vitro* plants that were graft-infected and grown as described in 3.2.1.1. Those plants that resulted infected were prepared for the rooting as described in 3.2.1.2 and transferred in pots. The *ex vitro* plants were further maintained in the greenhouse for two years and then the material was collected for the RNA extraction in the month of July where the seasonal phytoplasma colonisation starts to be high especially under greenhouse conditions (Pedrazzoli *et al.*, 2008).

In order to compare the expression levels in different tissues, for each genotype roots and leaves from three infected and healthy plants were respectively pooled and immediately frozen (3.2.2.2). The gene expression analysis was then directly performed on the RNA extracted from each pool using a one step quantitative reverse transcription PCR (one step qRT PCR) as described in 3.2.20.1. The genotypes used in the analysis are reported in Table 51. The study was focused first on the root tissue and only for H0909 and D2212 leaf tissue was also included in the analysis. As described in the housekeeping gene analysis (4.4.3) the behaviour of the primer pair for the control gene EF was not optimal due to primer dimer production when applied for the real-time analysis using directly RNA as template. Therefore, because no other good choices were available between the tested housekeeping genes, the 18S rRNA was employed as a second control in the normalisation of the results obtained from the *ex vitro* plants. Moreover, only the half of the genes was analysed (genes 3 to 22) on these plants while the remaining half will be object of a future study.

Table 51 Genotypes and tissues used in the real-time PCR analysis of the *ex vitro* plants

Genotypes	Tissue	'*Ca.* P. mali' strain
Malus sieboldii	Roots	AT2
Golden delicious	Roots	AP
H0909	Roots, Leaves	AT2
M9	Roots	AP
D2212	Roots, Leaves	AP

Gene 3

The gene 3 resulted to be significantly overexpresed in the roots of infected GD and M9 genotypes in both the 18S rRNA- and the eIF4-normalised data. In D2212 leaves, only the data normalised on the 18S rRNA showed a significant differential expression.
From the comparison of the general expression level, using GD healthy root as reference, M9 showed the highest expression relative to the other genotypes. In *M. sieboldii*, H0909 root and D2212 (in both leaf and root tissues) the level of gene 3 was comparable while in H0909 infected and healthy leaf it was slightly lower similar to that observed in GD infected root.

Gene 5

The gene 5 was significantly differentially expressed in all genotypes in all tissues. Only for M9 the t-test P value was not significant in the eIF4-normalised dataset. Interestingly, in all resistant genotypes the gene resulted to be overexpressed while in GD root it was down regulated. The general expression level was comparable in the infected root tissues of all genotypes except for GD where the level decreased. Similarly, the level in healthy root tissue was comparable between *M. sieboldii*, H0909 and D2212. In H0909 leaf the expression level in the two states was in the range of infected and healthy root respectively. In M9 no differences were seen after infection and the level remained unchanged but comaparable to that in D2212 leaf in which it resulted to be the highest among all the genotypes.

Gene 7

The differential gene expression of gene 7 was significant in M9, H0909 roots and leaves, and D2212 leaves in both the normalised datasets. In D2212 roots the level after infection was significantly different only in the eIF4-normalised data but the values obtained from the two datasets were considerably diverse. The general expression level was comparable in GD infected and healthy roots and H0909 infected and healthy leaves. In *M. sieboldii*, H0909 roots and D2212 leaves and roots the level was similar in both the infected and healthy state while M9 was the one with the highest expression level.

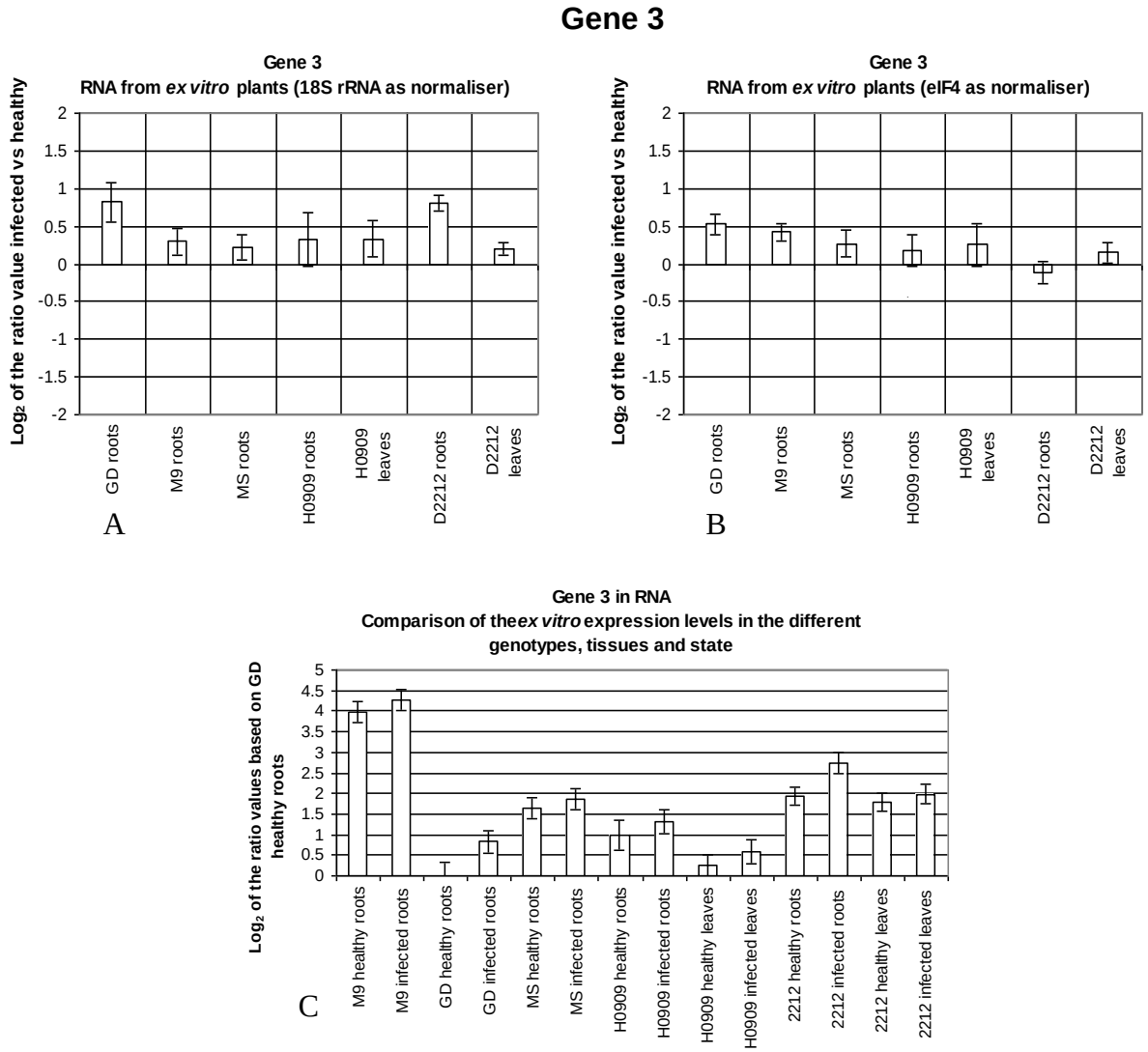


Figure 48 Gene 3 expression levels; A= normalised on 18S rRNA; B= normalised on eIF4; C= general expression level (normalisation on 18S rRNA). Blue coloured bars represent significant differences in the expression level between infected and healthy state.

Table 52 P values for the two-tailed t-test (α=0.05) on the Gene 3 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD roots *ex vitro*	M9 roots *ex vitro*	MS roots *ex vitro*	H0909 roots *ex vitro*	H0909 leaves *ex vitro*	D2212 roots *ex vitro*	D2212 leaves *ex vitro*	
4.22E-03	**4.94E-02**	1.93E-01	1.77E-01	8.25E-02	6.98E-02	**2.02E-02**	18S rRNA
5.70E-03	**1.68E-03**	1.51E-01	3.41E-01	3.05E-01	3.62E-01	2.15E-01	eIF4

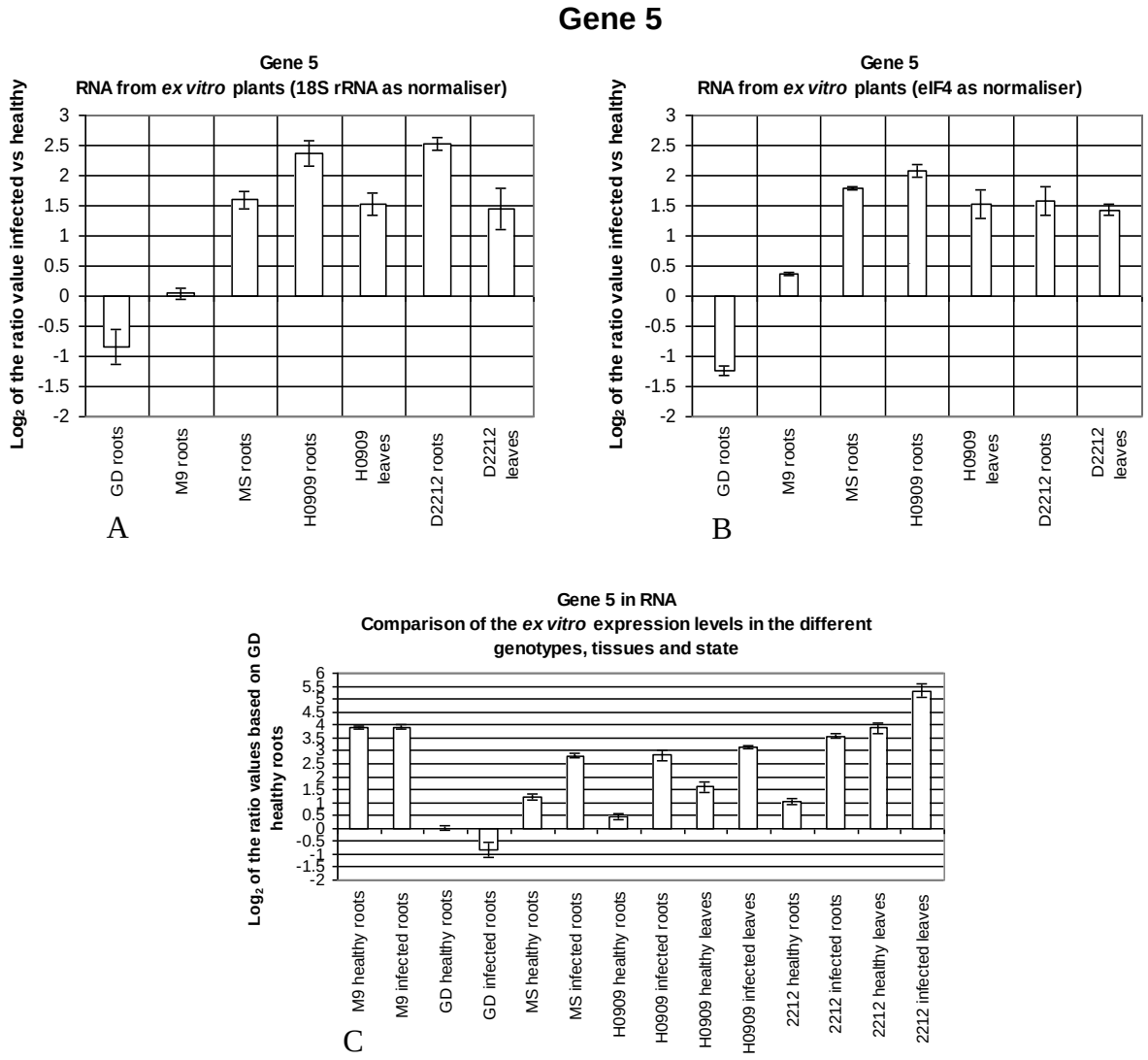


Figure 49 Gene 5 expression levels; A= normalised on 18S rRNA; B= normalised on eIF4; C= general expression level (normalisation on 18S rRNA). Blue coloured bars represent significant differences in the expression level between infected and healthy state.

Table 53 P values for the two-tailed t-test (α=0.05) on the Gene 5 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD roots *ex vitro*	M9 roots *ex vitro*	MS roots *ex vitro*	H0909 roots *ex vitro*	H0909 leaves *ex vitro*	D2212 roots *ex vitro*	D2212 leaves *ex vitro*	
2.92E-03	6.49E-01	**3.68E-03**	**9.36E-03**	**3.27E-03**	**8.57E-05**	**5.06E-03**	18S rRNA
7.89E-04	**4.12E-03**	**4.32E-05**	**9.51E-04**	**1.94E-04**	**1.92E-02**	**1.50E-04**	eIF4

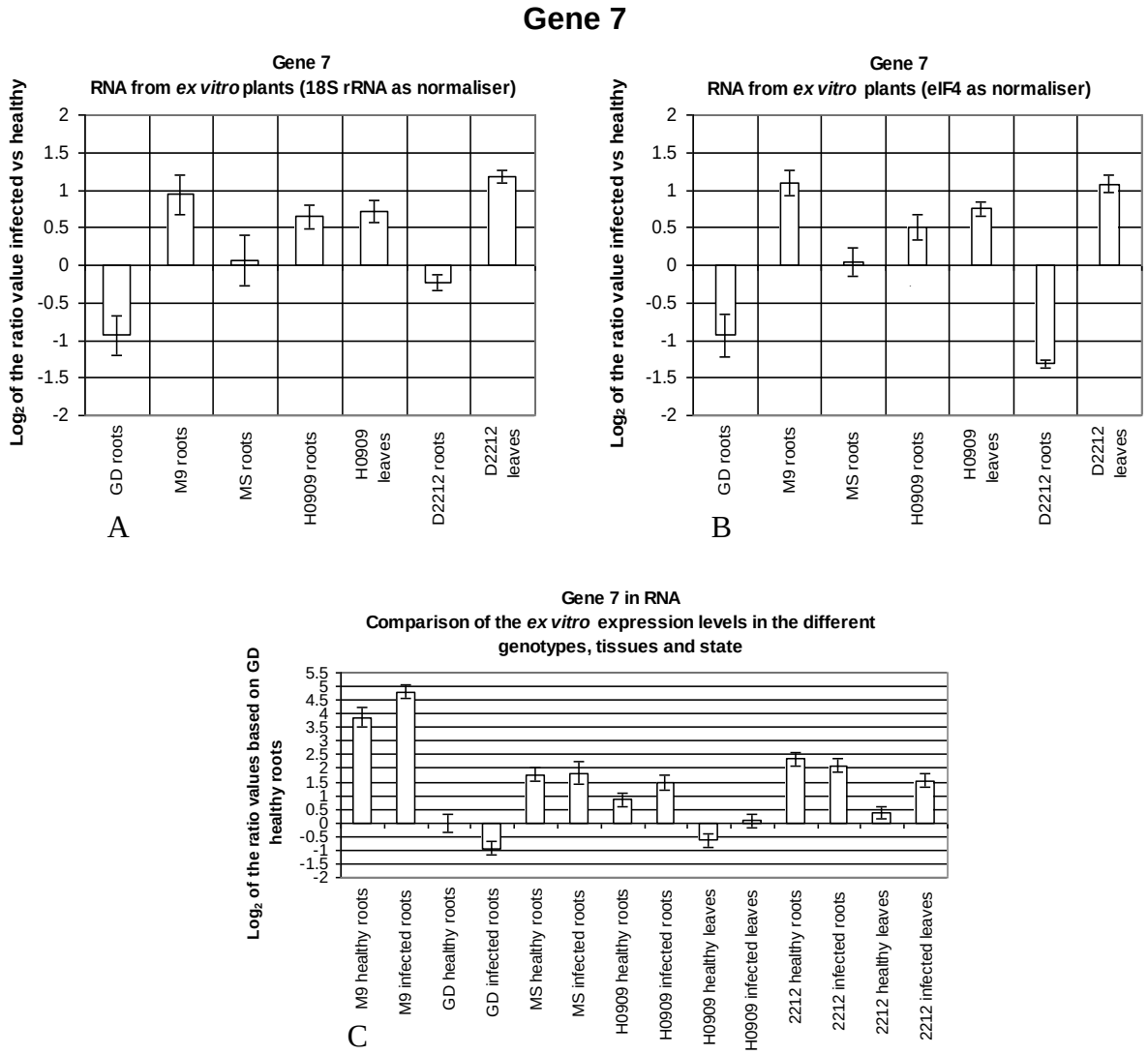


Figure 50 Gene 7 expression levels; A= normalised on 18S rRNA; B= normalised on eIF4; C= general expression level (normalisation on 18S rRNA). Blue coloured bars represent significant differences in the expression level between infected and healthy state.

Table 54 P values for the two-tailed t-test (α=0.05) on the Gene 7 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD roots *ex vitro*	M9 roots *ex vitro*	MS roots *ex vitro*	H0909 roots *ex vitro*	H0909 leaves *ex vitro*	D2212 roots *ex vitro*	D2212 leaves *ex vitro*	
5.65E-02	**1.18E-02**	8.24E-01	**8.42E-03**	**4.09E-03**	6.72E-02	**3.58E-04**	18S rRNA
1.85E-01	**1.10E-03**	6.83E-01	**1.92E-02**	**6.37E-04**	**1.59E-03**	**9.35E-03**	eIF4

Gene 9

The gene 9 was significantly differentially expressed in *M. sieboldii* roots, H0909 roots and leaves and D2212 leaves in both the 18S rRNA- and eIF4–normalised datasets. In GD and D2212 roots the level was significantly different only in the 18S rRNA- and eIF4-normalised datasets, respectively. However, in D2212 while almost no changes were seen in the 18S rRNA normalised data with eIF4 the gene resulted down-regulated. The general expression level was comparable in the root tissues of *M. sieboldii*, H0909 and D2212 in both healthy and infected state and in D2212 infected leaves. GD healthy and infected roots, H0909 infected leaves and D2212 healthy leaves showed similar levels. H0909 healthy leaves showed the lowest expression level whereas M9 infected and healthy roots the highest.

Gene 11

The gene 11 was significantly differentially expressed in GD and D2212 leaves in both the normalised datasets. Moreover, the eIF4 normalised data showed a significant difference between the healthy and infected state of M9 and *M. sieboldii* roots whereas in H0909 roots this was the case only for the 18S rRNA normalised dataset. The general expression level was in a comparable range in almost all the genotypes except in M9 that showed the highest level and *M. sieboldii* where the level was higher compared to the other genotypes.

Gene 14

The expression of gene 14 was significantly differential in H0909 leaves and D2212 roots in both the normalised datasets. However, the latter showed a contrasting behaviour between the two datasets resulting over-expressed with the 18S rRNA and down-regulated with eIF4 as normaliser. Therefore, the differential expression of gene 14 in D2212 roots was not considered valid although significant differences were found with both the normalisers. The eIF4 normalised data of GD root, M9 root and D2212 leaf were significantly different between the two states while for H0909 root this was the case for the 18S rRNA normalised data. The general expression level was the highest in M9 whereas in the other genotypes the level varied in a range between 0.8 and 2.5 times more compared to GD healthy roots used as reference.

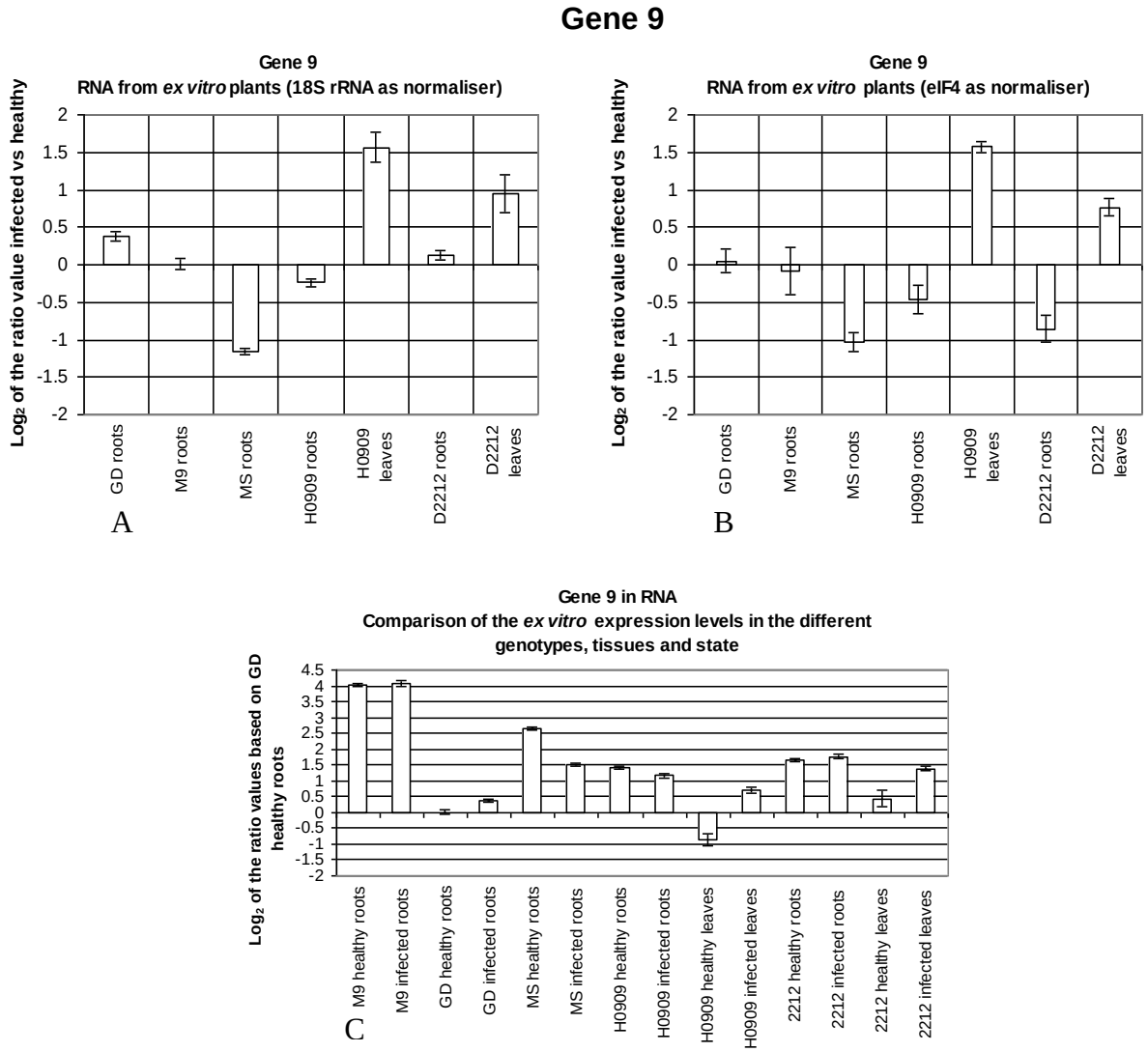
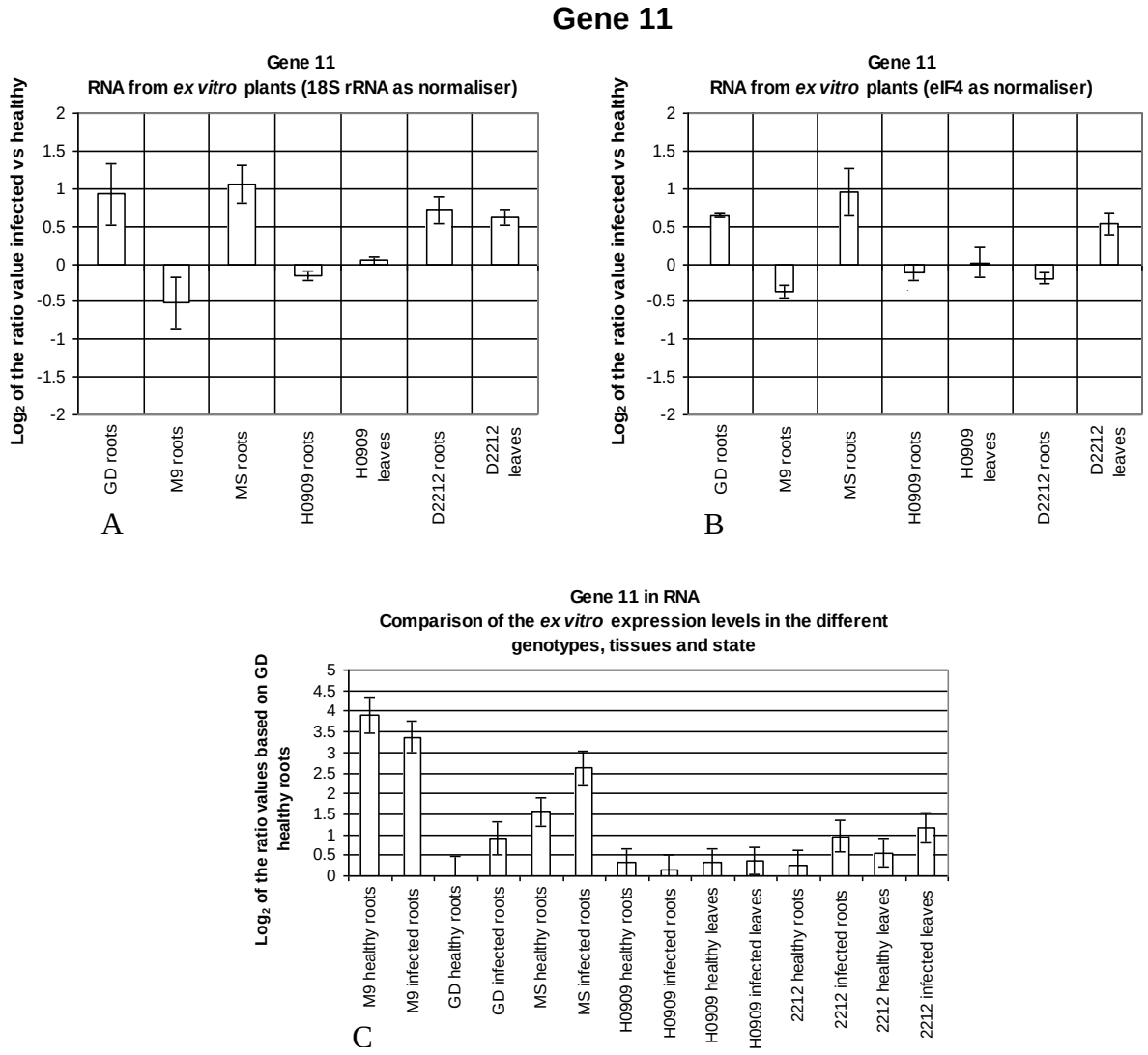


Figure 51 Gene 9 expression levels; A= normalised on 18S rRNA; B= normalised on eIF4; C= general expression level (normalisation on 18S rRNA). Blue coloured bars represent significant differences in the expression level between infected and healthy state.

Table 55 P values for the two-tailed t-test (α=0.05) on the Gene 9 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD roots *ex vitro*	M9 roots *ex vitro*	MS roots *ex vitro*	H0909 roots *ex vitro*	H0909 leaves *ex vitro*	D2212 roots *ex vitro*	D2212 leaves *ex vitro*	
1.07E-02	9.31E-01	**9.22E-04**	**1.62E-02**	**9.29E-05**	9.85E-02	**1.12E-03**	18S rRNA
6.95E-01	6.90E-01	**1.02E-03**	**2.86E-02**	**3.71E-04**	**3.37E-03**	**3.13E-04**	eIF4



<u>Figure 52</u> Gene 11 expression levels; A= normalised on 18S rRNA; B= normalised on eIF4; C= general expression level (normalisation on 18S rRNA). Blue coloured bars represent significant differences in the expression level between infected and healthy state.

<u>Table 56</u> P values for the two-tailed t-test (α=0.05) on the Gene 11 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD roots *ex vitro*	M9 roots *ex vitro*	MS roots *ex vitro*	H0909 roots *ex vitro*	H0909 leaves *ex vitro*	D2212 roots *ex vitro*	D2212 leaves *ex vitro*	
1.43E-02	1.94E-01	1.46E-01	**2.16E-02**	2.63E-01	1.31E-01	**1.04E-03**	18S rRNA
7.71E-04	**2.20E-02**	**7.86E-03**	2.29E-01	8.96E-01	5.28E-02	**2.09E-02**	eIF4

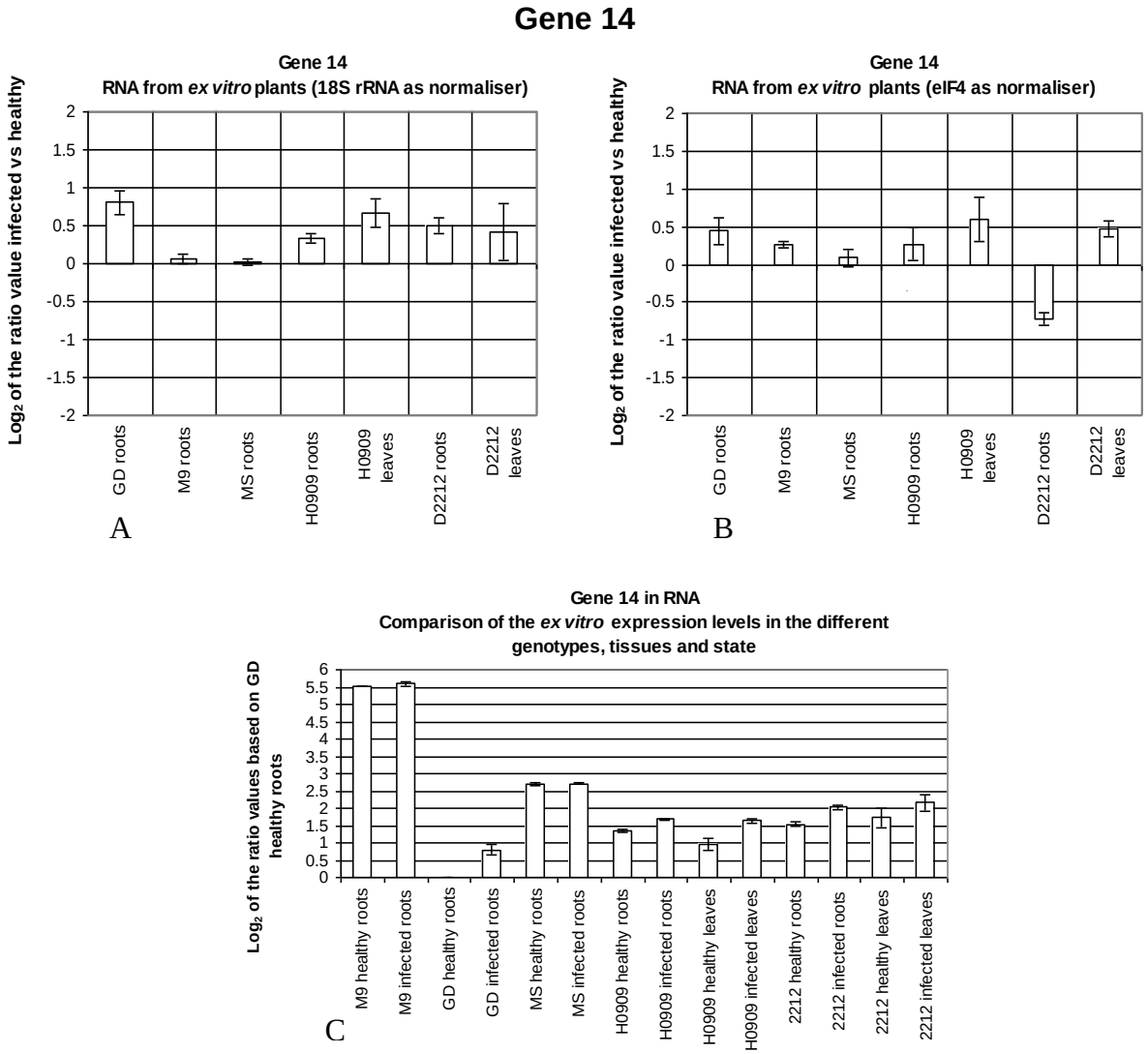


Figure 53 Gene 14 expression levels; A= normalised on 18S rRNA; B= normalised on eIF4; C= general expression level (normalisation on 18S rRNA). Blue coloured bars represent significant differences in the expression level between infected and healthy state.

Table 57 P values for the two-tailed t-test (α=0.05) on the Gene 14 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD roots *ex vitro*	M9 roots *ex vitro*	MS roots *ex vitro*	H0909 roots *ex vitro*	H0909 leaves *ex vitro*	D2212 roots *ex vitro*	D2212 leaves *ex vitro*	
1.06E-01	2.61E-01	6.64E-01	**1.38E-02**	**2.35E-03**	**2.04E-02**	1.22E-01	18S rRNA
4.34E-02	**3.58E-04**	3.89E-01	1.89E-01	**4.52E-02**	**4.13E-02**	**3.04E-02**	eIF4

Gene 15

The gene 15 was significantly differentially expressed in M9, H0909 leaves, D2212 roots and leaves. Again, the values obtained for D2212 roots with the two normalisers differed considerably but the gene 15 resulted down regulated with both datasets. In case of *M. sieboldii* roots only the 18S rRNA normalised data were significantly different in the two states. The general expression level was similar in *M. sieboldii,* H0909 and D2212 root tissues for both infected and healthy state. GD healthy and infected were comparable to D2212 infected leaves while H0909 healthy and infected leaves and D2212 healthy leaf showed an expression level lower than that of GD healthy root. Finally, M9 showed the highest level.

Gene 21

A confirmed differential expression of gene 21 was only detected in M9 with both normalisers. For the other genotypes, the 18S rRNA normalised data of *M. sieboldii* roots and H0909 leaves were significantly different between the two states while this was the case in the eIF4 normalised data for D2212 roots. Also here, the behaviour of the gene expression in D2212 root was contrasting between the data obtained from the two normalisers. The general level was comparable in all the genotypes except in GD healthy and infected that showed the lowest values and M9 healthy and infected that showed the highest expression level.

Gene 22

The gene 22 resulted significantly differentially expressed in D2212 leaves with both the controls showing up regulation after infection. In *M. sieboldii* roots normalised on eIF4, the difference was also significant but the relative high variability associated to this primer pair did not allow further considerations. The expression level was the lowest in GD healthy and infected roots and H0909 healthy and infected leaves while in the other genotypes and tissues the levels were comparable with the exception of M9 that showed the highest expression level.

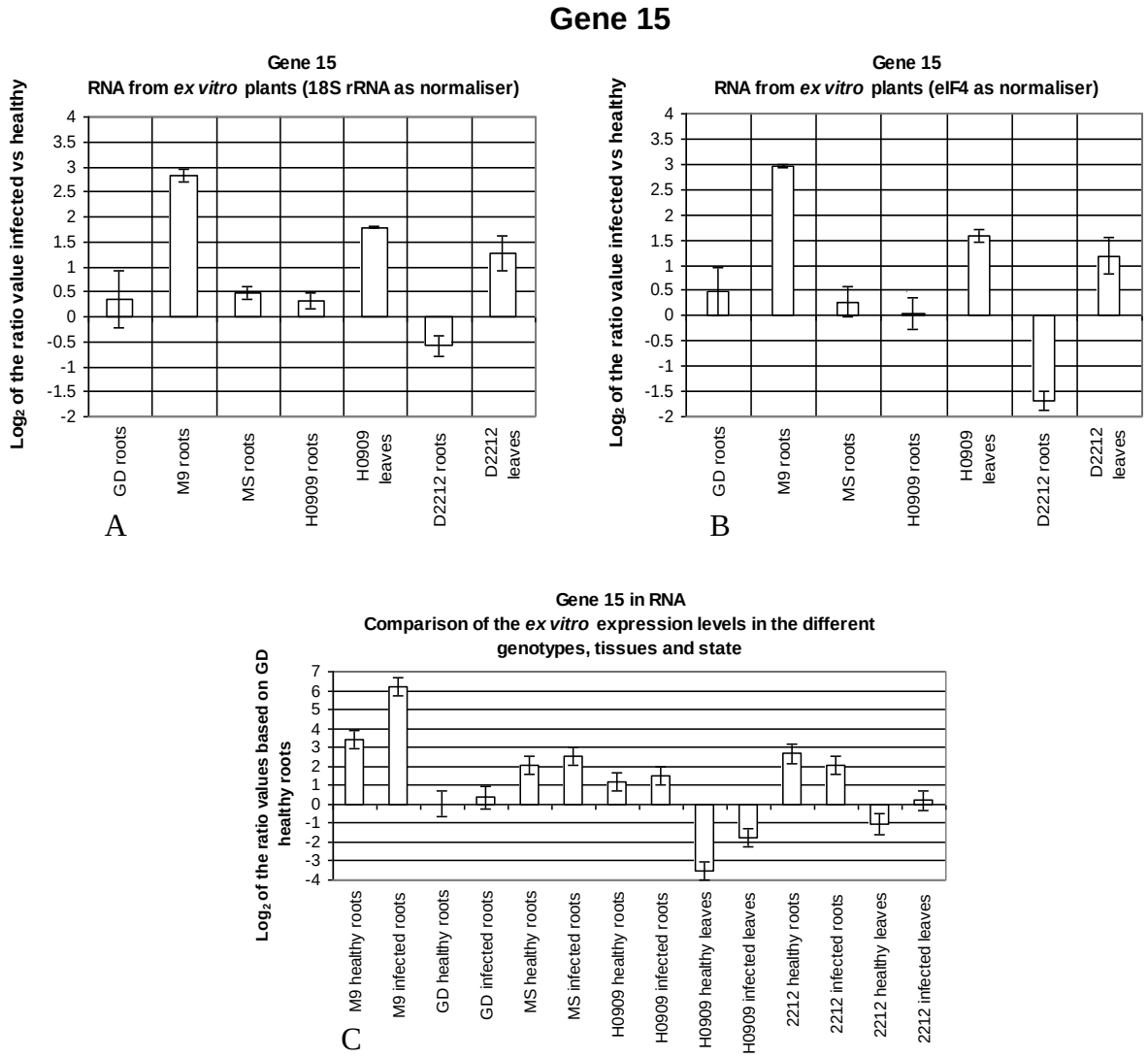


Figure 54 **Gene 15 expression levels; A= normalised on 18S rRNA; B= normalised on eIF4; C= general expression level (normalisation on 18S rRNA). Blue coloured bars represent significant differences in the expression level between infected and healthy state.**

Table 58 **P values for the two-tailed t-test (α=0.05) on the Gene 15 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.**

GD roots *ex vitro*	M9 roots *ex vitro*	MS roots *ex vitro*	H0909 roots *ex vitro*	H0909 leaves *ex vitro*	D2212 roots *ex vitro*	D2212 leaves *ex vitro*	
3.91E-01	**4.45E-02**	**2.42E-02**	8.58E-02	**2.74E-04**	**3.66E-02**	**4.37E-03**	18S rRNA
2.34E-01	**3.65E-05**	3.63E-01	8.74E-01	**2.51E-03**	**2.71E-04**	**2.94E-02**	eIF4

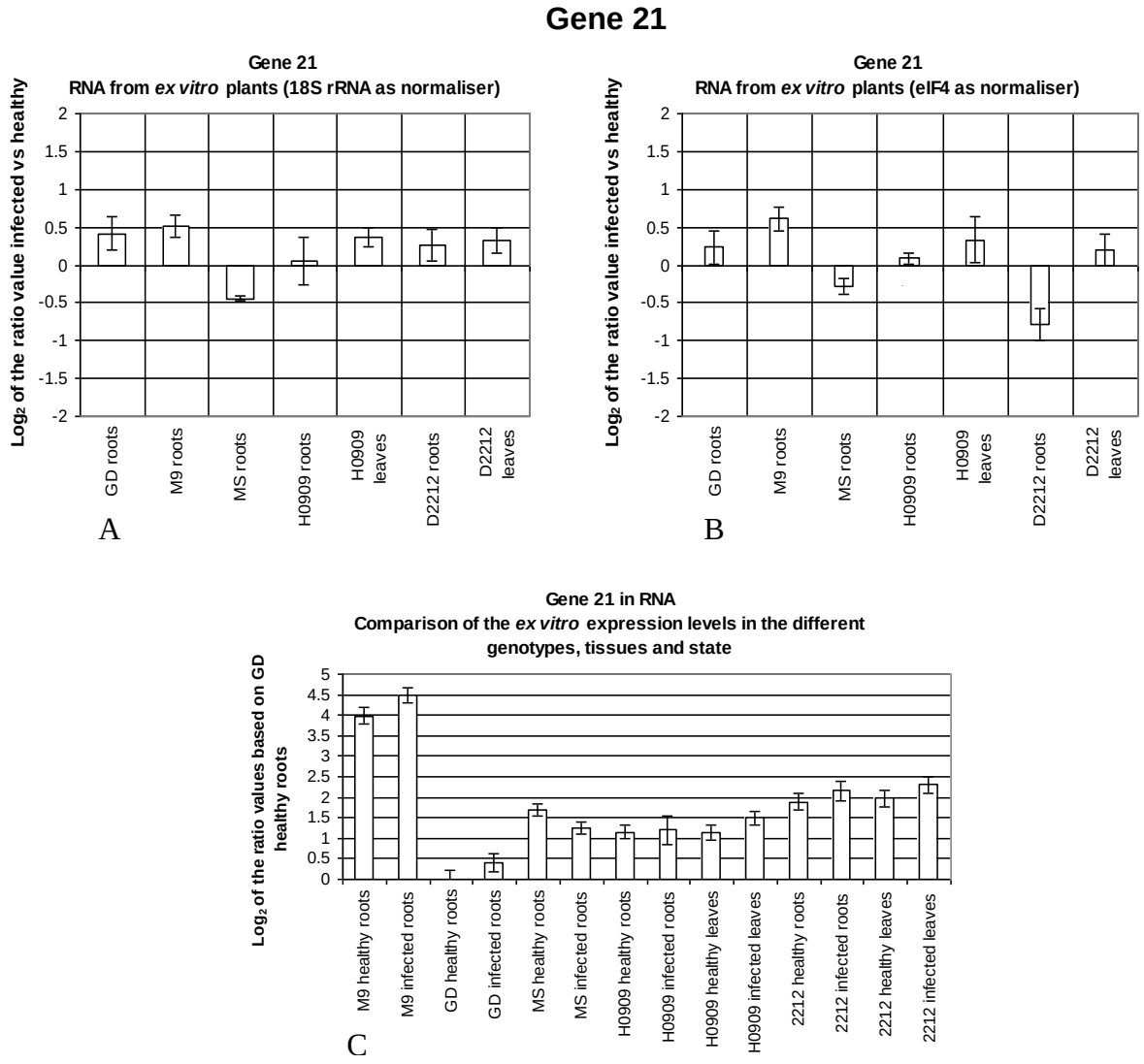


Figure 55 Gene 21 expression levels; A= normalised on 18S rRNA; B= normalised on eIF4; C= general expression level (normalisation on 18S rRNA). Blue coloured bars represent significant differences in the expression level between infected and healthy state.

Table 59 P values for the two-tailed t-test (α=0.05) on the Gene 21 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.

GD roots *ex vitro*	M9 roots *ex vitro*	MS roots *ex vitro*	H0909 roots *ex vitro*	H0909 leaves *ex vitro*	D2212 roots *ex vitro*	D2212 leaves *ex vitro*	
5.44E-02	**3.60E-02**	**2.34E-03**	7.67E-01	**4.90E-02**	1.26E-01	5.49E-02	18S rRNA
2.18E-01	**6.54E-03**	6.08E-02	2.86E-01	2.28E-01	**3.14E-03**	3.03E-01	eIF4

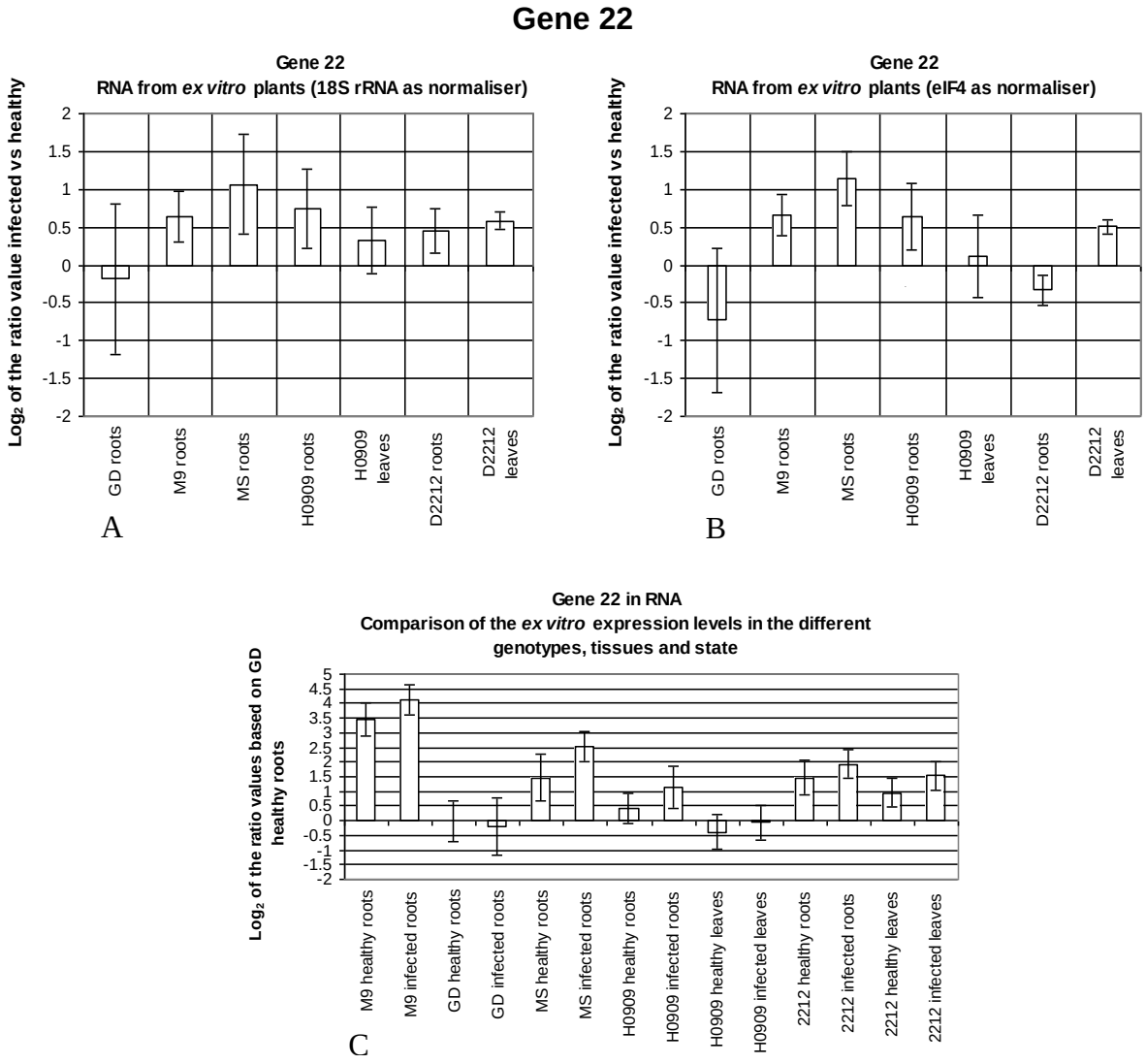


Figure 56 **Gene 22 expression levels; A= normalised on 18S rRNA; B= normalised on eIF4; C= general expression level (normalisation on 18S rRNA). Blue coloured bars represent significant differences in the expression level between infected and healthy state.**

Table 60 **P values for the two-tailed t-test (α=0.05) on the Gene 22 expression levels between infected and healthy state. Values in bold represent a significant difference between the two sets of expression values: the infected compared to the healthy.**

M9 roots *ex vitro*	GD roots *ex vitro*	MS roots *ex vitro*	H0909 roots *ex vitro*	H0909 leaves *ex vitro*	D2212 roots *ex vitro*	D2212 leaves *ex vitro*	
1.01E-01	7.52E-01	5.76E-02	2.60E-01	2.76E-01	9.09E-02	**1.41E-02**	18S
1.11E-01	2.53E-01	**1.04E-02**	2.39E-01	7.22E-01	1.24E-01	**2.71E-03**	eIF4

In the Table 61 are resumed the results of the real-time analysis on RNA of *ex vitro* plants.

Table 61 Summary of the gene expression analysis by qPCR on cDNA of *in vitro* plants. Significantly differentially expressed genes after normalisation with eIF4 or EF are indicated by arrows for upregulation (▲) or downregulation (▼)

Gene	normaliser	GD Root	M9 Root	MS Root	H0909 Root	H0909 Leaf	D2212 Root	D2212 Leaf
3	eIF4	▲	▲	-	-	-	-	-
	18S	▲	▲	-	-	-	-	▲
5	eIF4	▼	▲	▲	▲	▲	▲	▲
	18S	▼	-	▲	▲	▲	▲	▲
7	eIF4	-	▲	-	▲	▲	▼	▲
	18S	-	▲	-	▲	▲	-	▲
9	eIF4	-	-	▼	▼	▲	-	▲
	18S	▲	-	▼	▼	▲	▼	▲
11	eIF4	▲	▼	▲	-	-	-	▲
	18S	▲	-	-	▼	-	-	▲
14	eIF4	▲	▲	-	-	▲	▼	▲
	18S	-	-	-	▲	▲	▲	-
15	eIF4	-	▲	-	-	▲	▼	▲
	18S	-	▲	▲	-	▲	▼	▲
21	eIF4	-	▲	-	-	-	▼	-
	18S	-	▲	▼	-	▲	-	-
22	eIF4	-	-	▲	-	-	-	▲
	18S	-	-	-	-	-	-	▲

4.5 Marker development

The production of a resistant rootstock is a long process that requires several years. The classical breeding of apple requires about 20-30 years before yielding a new candidate genotype with the characteristics of interest. The introduction in the molecular biology of the concept of genetic markers and the establishment of techniques for their development highly accelerated the research and the selection of genotypes with traits of interest by marker assisted selection. There are several types of markers that give a different grade of resolution of the genomic structure. Microsatellite (simple sequence repeats, SSR) markers were chosen in the frame of the SMAP project because of their utility in gene mapping and cross-species amplification as well as due to the large availability of multi-allelic loci (Bisognin *et al*., 2009). In addition to SSR markers, single nucleotide polymorphism (SNP) markers should be developed in order to obtain further information for the establishment of a genetic map characterising possible differences between genotypes. An applied aspect of this thesis was the development of molecular markers for AP-resistance in expressed genes. Therefore, the sequence data obtained in the different cDNA-AFLP experiments were exploited to develop SNP and cDNA-AFLP-derived markers associated to a gene function in order to increase the information necessary for the development of a genetic map of the genotypes used in the project.

4.5.1 Single nucleotide polymorphism (SNP) marker development

The sequences obtained in this work from the cDNA-AFLP analysis on the healthy genotypes (4.1.1) and selected in the hybridisation screening (4.1.3.1) from the different genotypes were analysed for the individuation of single nucleotide polymorphisms (SNP) through sequence comparison with the Staden Package suite (3.2.16). In order to confirm the polymorphic character of these markers, specific primers were developed on the resulting putative SNPs and a SNaPshot analysis (SNaPshot Multiplex System; Applied biosystems) was performed at the IASMA Institute on the genomic DNA obtained from *Malus sieboldii* and M9. In this way, the genotype used as resistant in the project was compared with the susceptible one in order to associate the markers with one genome or the other. The putative SNPs were developed on the 118 sequences obtained from the macroarray screening (4.1.3.1) and then tested on the genomic DNA of *Malus sieboldii* and M9. Nine SNPs were individuated to be useful as markers (Table 62). Moreover, from the cDNA-AFLP analysis carried out for the individuation of differentially expressed genes after the phytoplasma infection 162 sequences were obtained on which a first preliminary SNP analysis was done and 155 putative SNPs were individuated. These will be further investigated within the SMAP project through SNaPshot analysis to confirm their polymorphic character.

Table 62 Primers used in the SNaPshot analysis that proved to be polymorphic at single nucleotide level

SNP name	Primer SNaPshot	Recognised SNP in the genotypes	
		M9	MS
c1-1	CAGGCGAAATCTTCTATCATC	**TT**	**CT**
c2-2	TTTTTTTGCGGCAGATGAAGCTTT	**CC**	**CG**
c2-3	TTTTTTTTTTTTCATCGAACCAGGGCAGTA	**GG**	**CC**
c7-5	TTTTTTTTTTTTTTTTTTTTTTTACGAGAAGGAATACAAGCCA	**GG**	**AG**
c11-2	TTTTTTGTTGCTTAGAGGATAATTGAAAGG	**CC**	**CT**
c11-4	TTTTTTTTTTTTTTTTTTTTTTTTTAGGGGAATCGAAATGGA	**GG**	**AA**
c11-9	TTTTTTTTTTTTTTTTTTTTTTTTATTTATCGCCACCAGCTTT	**AA**	**AG**
c12-2	TTTTTCTATCCTAGAGTTGCTGATAACAGG	**CT**	**TT**

Table 63 Sequences and relative SNPs individuated in the comparison of the cloned inserts derived from the different genotypes used in the first cDNA-AFLP analysis (4551, H0909, M9)

>C1
GAAGAACCCAGAGCACCTTGACTCTCAACAGAGGACATAAAAGGAAAACACTAGCAGCGCCACCGCCGGTCA AGCACCCAGTTCTGAGAATGAATGCCTAACAACGACACCGACACGAAAATACACAACCAGCAATCCAAAGAA ATACTTATAGAGAAGTTTTTACGACTAATATTATTATAA**[A/G]**GATGATAGAAGATTTCGCCTGCG
>C2
GTTCCCTCCGCCATATCCTAACGTGGATGCCACAGCGTTTTTTGCAGATGTCGCCT**[C/T]**GTCAGCAATAG CGGATGCGGCAGATGAAGCTTT**[C/G]**TCGGTGTAA**[A/G]**TACTGCCCTGGTTCGATGGCTTGGTATCTCC AGCATGCATGCTCATCTTATCAAACGAGCGGAGAATCG
>C7

CAACAGCCTTGGCAGCTCCAGTGCT**[A/G]**CTAGGAATGATGTTGAAGGAAGCTGCACGTCCACCTCTCCAG
TCCTTCATTGATGGACCGTCAACAGTCTTTTGGG**[A/T]**GGCAGTGATGGAGTGCACCGTGGTCATGAGACC
CTCAACAATTCCGAACCTGTCGTTGATAACCTTGGCAAGGGGAGCAAGGCAGTTGGTAGTGCAACTGGCATT
GGAAAGAATGTGAATGT**[C/T]**TGGCTTGTATTCCTTCTCGTTCACACCAACAACGAACATGGGAGCATCCT
TGCTGGGAGCAGAGATGATAACCTTCTTTGCACCTCCCTTGATATGGGCAGCAGCTTTCTCCTTGTCGGTGA
AAACACCGGTAGACTCAACAACAAAGTCAGCACCGGCCTCACCCCATG

>C11

CAAAAGTAGAGAGCCAGCAACTTATCCATAGTGAAGA**[A/C]****[A/C]**AAG**[C/T]**GCAGGC**[A/G]**G**[A/G]**
CCTTTCAATTATCCTCTAAGCAACCCTACACA**[C/T]**ATTGCCTC**[C/T]**CCCC**[C/T]**TCCATTTCGATTC
CCCTA**[A/C]**TCTCTACTCATCGTC**[A/G]**TCTTCCTCCTCGCCGCTCTC**[A/G]**TCCTCGTCTTCATCATC
[A/G]TCATTTACCTAGA**[C/T]**TTGGACTTGTAGA**[C/T]**TCTTCTTCATCAGCTTCGTTATTTCCATCAG
CTAATTGCTTGTTGT**[A/G]**TGCCTGAATGTTCTTGTTATATTCTACCTTCCTCTTGTCTGCCTTGGCTTGA
TAGGGAGCTTTCTCGGCATCTGACAACGATTTCCATTTATCGCCACCAGCTTT**[A/G]**CCAAC*GGCAGCGA
CCGACTT*GTTGTTTGGATGCTCCTTCTTGTACTTCTCT

>C12

TCCCGGGTAAACAAACTGTCTGTCATACA**[C/T]**GTACTGCGGGT**[A/G]**CCTGTTATCAGCAACTCTAGGA
TAGAAAGTCTTGTCAGTAACAACATTAGTGATGCGGGGAGCATATCCAACACCGTTAGCACGAGGTCTCTTG
GGCTGAGGCTTTGCTGCTTCAGCCACCTCTTCTTGTCTGCC*TTG

4.5.2 cDNA-AFLP-derived markers: K, C8 and 30

As previously described, the analysis of the constitutively expressed genes and the differentially expressed genes yielded three putative markers. The markers K and C8 were individuated in the analysis of the constitutively expressed genes and are associated to the polymorphic character of the respective primers (4.1.2.2 for K; 4.1.3.3 for C8). The marker 30 was developed from the differentially expressed genes and it was found to be associated only to *Malus sieboldii* and *Malus sieboldii*-derived genotypes (4.4.1.1).

4.5.2.1 Development of the marker K

The marker K was derived from the band 3HCGTw (4.1.2.2, Table 13). Specific primers (Kin) were developed on this cDNA fragment for the production of a 200 bp amplicon. Amplification of the genomic DNA extracted from *Malus sieboldii* and *Malus sieboldii*-related genotypes produced an 1800 bp PCR product revealing the presence of a 1600 bp long intronic region. Moreover, these primers showed a polymorphic character yielding the 1800 bp band in the *Malus sieboldii* related genotypes but giving no amplification in the *Malus* x *domestica* genotypes using an annealing temperature of 60°C. Nevertheless, lowering the temperature to 54°C made it possible to obtain a product also in *M.* x *domestica* genotypes (Figure 57). The 1800 bp bands amplified from *M. sieboldii* and from *M. x domestica* were cloned (3.2.9) and sequenced (3.2.15). The sequence analysis (3.2.16) showed homology between the two products. In order to gain information about the Kin primer annealing site new primers (Table 64) outside of the 200 bp region were developed from the 3HCGTw sequence.

Table 64 Primers developed on the 3HCGTw cDNA-AFLP fragment sequence upstream (forward) and downstream (reverse) to the Kin primer annealing sites.

Forward	Reverse
5' CGACTGCGTAGTGATCCGTCAG-3'	5'-GAGCGATGAGTCTGAGTAATAGAC-3'

A PCR product was obtained in all the genotypes allowing the sequencing of the genomic regions targeted by the Kin primers. Sequence analysis revealed that in the *M. x domestica* genotypes the 3' nucleotide corresponding to the Kin forward primer (Table 13) was associated to a SNP corresponding to a T instead of a C characteristic of the *Malus sieboldii* and *Malus sieboldii*-derived genotypes. The Table 65 summarises the results from the amplification of the genomic DNA of the most important genotypes used in the project.

Table 65 PCR amplification of the K marker in different genotypes at 60°C as annealing T

Genotype	Presence of K marker at 60°C
Malus sieboldii	+
H0909	+
4551	+
D2212	+
M9	-
Golden delicious	-
Gala	-

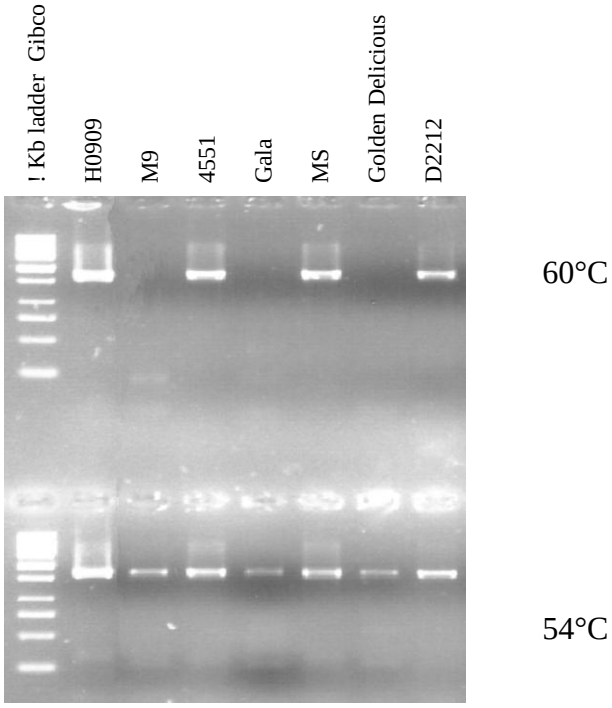


Figure 57 Agarose gel electrophoresis of PCR products obtained at different annealing temperatures with the specific primers designed on the Kin fragment with DNA of *Malus* x *domestica* and *Malus sieboldii* and *Malus sieboldii*-derived hybrids. The specific product size is 1800 bp.

4.5.2.2 Development of the marker C8

The primer pair C8 was derived from the contig 8 obtained from the sequencing of the clones individuated in the second differential screening for the comparison of constitutively expressed genes between H0909, 4551 and M9 (4.1.3.3; Table 18). The PCR carried out at 60°C using genomic DNA of the *Malus sieboldii*-derived genotypes showed in all genotypes a PCR product of an expected size of about 160 bp. In addition, a PCR product of about 800 bp was obtained in certain genotypes (Table 67). The amplification performed at 60°C on genomic DNA of *Malus* x *domestica* genotypes showed the small sized band (about 160 bp) but not the 800 bp fragment. As in the case of the marker K, lowering the annealing temperature to 54°C enabled the amplification of the 800 bp PCR product in all genotypes. In order to check if the difference was due to polymorphisms in the primer annealing-region new specific primers were designed (Table 66) external to the region amplified from the polymorphic primers. The new primers were modified for the USER PCR cloning method (3.2.9.2). In this way, the 800 bp PCR products obtained at 54°C using *M. sieboldii* and M9 genomic DNA as template were cloned. From the following sequencing two types of clones were obtained containing fragments with different size, one about 800 bp and the other about 160 bp in size. After sequence alignment the 160 bp sequence showed a 98% homology on its entire length with two regions of the 800 bp insert (Figure 58).

The sequence of the 800 bp fragments had a 94 % (562/593; E value= 0.0) identity to the sequence with accession number EB150617.1. This sequence is part of a group of expressed sequence tags (EST) forming the unigene Mdo.157 annotated as transcribed locus in *M. x domestica* genotypes. The sequences obtained from the cDNA-AFLP fragment and from the cloned PCR products amplified from the genomic DNA were assembled in a contig together with the annotated unigene Mdo.157 (which longest sequence is represented by the EB129771). The putative structure of the gene is reported in the Figure 58.

Table 66 USER-modified primers modified for the cloning of the PCR band obtained from genomic DNA for the marker C8

C8-derived USER forward primer	C8-derived USER reverse primer
5’-GGAGACA**U** TCCCGAATTTCCGACGATCC-3’	5’-GGGAAAG**U** CTTTTGAGAAGGAGCCGCGT-3’

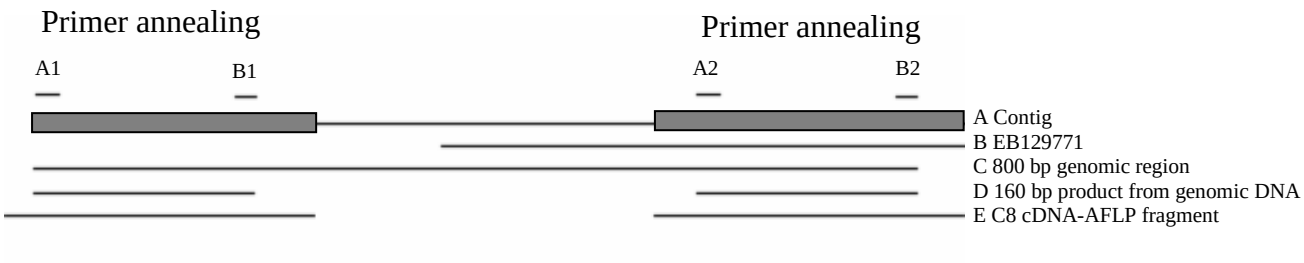


Figure 58 Schema of the C8 fragment structure. A: contig corresponding to sequences obtained from cloned fragments and the annealing with *Malus* sequences present in database; B: region corresponding to part of the sequence EB129771; C: 800 bp fragment sequenced from genomic DNA; D: localisation of the 160 bp sequences obtained from genomic DNA on the contig; E: localisation of the C8 cDNA-AFLP fragment on the contig. A1 and A2 = annealing site for the contig 8 forward primer; B1 and B2 = annealing site for the contig 8 reverse primer. Regions on the contig with high homology (>95 %) are highlighted by filled boxes.

The boxes in Figure 58 represent two regions of the Mdo.157 with high sequence homology (>95 %). Therefore, the primer pair associated to the C8 marker can anneal in two positions. In *M. sieboldii,* the primers have 100 % complementarity to the regions corresponding to A1, B1 and B2 while the region A2 contains a mismatch where a T is present instead of C at position 15 of the primer. In *M. x domestica*, this mismatch is also present in the region A1. Therefore, in *M. sieboldii* 100 % primer annealing can yield two PCR products: combination A1-B1 an 160 bp product and A1-B2 an 800 bp. On the contrary, in *M. x domestica* mismatches in both the A regions probably allow only the production of a 160 bp band due to the A1-B1 and A2-B2 combinations. Reducing the annealing temperaturet to 54 °C the less stringent reaction conditions will then allow the production of the 800 bp PCR product.

Table 67 PCR amplification of the C8 marker in different genotypes at 60°C annealing T

Genotype	Presence of C8 marker at 60°C
Malus sieboldii	+
H0909	+
4551	-
D2212	-
M9	-
Golden delicious	-
Gala	-

4.5.2.3 Development of the marker 30

The primer pair 30 developed for the real-time PCR analysis (Table 25) derived from the sequence corresponding to the band TGC1 was tested on the genomic DNA and cDNA of the *Malus sieboldii*, H0909, D2212, M9 and GD genotypes. A band was obtained only in the *Malus sieboldii* and *M. sieboldii*-derived genotypes but not in the *M. x domestica* genotypes also after lowering the annealing temperature in the PCR. In addition, primers developed externally to the

region used as target (Table 24) were also tested showing the same results as observed for the primer pair 30. Presently, no information is available on this marker and further analyses have to be carried out to better characterise the polymorphism associated to this gene.

4.5.2.4 Application of the marker K, C8 and 30 and their distribution in two different crossing populations

The markers K, C8 and 30 were applied in a screening on two different crossing populations: 'Gala' x *Malus sieboldii* and H0909 x M9. The progeny plants of these two crosses were previously characterised through SSR analysis in the frame of the SMAP project (Bisognin *et al.*, 2009). Because in the first crossing type the non-apomictic diploid *M.* x *domestica* genotype 'Gala' was used as mother only sexually derived triploid seedlings were obtained. On the contrary, in the second crossing the apomictic tetraploid genotype H0909 was used as mother. In this case, the seedlings produced were grouped in four classes depending on the combination of the parental haplotypes (Figure 3). The possible classes where: mother-like (produced asexually through apomixis; the haplotypes and ploidy are the same as the mother), hybrids I (genotypes with increased ploidy and presenting a paternal haplotype and all the maternal haplotypes), hybrid II (genotypes presenting half of the haplotypes of both the seeds and pollen parents) and autopollinated (genotypes generated by selfing) (Bisognin *et al.*, 2009). The identification of tetraploid autopollinated plants through SSR analysis caused a particular problem when a SSR allele was present twice (C. Bisognin, personal communication). The markers K, C8 and 30 applied through PCR on the progeny mentioned above showed that they were segregating in the population and that also in the case of autopollinated plants it was possible to distinguish between plants in cases where the SSR analysis was not clear. The comparison analysis between the marker segregation in hybrids II and in autopollinated plants showed a similar distribution indicating a free recombination of the haplotypes in the gametes formation. The differences observed between autopollinated and hybrid II genotypes in the number of plants presenting all the three markers or none of them can be explained with the fact that in the autopollinated plants the pollen derives from H0909 itself (characterised by having all the three markers) while in the hybrid II the father is M9 that lacks all the three markers.

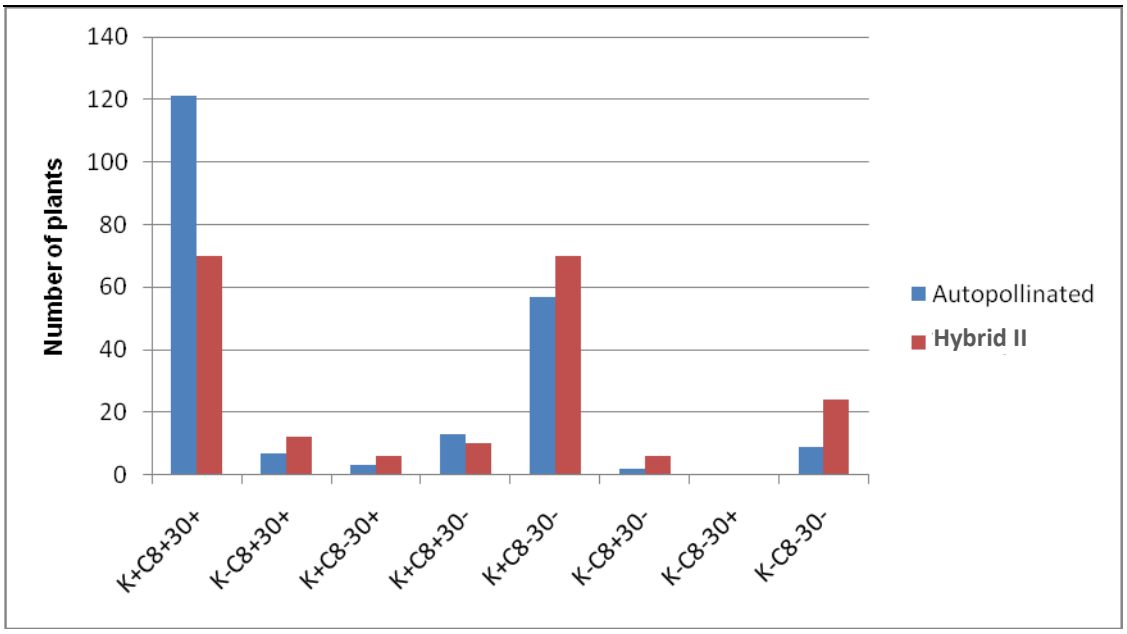


Figure 59 Distribution of the markers K, C8 and 30 in the progeny derived from the crossing H0909 x M9. The plants tested were those classified as true hybrids (blue bars) and autopollinated (red bars) in the SSR analysis.

The progeny plants derived from the crossings were also tested for their level of resistance against AP. The plants were graft-inoculated with '*Ca*. P. mali' in the greenhouse and the development of AP symptoms was recorded in the field. A cumulated disease index *(Seemüller et al.*, 2008; Bisognin *et al.*, 2008b) was established based on the type and severity of the symptoms on a four years observation period in the experimental orchard at JKI Dossenheim (Seemüller, unpublished data).

In order to find out if disease index and marker distribution in the population were correlated, the marker K, C8 and 30 were also tested on these plants. A preliminary screening was performed on 90 seedlings for the crossing 'Gala x *M. sieboldii* (Figure 60) and on 121 seedlings for the crossing H0909 x M9 (Figure 61). As shown in Figure 60 and Figure 61 the markers where distributed in almost all the disease index classes with comparable frequencies in the crossing with 'Gala' as mother. However, for the crossing with H0909 as mother the most represented class was that with disease index 3 though also here all the combinations were found.

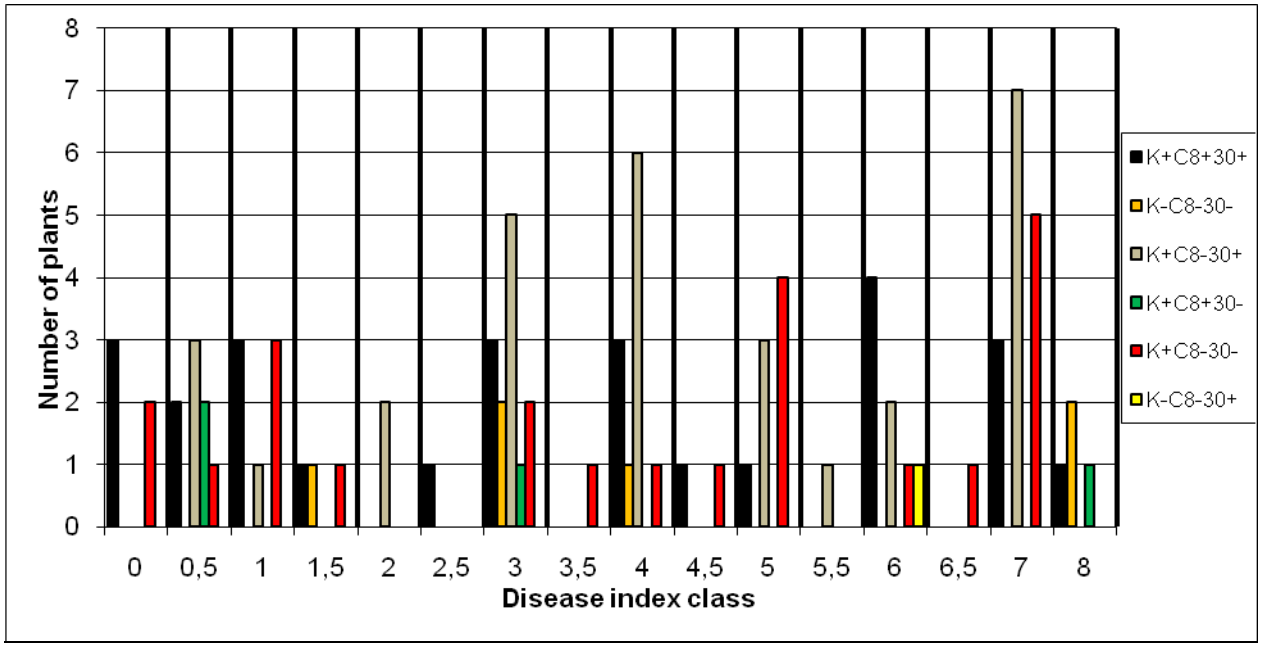


Figure 60 Distribution of the markers K, C8 and 30 in *M. x domestica* cv Gala x *M. sieboldii* progeny plants classified based on the disease index observed during AP-symptom development. Cumulated disease index ranging from 0 = no symptom to 10 = severe sypmptoms every year of observation or death of the plant

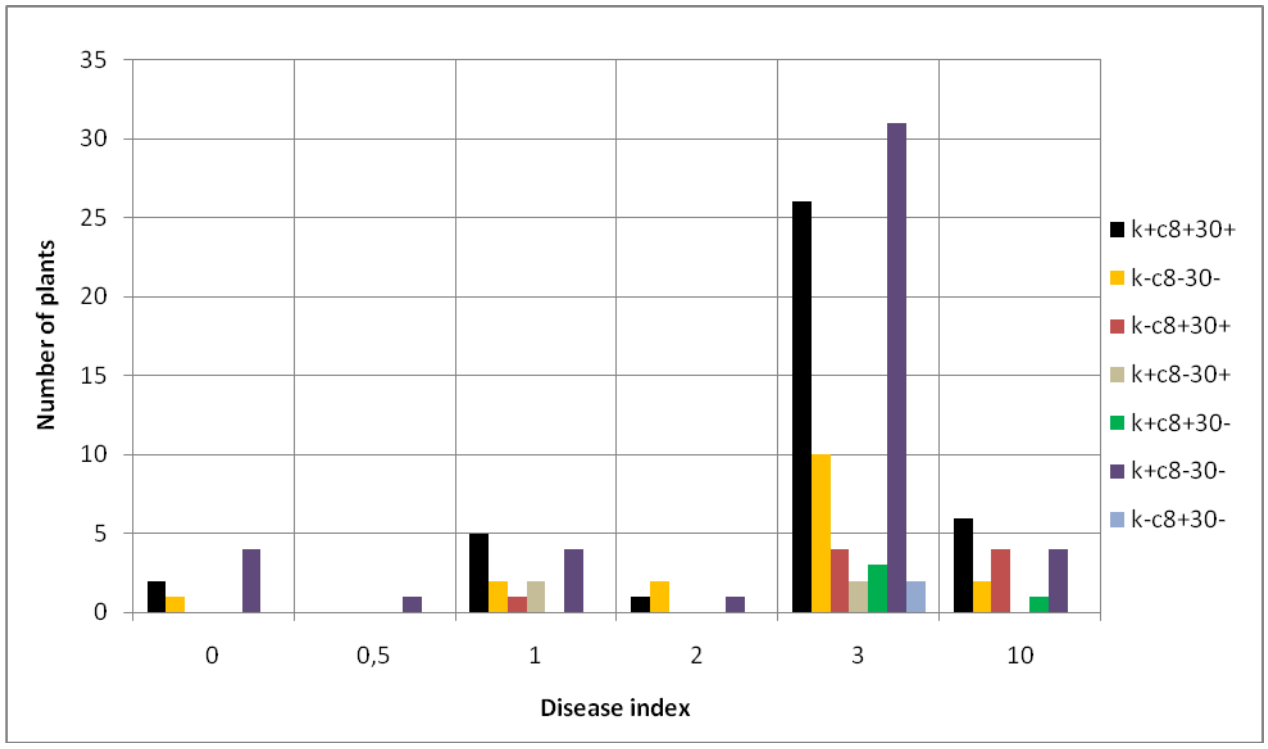


Figure 61 Distribution of the markers K, C8 and 30 in H0909 x M9 progeny plants classified based on the disease index observed during AP-symptom development. Cumulated disease index ranging from 0 = no symptom to 10 = severe sypmptoms every year of observation or death of the plant

5 Discussion

5.1 Nature of the resistance to apple proliferation

The apple proliferation disease represents a critical economic issue in important fruit growing areas of central Europe. In the last 15 years, a particularly affected zone was the Non Valley in the Trentino region (north Italy) where apple is intensively cultivated. Here, the extensive agricultural practice together with the particular environmental conditions allowed the causal agent '*Ca.* P. mali' to be efficiently transmitted causing an important spread of the disease (Branz, 2003). The predominant presence of mainly one apple cultivar (Golden delicious, GD), susceptible to AP, favoured the extended colonization of the apple trees (Branz, 2003). Moreover, like other fruit trees the apple cultivar scion is grafted and grown on a common rootstock. The use of rootstocks (M9, M11) selected for their high agronomic value but susceptible to AP disease allowed the phytoplasmas to survive in the root system over years once the scion was infected representing a permanent source of inoculum for the spread of AP.

At the beginning of this thesis apple proliferation-resistant apple genotypes with scarce agronomic value were available selected in a series of previous screenings based on phenotypic symptom observations on field-grown plants (Kartte and Seemüller, 1991). In this study, it was shown that in *M. sieboldii*-derived apomictic hybrids symptoms either never developed or developed temporarily with mild effects (during several years of field trials). It is known that '*Ca*. P. mali' is eliminated in the aerial parts due to degeneration of the phloem sieve tubes during the winter while it remains in the roots where intact sieve tubes are present throughout the year (Schaper and Seemüller, 1982). In spring, colonization of the stem may take place from the roots when new phloem is being formed (Schaper and and Seemüller, 1984; Seemüller *et al.*, 1984). This alternation in the phytoplasma distribution suggested a strategy in which resistant genotypes used as a rootstock could prevent the disease or reduce the impact in a scion cultivar. So far, the nature of the resistance mechanism is still unknown leaving open several questions. The difficulty to detect '*Ca.* P. mali' in the resistant hybrids indicated a low phytoplasma titre (Kartte and Seemüller, 1991), recently confirmed in a study based on real-time PCR analysis (Bisognin *et al.*, 2008b). However, in the same work it was shown that the phytoplasma concentration rely on host suitability and therefore cannot be considered a defining factor for resistance (Seemüller *et al.*, 2008; Bisognin *et al.*, 2008b). Nevertheless, a much lower pathogen titre was found in top-infected scions of trees grown on resistant rootstocks compared to top-infected trees on susceptible rootstocks (Bisognin *et al.*, 2008b). In the case of the resistant rootstocks, the low infection rate and low titre in the stems of trees seems to correlate with the low phytoplasma concentration in their roots. This characteristic, together with the poor host suitability, may have a negative effect on

the spread of the pathogen indicating that the low titre in the roots is likely to contribute to the resistance of *M. sieboldii*-derived stocks (Seemüller *et al.*, 2008).

Other factors are involved in the different responses to AP disease observed in the resistant compared to the susceptible genotypes. One role could be played by the different virulence of the phytoplasma strains which further complicates the characterisation of the resistance trait (Seemüller and Schneider, 2007). Considering the insufficiency of information about the resistance mechanism, structural or physiological differences determining the incapacity of the phytoplasmas to colonise the plant or the induction of resistance-related genes could be at the base of that mechanism. Differences in the susceptibility of the phloem cells of different apple genotypes to the phytoplasmas presence suggest that there could be an induced response against the pathogen characterised by callose deposition at sieve plates, necrosis and collapse of sieve elements in different grade and modality (Kartte and Seemüller, 1991; Lee *et al.*, 2000).

Studies on the gene expression are necessary for better understanding the nature of the resistance in the *M. sieboldii*-derived hybrids. In apple, in contrast to other organisms like *A. thaliana,* there is poor availability of information about genes involved in resistance mechanisms and more in general sequencing data are still missing or not jet annotated. In addition, several analysis tools developed in herbaceous plants are not directly applicable to woody plants like apple. Therefore, a specific setup of the method is often required which is a time consuming procedure.

The particular complexity of the '*Ca*. P. mali'-apple interaction, the poor knowledge on the resistance mechanism and the scarce availability of genetic data on apple plants made us to choose the cDNA amplified fragment length polymorphism (AFLP) as technique for the analysis of the expressed genes involved in this interaction.

5.2 cDNA-AFLP as technique for the gene expression pattern analysis

In this thesis, the comparison of constitutively expressed genes in healthy plants and the individuation of induced expressed genes in infected plants were both carried out through cDNA-AFLP procedures. This technique was initially developed for the analysis of expressed genes in different developmental stages of the studied organism (Bachem *et al.*, 1996). Further improvements of the method allowed broadening the range of its applicability. This together with its high sensitivity and the quality of information that can be obtained make the cDNA-AFLP a powerful technique for transcriptome analysis (Bachem *et al.*, 1996). The characterisation of the expression profiles associated with the experimental conditions gives the possibility to investigate the gene expression even in systems for which no information is available. Breyne *et al.* (2003) showed that using a modified cDNA-AFLP protocol it is possible to analyse the whole transcriptome allowing the individuation of interesting expressed sequences in a fast way.

Moreover, their results demonstrated that this technique can be as powerful as the microarray approach with the great advantage that in the cDNA-AFLP there is no necessity to have information on the sequences of the expressed genes or to know which metabolisms are involved in the studied system. These two last features were very important points for the choice of this technique as suitable for our analysis.

A prerequisite for the reliability of the cDNA-AFLP analysis is to have RNA of good quality as starting material to avoid biases in the production of the cDNA fragments. The relationship between the gene expression levels has to be maintained for the generation of transcription profiles truly informative. At the beginning of this work, the commonly used RNA extraction methods were not adequate for obtaining RNA of good quality from our plant tissues. Therefore, an extraction method originally developed for plants with high polysaccharide content was modified and optimised for the RNA extraction from woody materials and also employed for *in vitro* material (Moser *et al*., 2004).

The gene expression patterns can be influenced by environmental conditions, physiological state and time point at which the material is collected. Therefore, another important prerequisite is the production of plants that are grown under controlled conditions – a condition difficult to achieve with perennial woody plants like apple. Homogenous starting material would not only lower the variability due to environmental effects between different plants but also would reduce those differences that are not related to the plant-pathogen interaction. Therefore, for the generation of standardised material a previously established *in vitro* system for culturing and inoculating phytoplasma-infected apple was exploited.

5.2.1 Analysis of constitutively expressed genes

The knowledge and the data available at the beginning of this work were insufficient to define if the resistance observed in the group of *M. sieboldii*-derived hybrids was related to physiological differences at a constitutive level, compared to the susceptible genotypes, or induced after the phytoplasma infection. First, it was chosen to carry out a comparative analysis to investigate if constitutive differences are involved in the resistance to AP disease. The immediate accessibility to material coming from healthy field-grown plants was exploited for the set up of a cDNA-AFLP analysis for the characterisation of any differences. However, it was not clear if this procedure could be successfully applied for the analysis of genotypes with different ploidy degrees. In the cDNA-AFLP analysis expression patterns are usually directly compared between induced and non-induced state (or different developmental stages) of the same genotype allowing the individuation of differentially expressed genes. On the other hand, the comparison of constitutive expression patterns between different genotypes in a healthy state can produce many data that are complex to analyse. Nevertheless, in our system the cDNA-AFLP technique was applied in a non-standard way due to the particular choice of the genotypes to study: 4551, H0909 and M9. This approach

was justified by the genetic relationship between the genotypes used in the comparison. 4551 derives from the crossing *Malus sieboldii* x Laxton's superb while H0909 derives from the crossing 4556 x M9 where 4556 was obtained in the same crossing that produced 4551. Both genotypes H0909 and 4551 are apomictic hybrids and share as common parent the AP-resistant genotype *Malus sieboldii* (MS). Therefore, the cDNA-AFLP was applied comparing the expression patterns for the selection of bands present in both 4551 and H0909 (selection for MS-related transcripts) or in H0909 and M9 (selection for M9-related transcripts). The individuation of constitutively expressed genes common only to the apomictic genotypes was the main objective of this analysis. A second aim was the isolation of expressed sequences for the development of single nucleotide polymorphisms (SNPs) associated with the resistant hybrids. This latter was substantially a technical work related with the main project of which this thesis is part and the SNP analysis was further carried out by other researchers involved in the project for the development of genetic markers associated to the AP-resistant hybrids.

From the analysis of the cDNA-AFLP expression patterns several bands were individuated associated with the two resistant genotypes but not present in the susceptible genotype M9. After the re-amplification of the isolated cDNA fragments, several PCR products were obtained for each band. This was a major issue for the identification of the products which corresponded to the bands selected in the cDNA-AFLP analysis. This problem was known to be intrinsic to the cDNA-AFLP technique because of co-migration of fragments with the same size. In our case, the number of different PCR products from each band was higher as that observed in other cDNA-AFLP analyses performed on different systems (prof. Pezzotti personal communication). The primer combinations used for the analysis were selective for the first nucleotide of the adapted cDNA fragments (Figure 4) giving a much more complex pattern compared to that obtained by using more selective nucleotides. The choice of this technical approach derived from the observation that the patterns generated from experiments performed on other woody plants were poorly informative with primers selecting on the second or third nucleotide (prof. Pezzotti personal communication). Because of the high number of selected cDNA fragments that had to be sequenced it was decided to perform a pre-screening in order to limit the analysis of uninformative sequences. Polymorphisms in the sequence of the expressed genes can be at the level of the restriction site sequence recognized by the enzymes used in the procedure or at the nucleotide complementary to the selective 3' end of the primers employed. Differences between genotypes lead to the generation of cDNA-AFLP fragments of different size or to PCR amplification of homologous cDNA fragment with different primer combinations. This determines distinctive expression patterns though the gene is expressed in each genotype. To filter out such cases the differential screening was performed through hybridisation of a filter array prepared with the PCR products obtained from the isolated cDNA-AFLP bands. cDNA fragments differing in length or generated from another specific primer combination would give a hybridisation signal in presence of the complementary sequences

independently from which pattern these were obtained. In this way, it would be possible to individuate products present only in the two resistant genotypes 4551 and H0909. This first screening was performed in order to assess in a fast way the possibility to get more information from the cDNA-AFLP fragments selected from the comparison of the expression pattern. From the analysis, 6 spots common to 4551 and H0909 did not show hybridisation when using cDNA-AFLP template amplified from M9 as probe. The cloning of the corresponding bands produced a series of sequences that were subsequently tested through PCR for their presence in the genome of all the three genotypes. In only one case the specific primer pair Kin showed differences between the genotypes in the capability to amplify a genomic region. The primer designed on the other sequences showed a PCR product in all the genotypes tested and after sequence comparison differences between genotypes were found only at few nucleotide positions. These were further analysed for the development of SNPs. In none of the cDNA-AFLP fragments associated to the six spots individuated from the differential screening, sequences present only in the resistant genotypes were found. The fact that the probe produced from M9 did not hybridise to those spots could be due to a lower amount of product present in the probe because of a lower expression of the corresponding gene in this genotype. Nevertheless, differences were individuated at SNP level that could be used in the analysis of the crossing populations, an important information source for the breeding program aimed to develop an AP-resistant agronomic valuable rootstock.

One disadvantage of the array produced with spots containing more than one PCR product is that a hybridisation signal would be obtained even if only one product hybridised to the probe hiding possible differences related to the other fragments. Therefore, to gain more information on sequences derived from *Malus* genotypes and for further developing SNPs related with expressed sequences a deeper investigation was performed on the bands derived from the expression patterns analysis. A similar experiment was repeated this time using a high-density array produced with spots containing only one PCR product corresponding to a cloned cDNA-AFLP fragment. This approach was validated in a hybridisation test using the DIG-labelled PCR probe produced from the previously identifed Kin fragment. Spots associated to clones obtained from cDNA-AFLP bands derived from primer combination different to that corresponding to the Kin band were also individuated (Figure 14). This result showed that with this method it was possible to screen out fragments associated to different cDNA-AFLP bands but deriving from homologous gene products. Nevertheless, sequences associated to genes belonging to the same family or sharing domains with high similarity but involved in different pathways would be lost though they could play a role in the resistant genotypes. However, because we wanted to focuse on those products characterising the resistant plants, for each cDNA-AFLP template obtained from the different tissues and genotypes, hybridising probes were produced using all the specific primer combinations available. In this way, it was possible to better define differences related to the selective primers and to the specific tissue, root or sprout. The sequences of the clones individuated in the screening were

grouped in 23 clusters on which specific primers were developed and subsequently tested on the cDNA and the genomic DNA of 4551, H0909, M9, D2212 and *M. sieboldii*. All the primers showed a PCR product from the cDNA of each genotype while some primers did not work using genomic DNA as template. However, the products associated to a primer pair showed homologous sequences in each genotype. Nevertheless, from this analysis informative data could be further obtained from the polymorphisms associated to the restriction site and the selective primer that generated the cDNA-AFLP band. In addition, the sequences derived from clones individuated in the screening were employed to carry out the analysis for the SNP marker development. However, in one case it was possible to individuate through PCR a primer pair (C8) that showed polymorphism between *M. sieboldii*-derived and *M*. x *domestica* genotypes and also between different *M. sieboldii*-derived genotypes. These data are only a preliminary set of information that will be joined to the data derived from the analysis carried out by Bisognin in her doctoral thesis (part of the main project) using microsatellites markers on the same genotypes. (Bisognin *et al*., 2009) In addition, the data obtained from the hybridisation experiment could be source for a further analysis based on the selection of spots common in the H0909 and M9 but not in 4551 in order to individuate differences related to the *M*. x *domestica* genome. In the future, such markers could be associated to particular genomic regions linked with expressed genes.

The method used for the individuation of constitutive differences at a gene expression level led to the isolation of several sequences. About the half of these showed similarities with proteins of known function (Table 17) while the remaining sequences were similar to expressed genes or proteins of unknown function or showed no similarities with sequences present in databases. Interestingly, the genes with similarity to a 14-3-3-like protein and a high mobility group protein 1 (HMG1) identified in the first hybridization experiment were found also in the high-density macroarray indicating a reproducibility in the method used for the selection of differences at a nucleotide level. Nevertheless, for the products with known function, it was not possible to hypothesise a direct involvement in the phytoplasma-plant interaction based on their homology. However, 14-3-3-like protein, ankyrin repeat-containing protein and ser/threo protein kinase have ubiquitous functions also involved in signalling and plant defense (DeLille *et al.*, 2001; Sehnke *et al.*, 2002; Becerra *et al.*, 2004; Afzal *et al.*, 2008). Therefore, considering the complexity of this system it cannot be excluded that their expression could be influenced after infection as this cannot be excluded for the sequences with unknown putative function. However, further investigations are necessary in order to characterise their expression profiles in presence of the pathogen. In addition, the original material coming from field-grown plants was not an optimal source because of the non-standardised conditions at which the plants were grown. This might differently influence the physiology of the *M. sieboldii*-derived hybrids and the *M.* x *domestica* genotypes and eventually their expression profiles. However, several data are now available for the development of markers, a further important goal in the main project for the characterisation of the resistant genotypes.

5.2.2 Analysis of the putative markers individuated in the cDNA-AFLP analysis

The necessity to develop a resistant rootstock with good agronomic characteristics was one of the main goals of the SMAP project. Because the resistance was associated to a wild genotype with scarce agronomical traits the production of a new crossing carrying both the resistant and the agronomical traits was required. Several crosses were made in the SMAP project as reported in Bisognin *et al*., 2009, and this breeding activity is still ongoing. The screening of the breeding progeny for AP-resistance is done by graft-inoculation of the progeny seedling and requires several years of field observation. In order to speed up this screening process the research activity was focused on the development of molecular markers associated to the resistance. In bulk segregant analysis (BSA) molecular markers showing polymorphism between the parents of the population can be associated to a quantitative trait loci (QTL) like the resistance character (Lin *et al.*, 2004; Stankiewicz-Kosyl *et al.*, 2005; Lin *et al.*, 2006). The segregation analysis of the QTLs can then allow the selection of those genotypes that carry the desired traits (marker assisted selection; MAS). The association between molecular markers and QTLs requires the availability of a genetic map where the markers are distributed and are covering the whole genome allowing the definition of association groups corresponding to loci, regions or chromosomes. At the beginning of the SMAP project there was no availability of a genetic map of the *M. sieboldii* parent hindering the possibility to select interesting seedlings through BSA. Therefore, initially the positioning of SSR markers was a prerequisite for the establishment of a first genetic map. In a second time, the map should be improved with the positioning of other markers associated to genes with known function but more in general with SNPs derived from expressed sequences. This last aim was part of this thesis where SNPs associated to cDNA-AFLP fragments derived from constitutively expressed genes and from differentially expressed genes had to be developed. In the two cDNA-AFLP experiments a series of sequences were produced deriving from resistant and susceptible genotypes. On these sequences nine SNPs that showed a polymorphic character were obtained while other 155 putative SNPs were individuated. In addition, the differential screening on the constitutively expressed genes isolated in the first cDNA-AFLP allowed the individuation of two primer pairs (Kin and C8) that resulted to be polymorphic. Both showed temperature-dependent differences in the capability to amplify genomic regions in different genotypes. In order to characterise their potential as PCR markers, further analyses on the sequences associated to the targeted genomic regions revealed that the polymorphisms in both markers were associated to a SNP at the level of the annealing site of one of the two primers. Stringent conditions in the PCR reaction showed a PCR product only in *M. sieboldii* or *M. sieboldii*-derived genotypes while in *M. x domestica* the products were obtained only lowering the primer annealing temperature. In addition, another putative marker was individuated from the cDNA-AFLP analysis of differentially expressed genes. In this case, primers pairs designed for the real-time analysis on the internal region

corresponding to the gene 30 were not able to produce a PCR product either at RNA nor at genomic DNA level in the *M. x domestica* genotypes tested. Furthermore, primers developed on the cDNA-AFLP fragment associated to the gene 30 that were targeting regions up- and downstream of those associated to the real-time PCR primers failed to amplify this gene and also the corresponding genomic region in the susceptible genotypes. Further studies are necessary to verify if the gene 30 is missing in *M. x domestica* genotypes.

The power of K, C8 and 30 as markers was tested in two different crossing populations. The distribution of the markers in the different seedlings and their combinations indicated that there was free recombination between them (Figure 59). Nybom *et. al.* (2006), reported that for polyploid *Rosaceae* highly homologous genomes can be inherited preferentially in a pairwise manner. In this thesis from the results obtained with the markers K, C8 and 30 there are no indication for the occurrence of preferential pairing between individual members of the polyploid genome of *M. sieboldii* supporting also the observations made by Bisognin (unpublished data) using SSR markers.

5.2.3 Analysis of differentially expressed genes after phytoplasma-infection

5.2.3.1 *In vitro* system

A standardised material produced from plants grown in controlled conditions was necessary for the analysis of the differentially expressed genes after phytoplasma-infection. Therefore, it was decided to exploit the power of the *in vitro* technique for the further gene expression analysis. In addition, the traditional screening for the isolation of interesting traits in woody plants takes a considerable time, is labour-intensive and costly. In order to speed up the resistance screening in the *M. sieboldii*-derived hybrids and their progeny, an *in vitro* system was established with several *Malus* genotypes including those used in this thesis. *In vitro* grafting technique, also known as micrografting, was successfully employed in different woody plant species (Murashige T. *et al.*, 1972; Alskieff and Villermur, 1978) and efficient transmission of phytoplasmas through micrografting was described by Jarausch *et al.* (1999a). The possibility of having standardised material in association with an efficient infection process under controlled conditions was an optimal starting point for the analysis of each genotype employed in this thesis.

The *Malus* genotypes selected for the *in vitro* screening were the resistant genotypes *M. sieboldii,* H0909, D2212 and the susceptibles *M.* x *domestica* cvs. M9 and Golden Delicious. *M. sieboldii* was chosen as it was the original parental genotype carrying the resistance trait and used in the breeding program. The *M. sieboldii*-derived hybrids H0909 and D2212 were selected because of

their genetic pedigree and their different degree of resistance against AP. The susceptible *M.* x *domestica* M9 was chosen being one of the most diffused genotype employed as a rootstock whereas GD because it is a very important cultivar that is highly susceptible to AP. Although in this thesis the *in vitro* graftings on the selected cultures were performed mainly for the preparation of the infected material, interesting observations were derived from the analysis of the infection trials. The data collected (Table 19) showed that the percentage of successfull graftings (recorded as strong grafts) using infected GD as scion was only slightly higher in the *M.* x *domestica* genotypes compared to the *M. sieboldii*-derived ones. As reported in Bisognin *et al.* (2008a) this was an important evidence that *M.* x *domestica* and *M. sieboldii*-derived genotypes were cell compatible, an indispensable characteristic for the establishment of this system. Three months after graft-inoculation the pathogen was found to be efficiently transmitted in GD, M9 and D2212 while in *M. sieboldii* and in H0909 the transmission rate was very low. Over a longer period of observation, it was interesting to notice that for GD the inoculated plants remained infected for the entire period while for M9 after an initial reduction the number of plants that were still infected after one year remained unchanged. Differently for the apomictic genotypes none of the inoculated *M. sieboldii* and H0909 plants survived one year after the grafting while the number of D2212 infected plants decreased during the year of observation. The data obtained from the infection trial were joined to those generated in a similar experiment in the frame of the SMAP project. In the latter the '*Ca*. P. mali' AT-2 strain was employed for the investigations and the results were published together in Bisognin *et. al.*(2008a). However, the data available were not sufficient to analyse if the mortality was related to the '*Ca.* P. mali' AP strain and because several experiments were required to assess this question no further investigation were done in this thesis. Nevertheless, the most important aspect of our *in vitro* system was shown in a study complementary to this thesis (part of the same main project) where it was demonstrated that plants grown and infected in these conditions behave in a highly reproducible way (Bisognin *et al.*, 2008a). Moreover, analyses on the symptoms and the phytoplasma titre in infected plants showed also reproducible and reliable results. The resistant genotypes were successfully characterised in the *in vitro* system showing low phytoplasma titre and absence of symptoms similarly to that observed in *in vivo*-grown plants (Bisognin *et al.*, 2008b). Moreover, their behaviour after the phytoplasma infection could be monitored during a large period of time confirming the power of this method for the speed up of the '*Ca*. P. mali'-apple interaction analysis (Bisognin *et al*., 2008a). The demonstration that the *in vitro* system was a suitable platform for the analysis of the resistance against AP was crucial. This allowed the exploitation of this technique as a tool for the preparation of standardised healthy and '*Ca*. P. mali'-infected material (through micrografting) for its use in the analysis of the differentially expressed genes. The infection trials in this thesis were conducted using the AP subtype of '*Ca*. P. mali' but no infected plants were available for *M. sieboldii* and H0909. Nevertheless, contemporary to this study a series of graftings was performed by another group involved in the same main

project using the AT2 subtype of '*Ca.* P. mali' as inoculum where it was possible to obtain infected *M. sieboldii* and H0909 plants. These plants were used as material for the selection of differentially expressed genes in the phytoplasma-plant interaction. Seemüller and Schneider (2007) showed that there is a considerable genetic heterogeneity among '*Ca*. P. mali' strains. However, in their study AP and AT2 were classified in the same virulence group as strains inducing severe symptoms in plants grown *in vivo*. In addition, the suitability of this strain for the *in vitro* experiments was also shown by the data reported in Bisognin et al (2008a). However, because the relationship between the phytoplasma and its host appears to be a very complex system, it was reasonable to think that a relatively high number of expressed genes, not necessarily related to the resistance mechanism, would be differentially expressed.

5.2.3.2 Phytoplasma-plant interaction: a "black box"

In the past two decades, several studies based on different approaches were conducted for the investigation of the modifications that take place in the host in the phytoplasma-plant interaction (Kartte and Seemüller, 1991; Musetti *et al.*, 2000; Jagoueix-Eveillard *et al.*, 2001; Bertamini *et al.*, 2002; Carginale *et al*., 2004; Christensen *et al*., 2005). Nevertheless, still very little is known on this system, therefore, it can be considered as a black-box model because of the poor availability of information. Thus, an approach that allows obtaining more knowledge on this systems is required.
To investigate whether the resistance was associated to an induced plant defence-response to '*Ca*. P. mali' infection rather than to a constitutive character associated to physiological features, the cDNA-AFLP procedure was employed for the analysis of differences in the gene expression of resistant and susceptible plants. Since the time when this technique was improved to avoid the possibility of several transcript derived fragments (TDFs) arising from a single gene/cDNA (Breyne *et al.*, 2003) the cDNA-AFLP was exploited in several works. The technique was successfully employed for studying genes in systems where it is difficult (or not possible) to culture one organism on an artificial medium, which limits the use of the traditional genetic and molecular techniques (Wang *et al.*, 2009). Robustness, high-throughput and applicability in genome-wide expression analysis for gene discovery (where prior knowledge of sequences is not required or available) (Breyne and Zabeau, 2001; Breyne *et al*., 2002), rendered cDNA-AFLP particularly useful for the investigation of complex systems like fungi-plant interaction (Gabriels *et al.*, 2006; Polesani *et al.*, 2008; Wang *et al.*, 2010). Similarly, as the phytoplasmas-plant system is characterised by poor knowledge on the mechanisms involved during the infection and by high complexity in the type of interaction, the cDNA-AFLP should give the possibility to investigate the modification in the expression patterns in '*Ca*. P. mali'-infected apple plants. In the past, only two studies were performed focused on the individuation of differentially expressed genes associated to the phytoplasma-plant interaction. In both, similar approaches were adopted using either the differential display (DD) (Carginale *et al.*, 2004) or the cDNA-AFLP technique (Aldaghi, 2009).

Their results showed that the employment of methods based on the generation of patterns associated to the expression profiles can be a good choice for the study of this plant-pathogen interaction.

5.2.3.3 Selection of differentially expressed genes

The analysis of the expressed genes after the infection with '*Ca*. P. mali' allowed the individuation of several bands that resulted differentially expressed. However, the high complexity that characterised each pattern complicated the individuation of those cDNA-AFLP fragments that showed an up-regulation or a down-regulation in their expression. In order to reduce the number of sequences to analyse only those patterns characterised by showing an over expression in the infected resistant H0909 and *M. sieboldii* genotypes were considered. 63 bands were isolated, each one associated with more than one single product. The relatively low number of combinations used and the nature of the selective base applied in the method produced complex patterns with several sequences associated to each band. Experiments carried out in other plants showed that employment of primers selecting on the three first positions on the cDNA fragments gives the possibility to obtain bands that could be almost directly sequenced after blot excision and re-amplification (prof. Delledonne M. personal communication). The choice of the best combination of primers and the more effective type of selection (1, 2 or 3 selective nucleotides) needs to be set up in an empirical way. However, the constantly growing availability of sequences in database allows the development of previsional models for the selection of the most informative enzyme and primers to be employed in the analysis (Kivioja *et al.*, 2005; Stolting *et al.*, 2009). Although in this thesis the method could be optimised using more primer combinations, it was decided to proceed in the analysis of the bands individuated in the screening. After cloning and sequencing of the cDNA fragments, homology analysis with sequences present in databases showed a very broad range of metabolisms associated to the putative function of the identified genes. Although they had to be confirmed as differentially expressed, a first general look on the main pathways involved evidenced genes which encode for proteins associated to degradation, transport and folding of proteins, to plant development and morphogenesis, to cell signalling and to stress/pathogens response. The picture resulting from the sequence analysis reflects deep changes in the general plant metabolism as observed in previous studies ((Kartte and Seemüller, 1991; Lepka *et al.*, 1999; Musetti *et al.*, 2000; Carginale *et al.*, 2004; Christensen *et al.*, 2005; Musetti *et al.*, 2005). Nevertheless, the involvement of the genes individuated in the differential expression analysis in the plant response after '*Ca*. P. mali' infection had to be validated through specific studies on their expression level. Because of the high number of genes to be analysed the work was focused on those sequences that:

- resulted to be present in both H0909 and *M. sieboldii* resistant genotypes

- showed homology to genes that are involved in response to diseases
- related with putative changes in the plant metabolisms supposed to be involved after phytoplasmas infection
- were present in several clones obtained from the same cDNA-AFLP band

These sequences were analysed with the real-time PCR technique to study and validate their differential expression in the plant response to '*Ca*. P. mali'.

5.3 Quantitative reverse-transcription PCR analysis

5.3.1 Choice of the detection method

The real-time PCR (quantitative PCR, qPCR; quantitative reverse-transcription PCR, qRT-PCR) analysis is a powerful technique developed in the mid-1990s (Livak *et al.*, 1995; Gibson *et al.*, 1996; Heid *et al.*, 1996) that allow the determination of the initial concentration of a specific product through PCR. The detection method is based on the proportionality of a fluorescence signal to the quantity of product that is synthesized during the PCR allowing the monitoring of the reaction kinetic. Different approaches are available for the generation of the fluorescence signal. The most used are TaqMan probes (Higuchi *et al.*, 1992; Higuchi *et al.*, 1993; Livak *et al.*, 1995), Molecular Beacons (Tyagi and Kramer, 1996; Ortiz *et al.*, 1998), Scorpions primers (Whitcombe *et al.*, 1999) and SYBR Green I (Wittwer *et al.*, 1997; Woo *et al.*, 1998; Cane *et al.*, 1999). The first two requires the development of specific oligonucleotide probes that hybridise to the target giving a signal proportional to the quantity of probe bound (Molecular Beacons) or degraded (TaqMan). Scorpions are bi-functional molecules containing a PCR primer covalently linked to a probe that carryes a fluorophore interacting with a quencher which reduces fluorescence (Whitcombe *et al.*, 1999). During a PCR reaction, fluorophore and quencher are separated leading to an increase in light output from the reaction tube. In TaqMan probes, Molecular Beacons and Scorpion primers the coupling of a fluorogenic dye molecule and a quencher moeity to the same or different oligonucleotide substrates exploit the Förster Resonance Energy Transfer (FRET) for the generation of a fluorescence signal (Didenko, 2001). Differently, SYBR Green I binds to the double strand helix of the DNA giving a signal proportional to the quantity bound. TaqMan, Molecular Beacons, Scorpions primers are highly specific for their target and produce a high sensitive signal. They can also be used for multiplex analysis if different dyes are employed in association with the probe. The drawback is that in case several different targets have to be analysed the development of the probes is more complex, time consuming and expensive. Only recently, a Universal Probelibrary (Roche Applied Science) was developed for the study of multiple genes at the same time giving the advantage of high specificity and sensitivity but this method requires a well known

database of expressed genes which was not available for apple plants. In comparison, SYBR Green I is cheaper and can be used for the analysis of several genes without the need to develop specific probes. Its drawback is that it cannot be employed for multiplex detection and the presence of unspecific products in the reaction will give an additive signal influencing negatively the analysis. Nevertheless, this problem can be avoided using primers with high specificity for the annealing sites on the target. Therefore, this method was chosen for the study of the gene expression because the easiness to develop the procedure and the possibility to select primers which enabled the analysis of several genes at the same time with the same cycling conditions.

The necessity to investigate many targets in different genotypes and health status conditions required to adopt a method that could give fast results in a reproducible and reliable way. Therefore, a procedure in which the RNA extracted from the plants was directly used as template for the real-time PCR screening was optimised for our system. RNA of good quality is necessary for the reliability of the data produced (Fleige *et al.*, 2006) and the method developed for the RNA extraction from our plant materials was suitable for fulfilling this condition (Moser *et al.*, 2004).

The real-time PCR can be performed on cDNA or directly on total RNA (qRT-PCR). While the cDNA can be better conserved and is less prone to hydrolysis and degradation, the RNA has to be handled carefully. Moreover, the main disadvantage of the qRT-PCR method is that for each target a separated priming reaction is required hindering the possibility to return to the same preparation and amplify other targets at a later stage. It is also wasteful if only limited amounts of RNA are available (Bustin *et al.*, 2005). While it is possible to amplify more than one target in a single reaction tube (multiplex) (Wittwer *et al.*, 2001), this requires careful experimental design, optimisation of reaction conditions and excludes the use of SYBR Green I-based detection method. The advantage using directly RNA for the analysis is that the analysis can be performed on regions of the gene that are specifically retrotranscribed. Oligo dT-based cDNA synthesis can be characterised by bias in the 3' region to 5' region ratio (Wang *et al.*, 2000; Bustin, 2002; Bustin *et al.*, 2005; Singh *et al.*, 2005) if not fully completed affecting analysis having the 5' region as target. Although random priming can be used as method for the cDNA synthesis in case of mRNA targets present at low levels, the priming may not be proportional and hence its subsequent amplification not quantitative (Zhang and Byrne, 1999; Bustin *et al.*, 2005). In addition, random primed reactions are characterised by a narrow range of linearity compared to reactions using target-specific primers (Bustin and Nolan, 2004). Finally, although the RNA should be of very good quality a further advantage of the qRT-PCR on total RNA is that reliable data can be obtained also from samples with a certain grade of degradation if small regions and relative small amplicons are generated in the the analysis (Fleige and Pfaffl, 2006).

5.3.2 Real-time PCR data analysis

The sensitivity of this method characterise this powerful technique but the huge amount of data obtained through the real-time PCR has to be analysed in a way that allows producing reliable results. Several works were done with the purpose to establish a guideline for the data analysis. The different methods used have their advantages (easiness in the calculation) or disadvantages (results that over or under estimate the differences between the samples). Therefore, in this thesis in order to obtain more reliable data a method was developed modifying some aspects of the analysis presented in two published works: the DART-PCR method of Peirson *et.al* (2003) and the LinRegPCR method of Ramakers *et al.* (2003).

The choice of suitable internal controls for the investigated system is fundamental but it is not sufficient alone to allow for a reliable data analysis. Parameters like the determination of the reaction efficiency (E) and the selection of the range in which the PCR has a linear behaviour are important in order to be able to compare the different samples and states. Based on the linearity range of the reaction a threshold has to be set for the determination of the cycle at which the fluorescence in the sample reaches this value. This cycle is known as Ct and is a further important parameter for the comparison analysis of the samples. The DART-PCR and LinRegPCR methods calculate these parameters in different ways, therefore, in this thesis they were adopted as reference but modified to improve the data analysis. In the first method the calculation of the noise as described by the authors was not applicable in all of our curves and we needed a more defined way to determine the best linear range of each curve. In the case of the LinRegPCR method, in which an algorithm calculates the best linear regression that should fit with the linear phase of the reaction, we noticed that in several of our curves the range of linearity was wrongly determined. Recently, an improved LinRegPCR method was published supporting our observations (Ruijter *et al.*, 2009). Therefore, in this thesis a modified method was developed in which for each curve its derivative was calculated for the individuation of the range at which the fluorescence curve assumes the maximum slope. Subsequently, based on this range the threshold value was defined for each curve at the point where the derivative reaches its maximum. In the original LinRegPCR the efficiency is calculated based on the linear regression of the 3-6 fluorescence points that show the best correlation coefficient. We adopted this concept but the calculation of the linear regression with the best correlation coefficient was applied to the fluorescence values of the points generated from the instrument ranging around the fluorescence value corresponding to the threshold. In order to speed up the analysis and to facilitate the elaboration of the data a Visual basic® (Excel® module) algorithm was developed for the fast individuation of these parameters through the analysis of the raw fluorescence curves obtained in the real time PCR.

5.3.3 Evaluation of housekeeping genes as suitable controls in the '*Ca*. P. mali'-apple interaction

The need of internal controls for the data normalisation and the lack of information on the gene expression in apple plants was a significant problematic issue for the application of the real-time analysis. In the last years several works were published focused on the development of stable internal control to adopt in different plant systems (Czechowski *et al.*, 2005; Nicot et al., 2005; Exposito-Rodriguez *et al.*, 2008; Gutierrez *et al.*, 2008; Hu *et al.*, 2009). The most commonly used considered housekeeping genes like the ribosomal gene 18S rRNA (18S), the eukaryotic elongation factor 1 α (EF), the eukaryotic translation initiation factor 4A (eIF4), β tubulin (TUB), α-tubulin (TUA), ubiquitin and polyubiquitin genes (UBQ), actin (ACT), cyclophylin (CYP), glyceraldehydes-3-phosphate dehydrogenase (GAPDH). However, depending on which system, tissue, developmental stage is investigated, these genes, commonly considered to have a constant expression level, resulted to be differently stable and not universally suitable to be used as reference (Nicot *et al.*, 2005; Exposito-Rodriguez *et al.,* 2008; Gutierrez *et al.*, 2008; Hu *et al.*, 2009; Wan *et al.*, 2010). Therefore, a preliminary experiment was performed on our system in order to select between the housekeeping genes commonly employed as normalisers those which could be used as internal control. In several studies the eukaryotic translation initiation factor 4A (eIF4) and the elongation factor 1-α (EF) were showing the most stable expression level in different plant systems (Nicot *et al.*, 2005; Gutierrez *et al.*, 2008; Hu *et al.*, 2009; Wan *et al.*, 2010). Thus, in this work they were chosen as control candidates. Moreover, being involved in different, independently regulated cellular functions the possibility that they are co-regulated can be excluded, a condition necessary for the choice and comparison of different controls for the real-time analysis (Vandesompele *et al.*, 2002; Gutierrez et al., 2008). The further genes whose stability was assessed were GAPDH and α-tubulin (TUA). In addition, in the analysis carried out directly on RNA the ribosomal gene 18S rRNA (18S) was also selected as reference. Equal quantities of RNA were used in the qRT-PCR analyses and no differences in the 18S expression level were observed between the healthy and infected state. The analysis of the housekeeping genes expression level on RNA of *in vitro* healthy and infected plants and of leaf and root tissues of *ex vitro* healthy and infected plants was carried out using the 18S rRNA as normaliser. GAPDH and TUA resulted to be significantly differentially expressed in *in vitro* plants and in root tissues (Figure 25). In addition, GAPDH showed differential expression also in *in vitro* tissue whereas for eIF4α no variations in the three tissues were found. Although because of their low quality it was not possible to test the behaviour of the EF primer pair on the RNA a test on the cDNA using eIF4 as control showed that the expression levels of this gene remained unchanged after the '*Ca*. P. mali'-infection. Therefore, because for the *in vitro* plants the real-time analysis was carried out on cDNA, the expression data of the different genes were normalised with eIF4α and EF. Since the cDNA was produced using the oligo dT-based retrotranscription (RT) the use of the 18S rRNA as further control was

excluded. This method was chosen in order to have the most comparable data with the cDNA-AFLP analysis. In addition, the employment of random priming method for the production of cDNA (to allow the use of the 18S rRNA as control) was avoided to reduce the possibility to introduce biases due to a narrow linearity in the RT reaction (Bustin *et al.*, 2005). Moreover, the use of 18S rRNA as a control delivers reliable data but in some cases can produce narrow differences in the relative expression level compared to those obtained from genes with similar expression level to that of the target genes (Nicot *et al.*, 2005). However, in the case of *ex vitro* plants it was not possible to use the EF as internal control because of some technical problems related to the primer efficiency. The 18S rRNA and the eIF4α were therefore used for the data normalisation.

5.3.4 qRT-PCR gene expression analysis in *in vitro* and *ex vitro* plants

Although the *in vitro* system employed in this work is a very powerful tool for the screening of several genes involved in the '*Ca*. P.mali'-apple interaction a drawback of the system is the impossibility to analyse the gene expression in the root organ because not developed in our *in vitro* conditions. Additionally, the gene expression with environmental parameters like those that affect in field-grown plants can influence how the plants react against the pathogen. Bisognin *et al*. (2008a) showed that the resistant genotype can be characterised *in vitro* suggesting that the resistance can be efficiently studied in this system making it also suitable for our purposes. Nevertheless, to test the genes individuated in the screening in plants that were grown in conditions more similar to those present in field, we performed the same real-time analysis on healthy and infected plants derived from *in vitro* cultures and transferred to the greenhouse, termed as *ex vitro* plants. These were first maintained and infected through micrografting and then after induction of root formation were transferred from the in *vitro* condition in pots and further cultivated in greenhouse. In this way, we could analyse also the gene expression in roots but more in general the gene expression in a fully developed plant with environmental parameters more similar to those present in field. The greenhouse conditions ensured a better homogeneity of the material than field conditions.

5.4 Gene expression analysis of the '*Ca*. P. mali'-plant interaction in susceptible and resistant *Malus* genotypes

5.4.1 Real-time PCR data analysis of the cDNA-AFLP-derived differentially expressed genes

The real-time PCR analysis was performed to verify if the genes identified in the cDNA-AFLP analysis are really differentially expressed among healthy and infected status of a plant and, thus, if they are involved in the '*Ca.* P. mali'-plant interaction. The reliability of the obtained data could be evaluated in different ways. First, in real-time PCR analysis for each gene two different internal

controls were used as normalisers. If the same trend of gene expression was seen with both datasets, the real-time PCR data were judged reliable. Second, different tissue samples were analysed: the real-time PCR analysis was performed on freshly extracted *in vitro* material from the same healthy and infected genotypes used in the cDNA-AFLP. This can be considered as a biological replicate to the cDNA-AFLP analysis. In addition, gene expression data for a subset of genes was obtained from different organs of healthy and infected *ex vitro* plants. In this way, the behaviour of a gene was compared between plants maintained in different conditions in order to find out for which gene the expression level is influenced by the phytoplasma presence. A total of 20 genes were investigated and all of them were tested in *in vitro* tissues while only the half (gene 3, 5, 6, 7, 9, 11, 14, 15, 21, 22) was also analysed in *ex vitro* tissues. Further controls were applied to assess the reliability of the method employed for the individuation and the analysis of the differentially expressed candidate genes. In the cDNA-AFLP analysis the gene 24 was found to be overexpressed in all the infected *in vitro* genotypes. Therefore, it was chosen as a further control allowing the investigation of the gene expression in all the genotypes. In addition, the genes 25 and 35 were used as internal reference for the evaluation of the real-time analysis and procedure. Both genes were similar to a putative squalene monooxygenase and the corresponding cDNA-AFLP bands were individuated in two independent electrophoretic runs from the same specific primer combination. The purpose of this procedure was to evaluate if the cDNA-AFLP analysis was able to give reproducible data. Therefore, initially the cDNA-AFLP expression pattern of a primer combination (TCC) was characterised in a first run selecting and excising the differentially expressed bands. One of these generated the gene 25. Subsequently, in a new run the same bands were again individuated and excised from the gel. The band corresponding to the position from which the gene 25 originated was cloned and sequenced. The sequences were homologs and the new band was called gene 35 and analysed as an independent gene. In this case, the aim was to use it as internal control to test the reproducibility of our real-time system. Different primers as those for the gene 25 were designed for the real-time PCR experiment. The gene 25 is discussed more extensively later on. Here is noteworthy to mention that in *in vitro* plants the **gene 35** resulted to be over-expressed in MS AT2 and under-expressed in H0909 AT2. In GD AP (EF-normalised data) and M9 AP (eIF4-normalised data) it showed also down-regulation. Compared to the results obtained from the analysis of the gene 25 the same tendency was observed for MS and H0909 but not for M9 and GD. However, the general expression level in the plants analysed was comparable for gene 25 and gene 35. This result confirms the importance to have more than one internal control in order to increase the reliability of the data analysed through the method chosen. This was particular evident for the gene expression analysis in D2212 roots. High differences were seen for D2212 root between the data produced from the two internal controls. Although the relative expression between healthy and infected status were very similar in the two normalised datasets showing for each gene the same expression trend, in case of the 18S derived data they

were affected from higher variability causing the differences not to be statistical significant as it was in the case of eIF4α. Thus, 18S was not a suitable normaliser in D2212 roots although the reason for this is not clear.

In several cases, differences in the gene behaviour between tissues and genotypes were observed and in some cases the differential expression trend, characterising the cDNA-AFLP patterns of the genotypes, was not maintained in the qPCR analysis. This raised the question if the sequences chosen correspond to those that produced the signal in the cDNA-AFLP analysis. In addition, the *in vitro* plants used in the qRT-PCR derived from a new infection experiment which might have caused differences observed in the behaviour of the genes isolated from the cDNA-AFLP analysis in the first trial. Actually, the meaning to carry out the analysis on a new set of infected plants was to try to screen out false positives not related with the plant reaction to the phytoplasma infection. However, additional biological repetitions should be performed to define if those genes are indeed involved in the response to the pathogen. The qRT-PCR analysis on *ex vitro* plant tissues was performed with this purpose in order to gain more information for the individuation of such genes. This could give also the advantage to analyse the gene expression in plants that are grown in conditions more similar to those that can be found in the orchards.

From the real-time analysis, 12 genes resulted to be significantly differentially expressed in at least one genotype either *in vitro* plants or in *ex vitro* plants. Nevertheless, in few cases a tendency to over- or under-expression was observed though not supported by the statistical analysis due to the high variability associated to the real-time datasets. In *M. sieboldii* and H0909 *in vitro* plants, 30 and 16 % of the genes, respectively, confirmed the cDNA-AFLP analysis resulting to be over-expressed. Interestingly, in leaf tissue of H0909 *ex vitro* plants 67 % of the genes were up-regulated confirming the tendency observed in the cDNA-AFLP analysis. In this case, since the selection of the cDNA-AFLP differentially expressed bands was performed on green tissues only, the real-time data obtained from leaves were comparable. Overall, about 20% of the genes confirmed their cDNA-AFLP expression pattern in *M. sieboldii* or H0909 in either *in vitro* or *ex vitro* plants. On the contrary, 30 % of the genes showed down-regulation or were not differentially expressed. For the remaining 50 % of the genes a contrasting behaviour was observed.

5.4.1.1 Analysis of the genes with unknown function

Among the selected cDNA-AFLP fragments the genes 3, 14, 21, 22, 28, 30, 31 and 32 showed similarities with sequences present in database with an unknown function.

The **gene 3** showed no differential expression in tissues derived from infected *in vitro* plants. The data of the cDNA-AFLP were not confirmed by the real-time PCR analysis. On the contrary, it resulted to be up-regulated in root tissues of infected GD and M9 *ex vitro* plants. This result indicates a differential expression in roots of *M.* x *domestica* genotypes. Further investigations

need to be carried out to confirm the role of this gene in the response of susceptible genotypes to the phytoplasma infection.

The **gene 14** showed an expression pattern which could be attributed to differences between *M.* x *domestica* and *M. sieboldii*. Only in MS this gene was under-expressed in *in vitro* plants and not differentially expressed in *ex vitro* plants while in genotypes with a genomic contribution of *M.* x *domestica* it showed rather a tendency to be over-expressed. In *ex vitro* plants it showed over-expression in GD, M9 root tissue and D2212 leaf tissue (eIF4-normalised dataset), H0909 leaf-tissue (both controls) and H0909 root (18S-normalised dataset). In *in vitro* plants a strong over-expression was observed only in the GD genotype infected with the AP strain while no differential expression was seen in GD infected with strain AT-2. This result indicates a differential response of the plant to the different phytoplasma strains. Indeed, significant differences in phytoplasma concentration between the AP strain (also named PM4) and the AT-2 strain (named PM6) have already reported for these *in vitro* cultures by Bisognin *et al.* (2008a). The AP strain was multiplied to higher concentrations in the *in vitro* plants which might be reflected by the differential gene expression of gene 14. It would therefore be important to elucidate the function of gene 14 to interpret this result.

Contradictory results were obtained for the **gene 21**. Noteworthy is the expression in H0909 *in vitro* plants where this gene shows an up-regulation of about 2 fold after phytoplasma infection. In *ex vitro* plants the gene 21 resulted up-regulated in both the normalised datasets in M9 root as well as in leaves of H0909 (18S-normalised dataset) while it was under-expressed in roots of *M. sieboldii* and D2212 *ex vitro* plants. The tendency of over-expression in green tissues (*in vitro* plants, *ex vitro* leaves) of H0909 and D2212 indicate a confirmation of the cDNA-AFLP result in the real-time PCR analysis. A specific correlation of this gene to the phytoplasma infection needs to be confirmed.

The analysis on the *in vitro* plants was not possible for the **gene 22** due to the low quality of the fluorescence curves characterised by high background noise and by the formation of unspecific products. Moreover, also the data obtained for *ex vitro* plants exhibited a high variability which influenced the statistical analysis. Except for GD roots this gene showed a tendency to over-expression in *ex vitro* plants which was significant in MS root (eIF4-dataset) and in D2212 leaf tissue (both controls). As the data for susceptible *M.* x *domestica* genotypes GD and M9 were contradictory and statistically not significant, the real-time PCR analysis confirms the cDNA-AFLP data. This gene showed homology with a pepsin/ASP-protease domain at the amino acid sequence level. However, at the nucleotidic sequence level homology with a product of unknown function was found. Since conserved protease domains can be found in proteins involved in considerably different pathways and functions, the gene 22 was included in the group of genes with unknown function. From the qPCR analysis, this gene might be related to AP-resistance and it merits therefore further investigation.

Interesting observations can be derived for the genes 3, 14, 21, 22 from the comparison of the general expression level between the *ex vitro* genotypes in the different tissues. GD showed always the lowest expression level whereas in MS, H0909 and D2212 a comparable expression level was observed. Only for the gene 3 the expression level of H0909 is similar to that of GD. Differently, M9 showed about 4 times higher expression compared to GD and about 2 times higher expression compared to the other genotypes. This indicates fundamental differences in gene expression among the two cultivars of *M.* x *domestica* which further complicate the interpretation of the gene expression data between susceptible and resistant genotypes.

The genes 28, 30, 31, 32 were analysed only in *in vitro* plants. The **gene 28** was one of the most interesting among those that resulted to be differentially expressed. In GD infected *in vitro* plants the expression was about 4 fold higher compared to the healthy state while in MS the gene was about 2 fold over-expressed. Interestingly, in H0909 a down-regulation of about 2 fold was observed. From these results the gene 28 can be considered a candidate involved in the plant response after phytoplasma infection but more deeply investigation are necessary to understand if it is part of a general response or if it has a more specific function in the pathogen-plant interaction. Regarding the general expression level it is interesting to observe that the *M.* x *domestica* genotypes showed a higher general expression level compared to MS, H0909 and D2212.

The **gene 30** was a particular case because it was possible to study its expression only in the resistant genotypes showing under-expression only in MS AT2 with eIF4-normalised data. In GD and M9 no PCR product was obtained using this primer pair. Moreover, the general expression level resulted to be almost 6 times higher in MS and D2212 compared to that of H0909. It is noteworthy to mention that this gene derived by bands found differentially expressed only in H0909 and *M. sieboldii* (Table 21; TGC1 pattern 'short-run') and therefore isolated only from these two genotypes.

The **gene 31** showed differential expression in GD where it resulted to be about 3 fold over-expressed and H0909 where it showed an about 3 fold down-regulation. No differential expression was found in MS. Interestingly, the general expression level of M9 was lower compared to GD healthy and it reached in H0909 after infection the level of M9. This might indicate the influence of the allele of gene 31 deriving from M9 in the complex hybrid H0909 on its gene expression. On the contrary, in GD after infection and in healthy and infected state in MS and D2212 the expression level resulted to be 2-3 times higher than in GD healthy plants. This result can be interpreted in the way that the expression of gene 31 needs to be up-regulated after phytoplasma infection in GD whereas it has this high expression level constitutively in resistant genotypes. In this regard it is noteworthy to mention that in the *in vitro* AP-resistance screening performed by Bisognin *et al*. (2008a) on the same tissue cultures the genotypes H0909 was evaluated as less resistant than MS and D2212. Thus, although the real-time PCR data were not in agreement with the cDNA-AFLP data, gene 31 merits further investigation to elucidate its putative correlation to AP-resistance.

The last of the studied genes with unknown function was the **gene 32** resulting to be generally differentially expressed. Interesting is to observe that in GD and H0909 the gene was under-expressed while in MS it showed over-expression. H0909 shares part of its genome with the *M.* x *domestica* genotypes, therefore, it could show a similar behaviour as in GD for some genes. However, an over-expression of gene 32 was not only specific to *M. sieboldii.* An important difference became apparent between the expression patterns of this gene in GD infected with strain AP (under-expression) and strain AT-2 (over-expression). Thus, this gene showed a strain-specific differential expression in GD which is further supported by the general expression level which decreased after infection only in GD AP-infected plants. As discussed above, strain AT-2 showed lower multiplication efficiencies in the *in vitro* plants than strain AP (Bisognin *et al.*, 2008a). Further investigations might elucidate whether the product of gene 32 is involved in the reduced multiplication of strain AT-2 in GD as well as in *M. sieboldii*. Nevertheless, as an increasing number of reports (Seemüller and Schneider, 2007; Kube *et al.*, 2008; Schneider and Seemüller, 2009) demonstrate a high genetic variability among different '*Ca*. P. mali' strains, this result demonstrates for the first time that differences between strains can also be observed in the interaction with plant genes.

5.4.1.2 Analysis of the genes with known function

The genes 5, 6, 7, 9, 11, 15, 23, 24, 27, 29, 33 and 35 showed similarities with genes of known function. For the genes 5, 6, 7, 9, 11 and 15 the expression was studied in both the *in vitro* and *ex vitro* plants while for the genes 24, 27, 29, 33 and 35 only in *in vitro* plants.

The **gene 5** showed similarity to a class I chitinase, a basic chitinase in *Arabidopsis thaliana* also classified as pathogen related protein 3 (PR3) activated in response to fungi but more generally in the jasmonic acid- and ethylene-dependent systemic resistance and ethylene mediated signalling pathway. This gene resulted to be significantly differentially expressed in *in vitro* and *ex vitro* plants. In *in vitro* plants the gene showed over-expression in MS as well as in GD AP and GD AT2. On the contrary, a significant under-expression was seen in H0909 and a similar trend was observed for M9, both genotypes which share a part of their genome. In contrast to the observations in *in vitro* plants, the gene was under-expressed in *ex vitro* plants of GD while in all the resistant genotypes and all the tissues it was over-expressed. In M9 the gene resulted also to be over-expressed but only in the eIF4-normalised dataset. It is known (Christensen *et al.*, 2005) that after infection the phytoplasmas are activating the general response against pathogens or stress. However, no direct effects are expected from chitinase against phytoplasma because they lack a cell wall. Interestingly, in *in vitro* plants healthy H0909 and MS showed higher general expression level compared to the other genotypes whereas after infection in H0909 the level decreased in contrast to MS where there was an increment of the expression. In *ex vitro* plants the expression level in M9 root was higher compared to the other genotypes and did not change after

infection. Differently, in the resistant genotypes the expression increased to level comparable to that of M9.

The **gene 6** showed similarity to a chlorophyll a/b-binding protein. In previous works it was shown that the chlorophyll a/b-binding protein content decreased in phytoplasma-infected plants (Bertamini *et al.*, 2002; Bertamini *et al.*, 2003; Hren *et al.*, 2009). This was associated either to a down-regulation of genes encoding for chlorophyll a/b-binding proteins associated to the photosystem I and II (Hren *et al.*, 2009) or due to a general depletion in the protein and pigment content of infected leaves (Chang, 1998; Bertamini *et al.*, 2002; Bertamini *et al.*, 2003). The study of the expression level of the gene 6 in *in vitro* plants showed a down-regulation in MS (EF-normalised dataset) while in M9 (both controls) and H0909 (eIF4-normalised dataset) the expression level increased. Again, the genetically related genotypes exhibited similar gene expression pattern. In *ex vitro* plants it was not possible to obtain reliable results probably because of problems related to the primers used. In *in vitro* plants it is interesting to observe that the general expression level in M9 was between 6 and 11 times lower than that of the other genotypes.

The **gene 7** showed similarity with a COBRA-like protein. COBRA is a protein family involved in the morphological development of the cell wall contributing to the interaction between the membrane and the cell wall (Brady *et al.*, 2007) affecting the cell expansion and its orientation in roots (Schindelman *et al.*, 2001; Ko *et al.*, 2006). In *in vitro* plants it resulted under-expressed only in eIF4-normalised data of GD AP. In *ex vitro* plants in M9 root, H0909 (leaf and root) and D2212 leaf the gene resulted over-expressed in both the normalised datasets while in D2212 root (eIF4-normalised data) the gene was under-regulated. In *ex vitro* plants, also for this gene a higher general expression level can be observed in M9 root compared to the other genotypes and more in general it seems to be higher expressed in root tissues compared to leaf tissues.

It would be interesting to further investigate this gene to see if its involvement in the response to the phytoplasma infection will be confirmed. In this case we could hypothesise that the phytoplasmas presence induce modification in the expression of genes involved in the modification of the cell shape determining the morphological changes observed in phloematic tissues of symptomatic trees. In addition, Ko *et al.* (2006) showed that in *A. thaliana cob-5* mutants the absence of this gene triggers the expression of defence- and stress-related genes and that the mutant plant exhibits abnormal cell growth throughout the entire plant body and accumulate massive amounts of stress response chemicals such as anthocyanins and callose.

The **gene 9** showed similarity with a metallothionein. Metallothioneins may provide protection against heavy metal toxicity and oxidative stresses (Mir *et al.*, 2004; Wong *et al.*, 2004). In addition, they were shown to be involved in the response to a pathogenic attack and wounding (Choi *et al.*, 1996; Degenhardt *et al.*, 2005). In another work, a metallothionein gene was identified as differentially expressed upon phytoplasma infection in *Prunus* (Carginale *et al.*, 2004). The authors suggested that the presence of the phytoplasma would induce a higher protein turnover in the plant

increasing the availability of aminoacids essential for this pathogens which are dependending from their host for several metabolites necessary for their cellular functions. This could generate a high amount of heavy metals that would be toxic if not sequestered by proteins like the metallothionein. From the qRT-PCR analysis in *in vitro* plants it resulted to be significantly over-expressed only in H0909. In *ex vitro* plants over-expression was observed in H0909 and D2212 in leaf tissues (eIF4- and EF-normalised dataset) and GD root (18S-normalised data). In MS root and H0909 root the gene showed under-expression (both dataset). Noteworthy is the opposite behaviour in root and leaf tissues in the resistant genotypes. From the analysis of the general expression level it is interesting to observe that the gene 9 was present at a higher level in root compared to leaf tissues. In the latter the expression increases after infection and reaches in H0909 and D2212 a value comparable to that observed in root tissues. According to the hypothesis of Carginale *et al.* (2004) this signifies that the phytoplasma concentration in roots as well as in *in vitro* plants of resistant genotypes was not high enough to induce metallothionein expression. The low phytoplasma concentrations measured in roots of field-grown plants (Bisognin *et al.*, 2008b) as well as in tissue culture plants of *M. sieboldii* and D2212 (Bisognin *et al.*, 2008a) support this hypothesis. However, the over-expression of gene 9 in H0909 in *in vitro* as well as in *ex vitro* plants has to be explained by other factors.

The **gene 11** showed similarity to an aminopeptidase. The aminopeptidase function is related to the proteolysis and peptidolysis that acts in the protein metabolism and turnover. Previous studies suggested that higher protein degradation could be correlated with the phytoplasma infection because it lacks the enzymatic pathway for the production of basic molecules like aminoacids (Carginale *et al.*, 2004; Christensen *et al.*, 2005). Therefore, aminopeptidases could be differentially regulated after the infection of phytoplasmas. The gene was found to be over-expressed only in *in vitro* MS. In *ex vitro* plants the gene showed over-expression in GD root and D2212 leaf in both eIF4- and 18S-normalised data as well as in MS root (eIF4-normalised data). In *ex vitro* plants M9 and MS the gene 11 showed a higher general expression value compared to the other genotypes. These data are difficult to correlate to a higher or lower phytoplasma concentration in the plant as discussed above for gene 9. Further studies are needed to clarify the role of gene 11 in the phytoplasma-plant interaction.

The study of the expression level for the **gene 15** revealed a down-regulation in *in vitro* plants of D2212 and H0909 that was significant in both normalised datasets for D2212 and in the eIF4-dataset for H0909. Although a statistically significant up-regulation was seen in GD AT2 (EF-dataset), this was rather weak. A down-regulation of this gene was observed in *ex vitro* plants only in roots of D2212 (both normalised datasets) whereas it resulted over-expressed in M9 root, H0909 leaf and D2212 leaf in both the normalised datasets. A differential expression of gene 15 in *M. sieboldii* was absent in *in vitro* plants and only weak in roots of *ex vitro* plants (18S-normalised data). It is interesting to observe that the general expression level in root tissue was higher

compared to the expression level in leaf tissue where it seems to be activated after infection. In the BLAST analysis gene 15 showed similarity with a leucine rich repeat (LRR) family protein/extension family protein potentially involved in the regulation of cell wall expansion in response to signalling (Baumberger *et al.*, 2003). The data obtained in this thesis do not reveal a clear correlation of the expression pattern of this gene to AP-resistance or phytoplasma-infection. Therefore, further investigations are necessary to understand if this gene is indeed involved in the gene response after the phytoplasma infection.

The cDNA-AFLP bands from which the **gene 24** was identified were isolated from all the genotypes because this band indicated an over-expression in all the infected states. After sequencing of the corresponding clones a sequence common to all genotypes was identified which showed a similarity to an AKINgamma gene of *Arabidopsis thaliana*. AKINgamma is a subunit involved in the formation of a kinase complex SnRK1 (Gissot *et al.*, 2006). This kinase is putatively acting in plant-specific functions such as plant-pathogen interactions. In the real-time PCR analysis on *in vitro* plants a trend of a down-regulation of this gene was observed which was significant in GD AP and D2212 AP in EF-normalised data and in M9 AP and H0909 AT2 for the eIF4-normalised data. These results are in contrast to the expression pattern observed in the cDNA-AFLP analysis where the corresponding band was selected as up-regulated. Its general expression level resulted to be comparable in all the genotypes though in GD healthy and AT2 infected it was slightly higher. Assuming that gene 24 plays a general role in plant-pathogen interactions these contrasting data would indicate fluctuation in the expression of this gene because the biological replicates used for the cDNA-AFLP analysis and the real-time PCR analysis behaved differently. Further repetitions and investigations are necessary - also including the *ex vitro* plants - to investigate if this gene is really involved in a general response of the plant to the phytoplasma infection and at which period of the infection process it is activated.

The **gene 25** showed similarity to a squalene monoxygenase (squalene epoxidase 3 from *A. thaliana*; accession number NP_568033.1). This enzyme converts squalene into oxidosqualene which is the precursor of cyclic triterpenoids including membrane sterols, brassinosteroid phytohormones and non-steroidal triterpenoids (Xu *et al.*, 2004; Phillips *et al.*, 2006; Rasbery *et al.*, 2007). These are classes of compounds that are associated with several different cellular roles like components of the membrane (membrane sterols), plant growth and development (brassinosteroids) and disease resistance (saponins) (Phillips *et al.*, 2006), As shown in previous work on spiroplasmas, steroids are needed for the growth of spiroplasmas in axenic condition (Bové *et al.*, 1989). Also, '*Ca.* P. mali' has no metabolic pathway for steroid synthesis and is dependent on the steroids of the plant (Kube *et al.*, 2008). From the expression analysis the gene 25 resulted over-expressed in MS AT2 and under-expressed in H0909 AT2. Interestingly, the general expression level in GD healthy and infected was higher compared to the other genotypes and that a similar level was reached in MS plants after infection.

The **gene 27** was found in several clones and resulted to have homology with a copine like bonzai1 (BON1/CPN family) protein of *Arabidopsis thaliana*. This gene is known to be involved in the down-regulation of a disease resistance gene (R) SNC1. Yang *et al.* (2006) showed that this is a gene family that can be involved in the repression of several R genes and, therefore, its study in this thesis was interesting in order to see if there was differential regulation after the infection through phytoplasma. Only in M9 AP (eIF4-normalised data) and D2212 AP (18S-normalised data) a weak but significant under-regulation was shown. Also in this case the general expression level in GD healthy and infected *in vitro* plants was higher compared to that observed in the other genotypes. As cDNA-AFLP data and real-time PCR data are not in agreement, further studies are needed to verify a potential role of gene 27 in the infection process of '*Ca*. P. mali'. Assuming the repression of R genes by the expression of gene 27, its lower general expression level in AP-resistant genotypes would fit. Although in this scenario the under-expression in M9 needs to be explained, this would be an indication that resistance genes are involved in AP-resistance.

The **gene 29** showed homology to an Hsp70-like protein. Hsp70 molecular chaperones of plants are encoded by a multi-gene family whose members are developmentally regulated and differentially expressed in response to temperature stress and other conditions that interrupt normal protein folding or favour protein denaturation (Miemyk, 1997; Sung *et al.*, 2001). Stress-induced Hsp70s function to mitigate aggregation of stress-denatured proteins and to refold non-native proteins restoring their biological function (Sung *et al.*, 2001). It is proposed that, at least in tomato, Hsp70/Hsc70 is induced by avirulent strains of *R. solanacearum* as part of the defence response to chaperone newly synthesized defence proteins and to maintain cellular homeostasis essential for the execution of a defence response (Byth *et al.*, 2001). In the real-time PCR analysis the gene resulted to be down-regulated in GD AT2 and GD AP (significant in EF-normalised data). In the general expression analysis a lower level was observed in M9 (about 5-6 times), H0909 (about 3 times), MS and D2212 (about 2 times) compared to the GD healthy plants. Also in this case the real-time PCR data do not correspond to the cDNA-AFLP data in which the band corresponding to gene 29 was selected as over-expressed in resistant genotypes. However, the real-time PCR analysis revealed an important down-regulation of this gene in resistant genotypes – even when compared to the infected status of GD. Assuming a similar function of gene 29 in AP-resistance as in the system *R. solanacearum*/tomato (Byth *et al.*, 2001) the obtained data indicate a time point-dependent expression of this gene. In the infected resistant *in vitro* plants used for the real-time PCR experiment the plant defense could have reached a stable level and the expression of gene 29 would not have been necessary any more. A specific experiment in which the expression of gene 29 is studied at different time points after *in vitro* graft-inoculation is needed to verify this hypothesis.

The **gene 33** showed similarity to a glycolate oxidase. Glycolate oxidase is a peroxisomal flavin mononucleotide (FMN)-dependent oxidase that catalyses the oxidation of α-hydroxy acids to

the corresponding α-ketoacids. In green plants, glycolate oxidase is one of the enzymes in photorespiration, a pathway that results in reduced net photosynthesis (Lindqvist and Branden, 1989). In other works, it was shown that this enzyme is also involved in biotic and abiotic stresses playing a role in the regulation of H_2O_2 generation (Corpas *et al.*, 2001; del Rio *et al.*, 2003; Palma *et al.*, 2009). In the real-time PCR analysis of *in vitro* plants the gene 33 showed significant differential expression only in M9 AP (eIF4-normalised data) resulting down-regulated. An over-expression was observed with both normalisers in MS AT2 although it was not significant due to the high variability of the data. An up-regulation in infected plants of the resistance donor *M. sieboldii* confirm the cDNA-AFLP data and would be an indication that gene 33 is involved in the resistance to '*Ca.* P. mali'.

5.4.2 The phytoplasmas induce a general stress response in apple plants

The data obtained from the gene expression analysis were in several cases affected by variability. The cDNA-AFLP data and the real-time PCR data were often not in agreement. Although the gene expression pattern were analysed in stably infected *in vitro*-grown plants, this variation can be attributed in part to the different plant material used for the two analyses. As discussed above, cDNA-AFLP data have technical limits and need confirmation by an independent method. In this respect, qRT-PCR yields data that are more reliable and, therefore, the interpretation of the gene expression results was mainly based on the real-time PCR data. Although the *in vitro* system has been proven to be a powerful tool for the study of the resistance to AP in apple plants, the genes individuated as differentially expressed had to be confirmed in plants grown in conditions similar to those present in the orchards. Therefore, alternatively to further qRT-PCR analysis on the *in vitro* plants, *ex vitro* plants were used to assess the expression of the analysed genes. The interpretation of the data is also limited by the high number of genes for which no function could be assigned (genes 3, 14, 21, 22, 28, 30, 31, 32). E.g., the differential expression of gene 22 indicate a possible involvement in the resistance phenotype but no hypothesis of the resistance mechanism can be deduced because of its unknown function.

The determination of the entire genome sequence of '*Ca.* P. mali' (Kube *et al.*, 2008) revealed that the phytoplasma lacks a wide range of metabolic pathways and is dependent on a multitude of plant metabolites. Biochemical measurements of the compounds produced or degraded during the infection showed a broad spectrum of metabolisms involved in the plant-phytoplasma interaction (Kartte and Seemüller, 1991; Lepka *et al.*, 1999; Choi *et al.*, 2004). Consequently, genes involved in various metabolic pathways have been found to be differentially regulated after phytoplasma infection by other studies (Carginale *et al.*, 2004; Aldaghi, 2009). This further complicated the individuation of genes directly involved in the resistance response against '*Ca*. P. mali'.

The gene expression data obtained in this thesis demonstrate that '*Ca.* P. mali' induces a general stress response in apple plants. Based on the genes which had homology to known sequences genes associated with general plant defence, energy transport/oxidative stress response, protein metabolism and cellular growth were identified.

Genes 5, 24, 27 and 29 are related to the general plant defence induced upon pathogen attack or abiotic stress. It is not clear if these genes are activated by a general stress due to the presence of the phytoplasma (blocking of sieve tubes, depletion of metabolites etc.) or because of the identification of molecules that reveal its presence in the host. Although the genes 5 (chitinase I) and 29 (Hsp70) are homologous to proteins known to be involved in the general plant response, the genes 24 (AKINgamma) and 27 (BON1/CPN family) are good candidates for further investigations. AKIN proteins are involved in the assembly of the SnRK1 kinase complex which plays a role in plant-pathogen interaction (Hao *et al.*, 2003; Gissot *et al.*, 2006) and as cellular energy sensor regulating the resources allocation between plant organs (Schwachtje *et al.*, 2006; Lee *et al.*, 2009). Therefore, the gene 24 together with the gene 27 which has a role in the regulation of the disease resistance gene SNC1 merit further investigation as candidates for a specific answer in the resistant genotypes.

Genes 6, 9, 25 and 33 show homology to genes involved in energy transfer and oxidative metabolisms. Although differential expression was observed only in some genotypes this group could be interesting for further studies. Physiological studies in AP-diseased apple plants showed that the presence of the phytoplasma has a negative influx on the photosystems and in the discharge of the electron transfer chains (Bertamini *et al.*, 2002; Bertamini *et al.*, 2003). A resulting effect is the generation of reactive oxygen species (ROS) with subsequent production of hydrogen peroxide (H_2O_2). As mentioned above for the gene 33, glycolate oxidase can be involved in defensive mechanisms with generation of H_2O_2 (Corpas *et al.*, 2001; del Rio *et al.*, 2003; Palma *et al.*, 2009). H_2O_2 is involved in pathogen response (Baker and Orlandi, 1995) and activates a series of responses like the sterol pathway of which squalene monoxygenase (gene 25 and 35) is a key enzyme (Pose *et al.*, 2009). Because of the toxicity of H_2O_2 its effect is modulated in defence responses by ROS scavengers including metallothioneins (gene 9) which expression was shown to affect pathogen susceptibility to both bacterial blight and blast infection in rice (Oryza sativa) (Wong *et al.*, 2004). Interestingly, this gene in *ex vitro* plants of the resistant genotypes shows down-regulation in roots and up-regulation in leaf tissues.

The product of the gene 9 could be also associated, as well as the gene 11, to mechanisms related with protein degradation and aminoacid metabolisms. The phytoplasmas could specifically stimulate these pathways to render available basic molecules for their metabolism (Carginale *et al.*, 2004) or could imbalance the phloematic transport of nutrients unspecifically affecting the plant metabolism (Seemüller *et al.*, 1984; Lepka *et al.*, 1999).

The last functional group, to which the gene 7 and 15 belongs, was associated to genes that may be involved in the regulation of the plant development. Both showed differential expression in *ex vitro* plants not supported by the results in *in vitro* tissues. Further analyses are necessary to better characterise their expression behaviour. These genes could give insights on how the phytoplasma acts inducing dramatic changes in plant phenotype. Very little is known on this aspect of the AP-infection where the development of the symptoms suggest a strong modification of the signals and the pathways that regulate the cell growth but less is known on which pathways are involved.

Two genes (gene 14, 32) with unknown function showed a '*Ca.* P. mali' strain-specific differential expression. This indicates that the plant specifically reacts to the phytoplasma strain. Differences in virulence among '*Ca.* P. mali' strains (Seemüller and Schneider, 2007), as well as differences in phytoplasma concentration in the studied *in vitro* plants (Bisognin *et al.*, 2008a) support this strain-specific pathogen-plant interaction. This result further increases the complexity of the interaction between '*Ca.* P. mali' and the plant.

5.4.3 The resistance to apple proliferation

One major objective of this thesis was to analyse the nature of resistance to apple proliferation. As mentioned previously, this resistance is characterised by the absence of symptoms and a reduced concentration of the phytoplasma in the infected plant. This can be observed in the field (Seemüller *et al.*, 2008; Bisognin *et al.*, 2008b) where the root organ and the seasonal colonisation behaviour of the phytoplasma play an important role as well as in the *in vitro* system (Bisognin *et al.*, 2008a) where no roots are present. However, the nature of this resistance was completely unknown at the beginning of this thesis and the interaction of the phytoplasma with the resistant plant represented a "black box". Due to the complexity of the interaction between '*Ca.* P. mali' and the plant, the data obtained in this thesis did not allow elucidating the resistance mechanism, but they allowed shedding light into this "black box". As discussed in the previous chapter, a multitude of various plant genes is differentially expressed after phytoplasma infection. This was observed for the susceptible as well as for the resistant genotypes. However, no strong expression of a resistance-related gene was detected in the resistant genotypes. The picture which can be drawn from the data rather indicate two ways in which the resistance is manifested in the plant: 1. the general plant defense is specifically activated or 2. the production of metabolites necessary for the phytoplasma development is altered in resistant genotypes. The differential expression of the genes 5, 22, 27, 29 and 33 especially in the resistance donor *M. sieboldii* support the first hypothesis. The down-regulation of gene 6 and the differential regulation of gene 25 might be indications for the second hypothesis.

It is important to note that the gene expression was studied in homogenously infected plants. Thus, only genes which are constantely expressed during phytoplasma infection could be found

and not the genes which react at the beginning of the infection. Therefore, the resistant phenotype can also be due to a constitutive difference in the expression level of certain genes. A higher or lower constitutive expression level of certain genes could influence the susceptibility of the plant to pathogens (Gau *et al.*, 2004; Degenhardt *et al.*, 2005). The comparison of the differential expression between susceptible and resistant plants is not sufficient alone to give a picture of how the pathogen interacts with the host. Genes that appear not to be differentially expressed or that show a lower grade of induction or inhibition can have a constitutive expression level considerably higher or lower as the level shown by genotypes that react stronger after infection. In these cases, such genes will appear not or only slightly differentially expressed but their constitutive expression level could contribute to the resistance character against the pathogen attack (Degenhardt *et al.*, 2005).

The constitutive expression level of genes is of particular interest in the resistant genotypes as they are triploid or tetraploid. In this regard, the qRT-PCR analysis was very useful as it was able to reveal these differences of the general expression levels. Although the *in vitro* data were characterised in several cases by high variability the general expression level often appeared to be comparable between the genotypes. In addition, the comparison of the *in vitro* and *ex vitro* data observed for M9 (that has a 2n genome) with those for the resistant hybrids suggest that there is no correlation between ploidy and constitutive expression level. These observations are supported also by another study where gene expression analysis in a maize ploidy series showed that for several genes the expression level was not related to the ploidy level (Guo *et al.*, 1996). Thus, differences in general gene expression might be directly correlated to resistance and might be inherited in the SMAP breeding program to the recombinant progeny which has lower ploidy (3n). However, independent from the ploidy degree intrinsic differences in the constitutive expression levels of some genes were seen in the healthy state which indicate genotypic differences.

Interesting is also to observe that in *ex vitro* plants two genes have a different expression level between root and leaf tissues and that after infection these show an opposite tendency in their differential expression in the two tissues (genes 9 and 15).

In previous work it was shown that some plants could undergo a spontaneous remission of the symptoms, a phenomenon known as recovery. It was observed that in recovered plants the phytoplasma can be still present in the roots but cannot colonise the above ground part of the plant (Carraro *et al.*, 2004). In contrast, symptomatic trees were completely colonised. Variations in symptom expression are related to the presence or absence of phytoplasmas in the aerial part of the plants (Seemüller *et al.*, 1984). In recovered plants the phytoplasmas may or may not be eliminated from the bottom part (Schmid, 1965) but the diffusion of the pathogen in the phloem seems to be hindered. Little is known about the recovery mechanism. Musetti *et al* (2004) showed that in '*Ca*. P. mali'-infected leaves, but not in roots of the recovered plants, an overproduction of H_2O_2 was in association to a lower activity of guaiacol peroxidase (GPX) and a lower content of

reduced glutathione (GSH). Because GSH and GPX have an antioxidative role, in the same work the authors hypothesised that the low level of these compounds were correlated to the high level of H_2O_2 inducing an oxidative stress with the following production of reactive oxygen species (ROS). The evidence of ROS activity was supported by the observation of elevated malondialdehyde (MDA) content in the leaves of recovered plants. They suggested that the MDA derived from the biosynthesis of breakdown products due to lipid (membrane) peroxidation. This system was supposed to be at the base of the recovery allowing the elimination of the phytoplasmas from the canopy due to the toxicity of the produced ROS species (Musetti *et al.*, 2004; Musetti *et al.*, 2005).

The “recovery” phenotype resembles the “resistant” phenotype. It was therefore interesting to analyse if a similar mechanism is involved in the resistance observed in the *M. sieboldii*-derived genotypes. The expression pattern observed for the gene 25 (squalene monoxigenase) and the gene 33 (glycolate oxidase) would fit with a scenario where both genes are actors in a response resembling that presented by Musetti *et al.*. In peroxisomes, glycolate oxidase directly produce H_2O_2 (as byproduct of the photorespiration) which is removed by catalases (Corpas *et al.*, 2001; del Rio *et al.*, 2006; Palma *et al.*, 2009). Nevertheless, in stress conditions catalases can be repressed or not associated to peroxisomes (del Rio *et al.*, 1996; del Rio *et al.*, 2002) leading to the production of free H_2O_2. In addition, impaired content of chlorophyll a/b binding protein (gene 6 was down-regulated in MS) has a negative effect on the assembly of the photosystem II (PSII) and on the electron transport rate (de Bianchi *et al.*, 2008). In phytoplasma-infected plants it was shown that the content in photosystems I and II as well as that of chlorophyll a/b binding proteins is affected in leaves (Bertamini *et al.*, 2002; Bertamini *et al.*, 2003; Maust *et al.*, 2003; Albertazzi *et al.*, 2009; Hren *et al.*, 2009). In these conditions, CO_2 assimilation is impaired and an increased photorespiration leads to accumulation of glycolate that is transported to peroxisomes (Goyal and Tolbert, 1996; Fryer *et al.*, 1998; Maurino and Peterhansel, 2010). Finally, increased content in glycolate would generate higher content of H_2O_2 through the activity of the glycolate oxidase (Corpas *et al.*, 2001; del Rio *et al.*, 2006; Palma *et al.*, 2009). H_2O_2 has a signal function in many physiological processes, including hypersensitive response to pathogens, growth and development (Corpas *et al.*, 2001; Neill *et al.*, 2002; Neill *et al.*, 2002). It was shown that H_2O_2 stimulates the production of saponins (Hu *et al.*, 2003) which are synthetised starting from squalene and involving the activity of squalene epoxidase (gene 25). Saponins are known to have a role in the plant defence and to have antibacterial properties (Mandal *et al.*, 2005; Avato *et al.*, 2006). Among the gram-positive bacteria that were shown to be highly inhibited there was also *B. subtilis* which shares high homology on genomic structure with the phytoplasmas (Miyata *et al.*, 2002). Therefore, in case of the resistant *M. sieboldii* genotype we could hypothesise that the production of H_2O_2 is induced by the phytoplasma either through direct activation of the plant defense or through modification of photosynthetic pathways. H_2O_2 can act directly against the phytoplasma

itself or can induce the expression of further genes involved in the synthesis of pathogen defence-related molecules.

These responses activated in *M. sieboldii* by the phytoplasma presence would be at the base of its resistance. Comparing this behaviour to the modifications observed in recovered plants, it is intriguing to notice how the responses of resistant plants observed in this study resemble the phenomenon of recovery. Nevertheless, recovery occurs stochastically and especially old plants while in *M. sieboldii* it seems to be an inducible feature that is activated in a reproducible way resulting in resistance against 'Ca. P. mali'.

Further experiments are necessary in order to gain more information over the plant-phytoplasma interaction. One important aspect is that the qPCR method developed for the gene expression analysis proofed to be a reliable tool delivering results that can be repeated in a fast and easy procedure. This will allow the analysis of genes related to the pathways individuated in this thesis in order to elucidate the mechanisms involved.

5.5 Conclusions and outlook

The results obtained in this thesis showed that processes involved in the plant defense against pathogens are activated after phytoplasma infection. Genes 5, 24, 27 and 29 are involved in the general response to pathogens and differences between the susceptible and the resistant genotypes were also seen. In addition, the behaviour observed for the genes 14 and 22 merit further investigation to clarify if there is a specific answer in resistant genotypes against the '*Ca*. P. mali' and also if there are differences between '*Ca.* P. mali strains'. However, what can be observed from the experiments is a system in which the plant appears to react to the phytoplasma presence involving multiple mechanisms and metabolisms. Gene expression studies in such systems are difficult because of the high number of genes involved. Moreover, if the constitutive level of gene products has an influence on the capability of the phytoplasma to colonise the plant then an approach in which the differentially expressed genes are investigated will be not successful for the determination of those genes involved in the AP-defense. Differences of metabolites characterising *M. sieboldii* compared to *M. x domestica* can be at the base of the different susceptibility to '*Ca*. P. mali'. Therefore, although the genes that were found in this study were in line with previous work new insights were gained on the metabolisms or the mechanisms conferring the resistance character. The induced response observed in *M. sieboldii* makes us to hypothesise a type of resistance resembling the recovery phenomenon with the important characteristic that in *M. sieboldii* this would be an inducible, reproducible and inheritable feature. In addition, to prove that genes individuated as differentially expressed are really involved in the response and can be inherited maintaining their expression pattern in the apple offspring, an analysis on the segregating population has to be performed. Since it was observed that plants obtained from the crossings of the *M. sieboldii*-derived genotypes and the *M. x domestica* showed

different degrees of resistance (Bisognin *et al.*, 2008b), it is reasonable to think that this is characterised by a multigenic action instead of a monogenic feature. In such a case studies on the gene expression have to be conducted on a population of resistant plants in order to individuate common genic products involved in the response after the infection (Klink and Matthews, 2009; Manosalva *et al.*, 2009). A systematic screening of key enzymes involved in the reaction to pathogenic attacks could be carried out in a relatively high number of plants using as approach the qRT-PCR method employed in this thesis. Alternatively, the construction of a microarray containing all the possible genes expressed from *M. sieboldii* could be another approach for the individuation of differentially expressed genes but since no sequence information is available on this genotype, a preliminary cloning of the whole transcriptome would be necessary. This would be a very expensive and time consuming approach compared to a real time PCR approach.

Further possible information on the way the phytoplasma colonises the plant could be obtained from the study of its genome. The individuation of which plant metabolisms are required and if there are virulence factors that are recognised by the plants could give further insights on the pathogen-host interaction.

This thesis was initiated when very poor information was available on the phytoplasma-apple interaction. The qRT-PCR method developed here for the fast analysis of the gene expression is a tool that showed a great potential for the study of the '*Ca*. P. mali'-apple interaction and can be employed for future screening of the expression levels of genes putatively involved in the response mechanisms. Even if observed differences in gene expression were not statistically significant due to high variability of the data for some genes important indications about their expression behaviour could be obtained.

The expression level analysis of genes known to be related to defence mechanisms can give insights on the differences that render *M. sieboldii* plants resistant compared to *M.* x *domestica*. This can be an outlook of this work where the qRT PCR method can be employed for the direct comparison of specific metabolic pathways. The increasing availability of information on the genomic sequences (IASMA is completing the sequencing of the *M.* x *domestica* cv Golden delicious) can represent a source for the individuation of putative genes related to metabolic differences with the resistant genotypes. Studies focused on the comparison of such metabolisms could be conducted for the determination if such differences are involved in the response mechanism against the pathogen.

The general picture resulting from this thesis shows that a broad range of modifications take place in the plant metabolisms after the infection of the phytoplasma and that there are variations observed in the expression of genes involved in the general answer against stress and pathogens that could be at the base of the resistance character in *M. sieboldii*. In addition, plants that result resistant could be advantaged having a general expression level of gene products that better absorbe the impact due to the phytoplasma presence or that are creating physiological conditions

hindering the successful colonization of the host by the phytoplasmas. This aspect has to be further investigated. An approach focused on the characterization of the expression level of the genes in relation to the capability of the phytoplasma to colonise the plant would be necessary instead of studying the variations in the gene expression that occur after the infection. Important would be also the design of an experiment where the infection process is monitored through a period. A time course analysis would be necessary to investigate different activations of genes in the early phases after graft-inoculation. Genes expressed in the initial steps of the phytoplasma-plant interaction could play an important role in the definition of the resistance. In this thesis, it was not possible to investigate this aspect because the genes were studied in the late period where they could have reached a steady state and other genes could have been involved in the process only in a transitional way. Moreover, genes that were found to be differentially expressed should be further analysed in a higher number of plants in other genotypes to confirm their differential expression after the phytoplasma infection. In this view, the characterisation of the resistant plants derived from the new crossing will be crucial for the selection of candidate genotypes for the development of a new rootstock with sufficient resistant traits and good agronomic characteristics. Therefore, the markers K, C8 and 30 together with the SNP markers will add further high quality markers to the SSR data for the genomic analysis of the breeding progeny and the mapping of the resistance trait. In addition, the association of these markers with expressed genes will represent a further advantage allowing gaining information about the mechanisms or pathways that are inherited and the phenotypical traits observed in the seedlings. This would be not only an advantage for the selection of the resistant traits but also for the selection of other interesting traits like those associated to agronomical values.

6 Literature

Afzal AJ, Wood AJ, Lightfoot DA (2008) Plant receptor-like serine threonine kinases: roles in signaling and plant defense. Mol Plant Microbe Interact **21:** 507-517

Ahrens U, Seemüller E (1992) Detection of DNA of plant pathogenic mycoplasmalike organisms by a polymerase chain reaction that amplifies a sequence of the 16S rRNA gene. Phytopathology **82:** 828-832

Albertazzi G, Milc J, Caffagni A, Francia E, Roncaglia E, Ferrari F, Tagliafico E, Stefani E, Pecchioni N (2009) Gene expression in grapevine cultivars in response to Bois Noir phytoplasma infection. Plant Science **176:** 792-804

Aldaghi M (2009) Identification of host genes potentially implicated in the "Malus pumila, rootstock MM106" "Candidatus Phytoplasma mali" interactions. Universite de Gembloux

Alskieff J, Villermur P (1978) Greffage *in vitro* d'apex sur des plantules decapitées de pommier (*Malus pumila* Mill.). Compte Rendu de l´Académie des Sciences, Série D **287:** 1115-1118

Avato P, Bucci R, Tava A, Vitali C, Rosato A, Bialy Z, Jurzysta M (2006) Antimicrobial activity of saponins from Medicago sp.: structure-activity relationship. Phytother Res **20:** 454-457

Bachem CW, van der Hoeven RS, de Bruijn SM, Vreugdenhil D, Zabeau M, Visser RG (1996) Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. Plant J **9:** 745-753

Bai X, Zhang J, Ewing A, Miller SA, Jancso Radek A, Shevchenko DV, Tsukerman K, Walunas T, Lapidus A, Campbell JW, Hogenhout SA (2006) Living with genome instability: the adaptation of phytoplasmas to diverse environments of their insect and plant hosts. J Bacteriol **188:** 3682-3696

Baker CJ, Orlandi EW (1995) Active oxygen in plant pathogenesis. Annu. Rev. Phytopathol. **33:** 299-321

Baldo A, Norelli JL, Farrell RE, Jr., Bassett CL, Aldwinckle HS, Malnoy M (2010) Identification of genes differentially expressed during interaction of resistant and susceptible apple cultivars (Malus x domestica) with Erwinia amylovora. BMC Plant Biol **10:** 1

Barbara DJ, Morton A, Clark MF, and Davies DL (2001) Molecular variation in immunodominant membrane proteins from phytoplasmas. Acta Horticulturae **550:** 405-408

Barbara DJ, Morton A, Clark MF, Davies DL (2002) Immunodominant membrane proteins from two phytoplasmas in the aster yellows clade (chlorante aster yellows and clover phyllody) are highly divergent in the major hydrophilic region. Microbiology **148:** 157-167

Baric S, Dalla-Via J (2004) A new approach to apple proliferation detection: a highly sensitive real-time PCR assay. J Microbiol Methods **57:** 135-145

Baric S, Kerschbamer C, Vigl J, Dalla Via J (2007) Translocation of apple proliferation phytoplasma via natural root grafts – a case study. European Journal of Plant Pathology **121:** 207-211

Batjer LP, Schneider H (1960) Relation of pear decline to rootstocks and sieve-tube necrosis. Proc. Am. Soc. Hortic. Sci. **76:** 85-97

Baumberger N, Doesseger B, Guyot R, Diet A, Parsons RL, Clark MA, Simmons MP, Bedinger P, Goff SA, Ringli C, Keller B (2003) Whole-genome comparison of leucine-rich repeat extensins in Arabidopsis and rice. A conserved family of cell wall proteins form a vegetative and a reproductive clade. Plant Physiol **131:** 1313-1326

Becerra C, Jahrmann T, Puigdomenech P, Vicient CM (2004) Ankyrin repeat-containing proteins in Arabidopsis: characterization of a novel and abundant group of genes coding ankyrin-transmembrane proteins. Gene **340:** 111-121

Berg M, Davies DL, Clark MF, Vetten HJ, Maier G, Marcone C, Seemuller E (1999) Isolation of the gene encoding an immunodominant membrane protein of the apple proliferation phytoplasma, and expression and characterization of the gene product. Microbiology **145 (Pt 8):** 1937-1943

Bertamini M, Grando MS, Muthuchelian K, Nedunchezhian N (2002) Effect of phytoplasmal infection on photosystem II efficiency and thylakoid membrane protein changes in field grown apple (Malus pumila) leaves. Physiological and Molecular Plant pathology **61:** 349-356

Bertamini M, Grando MS, Nedunchezhian N (2003) Effects of phytoplasma infection on pigments, chlorophyll-protein complex and photosynthetic activities in field grown apple leaves. Biologia plantarum **47:** 237-242

Bertamini M, Muthuchelian K, Grando MS, Nedunchezhian N (2002) Effects of phytoplasma infection on growth and photosynthesis in leaves of field grown apple (Malus pumila mill. cv. Golden delicious). Photosynthetica **40:** 157-160

Bisognin B, Ciccotti A, Salvadori A, Moser M, Grando MS, and Jarausch W (2008a) In vitro screening for resistance to apple proliferation in Malus ssp. Plant Pathology **57:** 1163-1171

Bisognin C, Schneider B, Salm H, Grando MS, Jarausch W, Moll E, Seemuller E (2008b) Apple Proliferation Resistance in Apomictic Rootstocks and Its Relationship to Phytoplasma Concentration and Simple Sequence Repeat Genotypes. Phytopathology **98:** 153-158

Bisognin C, Schneider B, Salm H, Grando MS, Jarausch W, Moll E, Seemüller. aE (2008b) Apple proliferation resistance in apomictic rootstocks and its relationship to phytoplasma concentration and simple sequence repeat genotypes. Phytopathology **98:** 153-158

Bisognin C, Seemüller E, Citterio S, Velasco R, Grando MS, Jarausch W (2009) Use of SSR markers to assess sexual vs. apomictic origin and ploidy level of breeding progeny derived from crosses of apple proliferation-resistant Malus sieboldii and its hybrids with *Malus* × *domestica* cultivars Plant breeding **128:** 507-513

Bonnet F, Saillard C, Kollar A, Seemüller E, Bové JM (1990) Detection and differentiation of the mycoplasma-like organism associated with the apple proliferation disease using cloned DNA probes. Molecular Plant-Microbe Interactions**:** 438-443

Bové JM, Carle P, Garnier M, Laigret F, Renaudin J, and Saillard C (1989) Molecular and cellular biology of spiroplasmas. *In* e R.F. Whitcomb and J. G. Tully, ed, The Mycoplasmas, Vol 5. Academic Press, New York, pp 243-364

Bové JM, Garnier M (1998) Walled and wall-less eibacteria from plants: sieve-tube-restricted plant pathogens. Plant Cell, Tissue and Organ Culture **52:** 7-16

Bové JM, Garnier M (2003) Phloem- and xylem-restricted plant pathogenic bacteria. Plant Science **164:** 423-438

Bove JM, Renaudin J, Saillard C, Foissac X, Garnier M (2003) Spiroplasma citri, a plant pathogenic molligute: relationships with its two hosts, the plant and the leafhopper vector. Annu Rev Phytopathol **41:** 483-500

Brady SM, Song S, Dhugga KS, Rafalski JA, Benfey PN (2007) Combining expression and comparative evolutionary analysis. The COBRA gene family. Plant Physiol **143:** 172-187

Branz A (2003) Il monitoraggio sulla diffusione degli scopazzi del melo in Valle di None di Sole. *In*, Presented at the meeting " La frutticoltura delle valli del Noce:6a giornata tecnica, Feb.18, Cles, Italy

Braun EJ, Sinclair WA (1978) Translocation in phloem necrosis-diseased American elm seedlings. Phytopathology **68:** 1733-1737

Breyne P, Dreesen R, Cannoot B, Rombaut D, Vandepoele K, Rombauts S, Vanderhaeghen R, Inze D, Zabeau M (2003) Quantitative cDNA-AFLP analysis for genome-wide expression studies. Mol Genet Genomics **269:** 173-179

Breyne P, Dreesen R, Vandepoele K, De Veylder L, Van Breusegem F, Callewaert L, Rombauts S, Raes J, Cannoot B, Engler G, Inze D, Zabeau M (2002) Transcriptome analysis during cell division in plants. Proc Natl Acad Sci U S A **99:** 14825-14830

Breyne P, Zabeau M (2001) Genome-wide expression analysis of plant cell cycle modulated genes. Curr Opin Plant Biol **4:** 136-142

Bustin SA (2002) Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. J Mol Endocrinol **29:** 23-39

Bustin SA, Benes V, Nolan T, Pfaffl MW (2005) Quantitative real-time RT-PCR--a perspective. J Mol Endocrinol **34:** 597-601

Bustin SA, Nolan T (2004) Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. J Biomol Tech **15:** 155-166

Byth H-A, Karolina GK, B. aL (2001) Virulence-dependent induction of Hsp70/Hsc70 in tomato by Ralstonia solanacearum. Plant Physiology and Biochemistry **39:** 697-705

Cainelli C, Bisognin C, Vindimian ME, Grando MS (2004) Genetic variability of AP phytoplasmas detected in the apple growing area of Trentino (north Italy). Acta Horticulturae **657:** 425-430

Callis J, Carpenter T, Sun CW, Vierstra RD (1995) Structure and evolution of genes encoding polyubiquitin and ubiquitin-like proteins in Arabidopsis thaliana ecotype Columbia. Genetics **139:** 921-939

Cane PA, Cook P, Ratcliffe D, Mutimer D, Pillay D (1999) Use of real-time PCR and fluorimetry to detect lamivudine resistance-associated mutations in hepatitis B virus. Antimicrob Agents Chemother **43:** 1600-1608

Carginale V, Maria G, Capasso C, Ionata E, La Cara F, Pastore M, Bertaccini A, Capasso A (2004) Identification of genes expressed in response to phytoplasma infection in leaves of Prunus armeniaca by messenger RNA differential display. Gene **332:** 29-34

Carraro L, Ermacora P, Loi N, R. aO (2004) The recovery phenomenon in apple proliferation-infected apple trees. Journal of Plant Pathology **86:** 141-146

Catlin PB, Olsson EA, Beutel JA (1975) Reduced translocation of carbon and nitrogen from leaves with symptoms of pear curl. J.Am.Soc.Hortic.Sci. **100:** 184-187

Chang C-J, Lee IM (1995) Pathogenesis of diseases associated with mycoplasma-like organisms. *In* RPS U.S.Singh, and K.Kohmoto, eds., ed, Pathogenesis and Host Specificity in Plant Diseases, Vol 1. Elsevier Science publishing Co., New York, pp 237-246

Chang CJ (1998) Pathogenicity of aster yellows phytoplasma and *Spiroplasma citri* on periwinkle. Phytopathology **88:** 1347-1350

Chen T-A, Lei JD, and Lin CP (1992) Detection and identification of plant and insect mollicutes. *In* e R.F. Whitcomb and J. G. Tully, ed, The Mycoplasmas., Vol 5. Academic Press, New York, pp 393-424

Choi D, Kim HM, Yun HK, Park JA, Kim WT, Bok SH (1996) Molecular cloning of a metallothionein-like gene from Nicotiana glutinosa L. and its induction by wounding and tobacco mosaic virus infection. Plant Physiol **112:** 353-359

Choi YH, Tapias EC, Kim HK, Lefeber AW, Erkelens C, Verhoeven JT, Brzin J, Zel J, Verpoorte R (2004) Metabolic discrimination of Catharanthus roseus leaves infected by phytoplasma using 1H-NMR spectroscopy and multivariate data analysis. Plant Physiol **135:** 2398-2410

Christensen NM, Axelsen KB, Nicolaisen M, Schulz A (2005) Phytoplasmas and their interactions with hosts. Trends Plant Sci **10:** 526-535

Ciccotti AM, Bianchedi PL, Bragagna P, Deromedi M, Filippi M, Forno F, and Mattedi L (2008a) Natural and experimental transmission of Candidatus Phytoplasma mali by root bridges. Acta Horticulturae **781:** 459-464

Ciccotti AM, Bisognin C, Battocletti I, Salvadori A, Herdemertens M, Jarausch W (2008b) Micropropagation of Malus sieboldii hybrids resistant to apple proliferation disease. Agronomy research **6:** 445-458

Corpas FJ, Barroso JB, del Rio LA (2001) Peroxisomes as a source of reactive oxygen species and nitric oxide signal molecules in plant cells. Trends Plant Sci **6:** 145-150

Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiol **139:** 5-17

Davis RE, Lee IM (1993) Cluster-specific polymerase chain reaction amplification of 16S rDNA sequences for detection and identification of mycoplasmalike organisms. Phytopathology **83:** 1008-1011

de Bianchi S, Dall'Osto L, Tognon G, Morosinotto T, Bassi R (2008) Minor antenna proteins CP24 and CP26 affect the interactions between photosystem II subunits and the electron transport rate in grana membranes of Arabidopsis. Plant Cell **20:** 1012-1028

Deeley J, Stevens WA, and Fox RTV (1979) Use of Diene's stain to detect plant diseases induced by mycoplasma-like organisms. Phytopathology **69:** 1169-1171

Degenhardt J, Al-Masri AN, Kurkcuoglu S, Szankowski I, Gau AE (2005) Characterization by suppression subtractive hybridization of transcripts that are differentially expressed in leaves of apple scab-resistant and susceptible cultivars of Malus domestica. Mol Genet Genomics **273:** 326-335

del Rio LA, Corpas FJ, Sandalio LM, Palma JM, Barroso JB (2003) Plant peroxisomes, reactive oxygen metabolism and nitric oxide. IUBMB Life **55:** 71-81

del Rio LA, Corpas FJ, Sandalio LM, Palma JM, Gomez M, Barroso JB (2002) Reactive oxygen species, antioxidant systems and nitric oxide in peroxisomes. J Exp Bot **53:** 1255-1272

del Rio LA, Palma JM, Sandalio LM, Corpas FJ, Pastori GM, Bueno P, Lopez-Huertas E (1996) Peroxisomes as a source of superoxide and hydrogen peroxide in stressed plants. Biochem Soc Trans **24:** 434-438

del Rio LA, Sandalio LM, Corpas FJ, Palma JM, Barroso JB (2006) Reactive oxygen species and reactive nitrogen species in peroxisomes. Production, scavenging, and role in cell signaling. Plant Physiol **141:** 330-335

DeLille JM, Sehnke PC, Ferl RJ (2001) The arabidopsis 14-3-3 family of signaling regulators. Plant Physiol **126:** 35-38

Didenko VV (2001) DNA probes using fluorescence resonance energy transfer (FRET): designs and applications. Biotechniques **31:** 1106-1116, 1118, 1120-1101

Diehl F, Beckmann B, Kellner N, Hauser NC, Diehl S, Hoheisel JD (2002) Manufacturing DNA microarrays from unpurified PCR products. Nucleic Acids Res **30:** e79

Doi Y, Teranaka M, Yora K, and Asuyama H (1967) *Mycoplasma* or PLT-group-like microorganisms found in the phloem elements of plants infected with mulberry dwarf, potato witches-broom, aster yellow, or paulownia witches-broom. Annals of the Phytopathological Society of Japan **33:** 259-266

Dorak MT (2006) Real-time PCR. Taylor & Francis Ltd

Eden-Green SJ, Waters H (1982) Collection and properties of phloem sap from healthy and LY-diseased coconut palms in Jamaica. Phytopathology **72:** 667-672

Exposito-Rodriguez M, Borges AA, Borges-Perez A, Perez JA (2008) Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. BMC Plant Biol **8:** 131

Favali MA, Musetti R, Benvenuti S, Bianchi A, Pressacco L (2004) Catharanthus roseus L. plants and explants infected with phytoplasmas: alkaloid production and structural observations. Protoplasma **223:** 45-51

Firrao G, Gobbi E, and Locci R (1994) Rapid diagnosis of apple proliferation mycoplasma-like organism using a polymerase chain reaction procedure. Plant Pathology **43:** 669-674

Fleige S, Pfaffl MW (2006) RNA integrity and the effect on the real-time qRT-PCR performance. Mol Aspects Med **27:** 126-139

Fleige S, Walf V, Huch S, Prgomet C, Sehm J, Pfaffl MW (2006) Comparison of relative mRNA quantification models and the impact of RNA integrity in quantitative real-time RT-PCR. Biotechnol Lett **28:** 1601-1613

Frisinghelli C, Delaiti L, Grando MS, Forti D, E. aVM (2000) Cacopsylla costalis (Flor 1861), as a Vector of Apple Proliferation in Trentino. Journal of Phytopathology **148:** 425-431

Fryer MJ, Andrews JR, Oxborough K, Blowers DA, Baker NR (1998) Relationship between CO2 Assimilation, Photosynthetic Electron Transport, and Active O2 Metabolism in Leaves of Maize in the Field during Periods of Low Temperature. Plant Physiol **116:** 571-580

Gabriels SH, Takken FL, Vossen JH, de Jong CF, Liu Q, Turk SC, Wachowski LK, Peters J, Witsenboer HM, de Wit PJ, Joosten MH (2006) cDNA-AFLP combined with functional analysis reveals novel genes involved in the hypersensitive response. Mol Plant Microbe Interact **19:** 567-576

Gachon C, Mingam A, Charrier B (2004) Real-time PCR: what relevance to plant studies? J. Exp. Bot. **55:** 1445-1454

Gasparich GE, Hackett KJ, Stamburski C, Renaudin J, Bove JM (1993) Optimization of methods for transfecting Spiroplasma citri strain R8A2 HP with the spiroplasma virus SpV1 replicative form. Plasmid **29:** 193-205

Gau AE, Koutb M, Piotrowski M, Kloppstech K (2004) Accumulation of Pathogenesis-Related Proteins in the Apoplast of a Susceptible Cultivar of Apple (*Malus domestica* cv. Elstar) after Infection by *Venturia inaequalis* and Constitutive Expression of PR Genes in the Resistant Cultivar Remo. Eur J Plant Path **110:** 703-711

Giannotti JG, Morvan G, Vago C (1968) Micro-organismes de type mycoplasme dans le cellules libériennes de *Malus sylvestris* L. atteint de la maladie de prolifération. Compte Rendu de l´Académie des Sciences, Série D **276:** 76-77

Gibson UE, Heid CA, Williams PM (1996) A novel method for real time quantitative RT-PCR. Genome Res **6:** 995-1001

Ginzinger DG (2002) Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream. Exp Hematol **30:** 503-512

Gissot L, Polge C, Jossier M, Girin T, Bouly JP, Kreis M, Thomas M (2006) AKINbetagamma contributes to SnRK1 heterotrimeric complexes and interacts with two proteins implicated in plant pathogen resistance through its KIS/GBD sequence. Plant Physiol **142:** 931-944

Goyal A, Tolbert NE (1996) Association of glycolate oxidation with photosynthetic electron transport in plant and algal chloroplasts. Proc Natl Acad Sci U S A **93:** 3319-3324

Gundersen DE, Lee I-M (1996) Ultrasensitive detection of phytoplasmas by nested-PCR assays using two universal primer pairs. Phytopathology Mediterranean **35:** 114-151

Gundersen DE, Lee I-M, Schaff DA, Harrison NA, Chang CJ, Davis RE, and Kingsbury DT (1996) Genomic diversity and differentiation among phytoplasma strains in 16S rRNA group I (aster yellows and related phytoplasmas) and III (X-disease and related phytoplasmas). International Journal of Systematic Bacteriology **46:** 64-75

Guo M, Davis D, Birchler JA (1996) Dosage effects on gene expression in a maize ploidy series. Genetics **142:** 1349-1355

Gutierrez L, Mauriat M, Guenin S, Pelloux J, Lefebvre JF, Louvet R, Rusterucci C, Moritz T, Guerineau F, Bellini C, Van Wuytswinkel O (2008) The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcription-polymerase chain reaction (RT-PCR) analysis in plants. Plant Biotechnol J **6:** 609-618

Hao L, Wang H, Sunter G, Bisaro DM (2003) Geminivirus AL2 and L2 proteins interact with and inactivate SNF1 kinase. Plant Cell **15:** 1034-1048

Heid CA, Stevens J, Livak KJ, Williams PM (1996) Real time quantitative PCR. Genome Res **6:** 986-994

Higuchi R, Dollinger G, Walsh PS, Griffith R (1992) Simultaneous amplification and detection of specific DNA sequences. Biotechnology (N Y) **10:** 413-417

Higuchi R, Fockler C, Dollinger G, Watson R (1993) Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. Biotechnology (N Y) **11:** 1026-1030

Hren M, Nikolic P, Rotter A, Blejec A, Terrier N, Ravnikar M, Dermastia M, Gruden K (2009) 'Bois noir' phytoplasma induces significant reprogramming of the leaf transcriptome in the field grown grapevine. BMC Genomics **10:** 460

Hu R, Fan C, Li H, Zhang Q, Fu YF (2009) Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. BMC Mol Biol **10:** 93

Hu X, Neill S, Cai W, Z. aT (2003) Hydrogen peroxide and jasmonic acid mediate oligogalacturonic acid-induced saponin accumulation in suspension-cultured cells of *Panax ginseng*. Physiologia Plantarum **118:** 414-421

IRPCM Phytoplasma/Spiroplasma Working Team - Phytoplasma Taxonomy Group *et al.* (2004) 'Candidatus Phytoplasma', a taxon for the wall-less, non-helical prokaryotes that colonize plant phloem and insects. Int J Syst Evol Microbiol **54:** 1243-1255

Jagoueix-Eveillard S, Tarendeau F, Guolter K, Danet JL, Bove JM, Garnier M (2001) Catharanthus roseus genes regulated differentially by mollicute infections. Mol Plant Microbe Interact **14:** 225-233

Jarausch B, Fuchs A, Schwind N, Krczal G, Jarausch W (2007) Cacopsylla picta as most important vector for Candidatus Phytoplasma mali in Germany and neighbouring regions. Bulletin of insectology **60:** 189-190

Jarausch B, Schwind N, Jarausch W, G. aK (2003) First Report of *Cacopsylla picta* as a Vector of Apple Proliferation Phytoplasma in Germany. Plant disease **87:** 101

Jarausch W, Lansac M, Bliot C, and Dosba F (1999a) Phytoplasma transmission by in vitro graft inoculation as basis for a preliminary screening method for resistance in fruit trees. Plant Pathology **48:** 283-287

Jarausch W, Lansac M, Dosba F (1996) Long-term maintenance of non-culturable apple proliferation phytoplasmas in their micropropagated natural host plant. Plant Pathology **45:** 778-786

Jarausch W, Peccerella T, Schwind N, Jarausch B, and Krczal G (2004b) Establishment of a quantitative real-time PCR assay for the quantification of apple proliferation phytoplasmas in plants an insects. Acta Horticulturae **657:** 415-420

Jarausch W, Saillard C, Dosba F, Bové JM (1994a) Differentiation of mycoplasmalike organisms (MLOs) in European fruit trees by PCR using specific primers derived from the sequence of a chromosomal fragment of the apple proliferation MLO. Appl Environ Microbiol **60:** 2916-2923

Jarausch W, Saillard C, Helliot B, Garnier M, Dosba F (2000) Genetic variability of apple proliferation phytoplasmas as determined by PCR-RFLP and sequencing of a non-ribosomal fragment. Mol Cell Probes **14:** 17-24

Kartte S, Seemüller E (1991) Histopathology off Apple Proliferation in *Malus* Taxa and Hybrids of Different Susceptibility. Journal of Phytopathology **131:** 149-160

Kartte S, Seemüller E (1991) Susceptibility of Grafted *Malus* Taxa and Hybrids to Apple Proliferation Disease. Journal of Phytopathology **131:** 137-148

Kirkpatrick BC (1992) Mycoplasma-like organisms: plant and invertebrate pathogens. *In* The prokaryotes, Ed 2nd edn. Springer, New York, pp 4050–4067

Kison H, Schneider B, Seemüller E (1994) Restriction fragment length polymorphism within the apple proliferation mycoplasmalike organism. Journal of Phytopathology **141:** 395-401

Kivioja T, Arvas M, Saloheimo M, Penttila M, Ukkonen E (2005) Optimization of cDNA-AFLP experiments using genomic sequence data. Bioinformatics **21:** 2573-2579

Klink VP, Matthews BF (2009) Emerging approaches to broaden resistance of soybean to soybean cyst nematode as supported by gene expression studies. Plant Physiol **151:** 1017-1022

Ko JH, Kim JH, Jayanty SS, Howe GA, Han KH (2006) Loss of function of COBRA, a determinant of oriented cell expansion, invokes cellular defence responses in Arabidopsis thaliana. J Exp Bot **57:** 2923-2936

Koltunow AM (1993) Apomixis: Embryo Sacs and Embryos Formed without Meiosis or Fertilization in Ovules. Plant Cell **5:** 1425-1437

Krczal G, Krczal H, and Kunze L (1988) *Fieberella florii* (Stal), a vector of apple proliferation agent. Acta Horticulturae **235:** 99-105

Kube M, Schneider B, Kuhl H, Dandekar T, Heitmann K, Migdoll AM, Reinhardt R, Seemuller E (2008) The linear chromosome of the plant-pathogenic mycoplasma 'Candidatus Phytoplasma mali'. BMC Genomics **9:** 306

Kunze L (1989) Apple proliferation. *In* Virus and Virus-like Diseases of Pome Fruits and Simulating Noninfectious Disorders. P. R. Fridlund, ed. Wash State University Press, Pullman., pp 99-113

Lee I-M, Davis RE (1992) Mycoplasmas which infect plants and insects. *In* RN J. Maniloff, McElhansey, L.R. Finch, and J.B. Baseman, eds., ed, Mycoplasmas: Molecular Biology and Pathogenesis. American Society for Microbiology, Washington, D.C., pp 379-390

Lee IM, Davis RE, Gundersen-Rindal DE (2000) Phytoplasma: phytopathogenic mollicutes. Annu Rev Microbiol **54:** 221-255

Lee IM, Gundersen-Rindal DE, Davis RE, and Bartoszyk IM (1998b) Revised classification scheme of phytoplasmas based on RFLP analyses of 16S rRNA and ribosomal protein gene sequences. International Journal of Systematic Bacteriology **48:** 1153-1169

Lee IM, Gundersen-Rindal DE, Davis RE, Bottner KD, Marcone C, Seemuller E (2004a) 'Candidatus Phytoplasma asteris', a novel phytoplasma taxon associated with aster yellows and related diseases. Int J Syst Evol Microbiol **54:** 1037-1048

Lee IM, Gundersen DE, Davis RE, Chiykowski LN (1992) Identification and analysis of a genomic strain cluster of mycoplasmalike organisms associated with Canadian peach (eastern) X disease, western X disease, and clover yellow edge. J Bacteriol **174:** 6694-6698

Lee IM, Gundersen DE, Hammond RW, and Davis RE (1994) Use of mycoplasmalike organism (MLO) group-specific oligonucleotide primers for nested-PCR assays to detect mixed-MLO infections in a single host plant. Phytopathology **84:** 559-566

Lee IM, Martini M, Bottner KD, Dane RA, Black MC, and Troxclair N (2003) Ecological implications from a molecular analysis of phytoplasmas involved in an aster yellows epidemic in various crops in Texas. Phytopathology **93:** 1368-1377

Lee IM, Martini M, Marcone C, Zhu SF (2004b) Classification of phytoplasma strains in the elm yellows group (16SrV) and proposal of 'Candidatus Phytoplasma ulmi' for the phytoplasma associated with elm yellows. Int J Syst Evol Microbiol **54:** 337-347

Lee KW, Chen PW, Lu CA, Chen S, Ho TH, Yu SM (2009) Coordinated responses to oxygen and sugar deficiency allow rice seedlings to tolerate flooding. Sci Signal **2:** ra61

Lee LG, Connell CR, Bloch W (1993) Allelic discrimination by nick-translation PCR with fluorogenic probes. Nucleic Acids Res **21:** 3761-3766

Lepka P, Stitt M, And Seemüller E (1999) Effect of phytoplasmal infection on concentration and translocation of carbohydrates and amino acids in periwinkle and tobacco. Physiological and Molecular Plant pathology **55:** 59-68

Lin F, Kong ZX, Zhu HL, Xue SL, Wu JZ, Tian DG, Wei JB, Zhang CQ, Ma ZQ (2004) Mapping QTL associated with resistance to Fusarium head blight in the Nanda2419 x Wangshuibai population. I. Type II resistance. Theor Appl Genet **109:** 1504-1511

Lin F, Xue SL, Zhang ZZ, Zhang CQ, Kong ZX, Yao GQ, Tian DG, Zhu HL, Li CJ, Cao Y, Wei JB, Luo QY, Ma ZQ (2006) Mapping QTL associated with resistance to Fusarium head blight in the Nanda2419 x Wangshuibai population. II: type I resistance. Theor Appl Genet **112:** 528-535

Lindqvist Y, Branden CI (1989) The active site of spinach glycolate oxidase. J Biol Chem **264:** 3624-3628

Livak KJ, Flood SJ, Marmaro J, Giusti W, Deetz K (1995) Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. PCR Methods Appl **4:** 357-362

Lorenz KH, Dosba F, Poggi-Pollini C, Llacer G, Seemüller E (1994) Phytoplasma diseases on *Prunus* species in Europe are caused by genetically similar organisms. Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz **101:** 567-575

Mandal P, Sinha Babu SP, Mandal NC (2005) Antimicrobial activity of saponins from Acacia auriculiformis. Fitoterapia **76:** 462-465

Manosalva PM, Davidson RM, Liu B, Zhu X, Hulbert SH, Leung H, Leach JE (2009) A germin-like protein gene family functions as a complex quantitative trait locus conferring broad-spectrum disease resistance in rice. Plant Physiol **149:** 286-296

Marcone C, Neimark H, Ragozzino A, Lauer U, and Seemüller E (1999) Chromosome Sizes of Phytoplasmas Composing Major Phylogenetic Groups and Subgroups. Phytopathology **89:** 805-810

Martini M, Lee IM, Bottner KD, Zhao Y, Botti S, Bertaccini A, Harrison NA, Carraro L, Marcone C, Khan AJ, Osler R (2007) Ribosomal protein gene-based phylogeny for finer differentiation and classification of phytoplasmas. Int J Syst Evol Microbiol **57:** 2037-2051

Marwitz R, Petzold H, und Özel M (1974) Untersuchungen zur Übertragbarkeit des möglichen Erregers der Triebsucht des Apfels auf einen krautigen Wirt. Phytopath. Z.**:** 85-91

Mattedi L, Forno F, Cainelli C, Grando MS, and Jarausch W (2008) Research on *Candidatus* Phytoplasma mali transmission by insect vectors in Trentino. Acta Horticulturae **781:** 369-374

Mattedi L, Forno F, Cainelli C, Grando MS, Jarausch W (2007) Transmission of Candidatus Phytoplasma mali by psyllid vectors in Trentino. . IOBC/WPRS Bulletin **30:** 267-272

Maurino VG, Peterhansel C (2010) Photorespiration: current status and approaches for metabolic engineering. Curr Opin Plant Biol **13:** 249-256

Maust BE, Espadas F, Talavera C, Aguilar M, Santamaria JM, and Oropeza C (2003) Changes in Carbohydrate Metabolism in Coconut Palms Infected with the Lethal Yellowing Phytoplasma. Phytopathology **93:** 976-981

Mayer CJ, Jarausch B, Jarausch W, Jelkmann W, Vilcinskas A, Gross J (2009) Cacopsylla melanoneura has no relevance as vector of apple proliferation in Germany. Phytopathology **99:** 729-738

McCoy RE, Caudwell A, Chang CJ, Chen TA, Chiykowski LN, Cousin MT, Dale JL, De Leeuw GTN, Golino DA, Hackett KJ, Kirkpatrick BC, Marwitz R, Petzold H, Sinha RC, Sugiura M, Withcomb RF, Yang IL, Zhu BM, Seemüller E (1989) Plant diseases associated with mycoplasma-like organisms. *In* WaT eds., ed, The Mycoplasmas, Vol V. Academic press, San Diego CA, USA, pp 545-640

Mehle N, Brzin J, Boben J, Hren M, Frank J, Petrovic N, Gruden K, Dreo T, Zezlina I, Seljak G, and Ravnikar M (2007) First report of '*Candidatus* Phytoplasma mali' in *Prunus avium*, *P. armeniaca* and *P. domestica*. Plant Pathology **56:** 721

Miemyk J (1997) The 70 kDa stress-related proteins as molecular chaperones. Trends Plant Sci **2:** 180-187

Mir G, Domenech J, Huguet G, Guo WJ, Goldsbrough P, Atrian S, Molinas M (2004) A plant type 2 metallothionein (MT) from cork tissue responds to oxidative stress. J Exp Bot **55:** 2483-2493

Miyata S, Furuki K, Sawayanagi T, Oshima K, Kuboyama T, Tsuchizaki T, Ugaki M, and Namba S (2002) Gene Arrangement and Sequence of str Operon of Phytoplasma Resemble Those of Bacillus More Than Those of Mycoplasma. Journal of General Plant Pathology **68:** 62-67

Montrichard F, Renard M, Alkhalfioui F, Duval FD, Macherel D (2003) Identification and differential expression of two thioredoxin h isoforms in germinating seeds from pea. Plant Physiol **132:** 1707-1715

Morton A, Davies DL, Blomquist CL, and Barbara DJ (2003) Characterization of homologues of the apple proliferation immunodominant membrane protein gene from three related phytoplasmas. . Molecular Plant Pathology **4:** 109-114

Moser C, Gatto P, Moser M, Pindo M, Velasco R (2004) Isolation of functional RNA from small amounts of different grape and apple tissues. Mol Biotechnol **26:** 95-100

Murashige T., Bitters W.P., Rangan T.S., Nauer E.M., Roistacher C., P.B. H (1972) A technique of shoot apex grafting and its utilization towards recovering virus-free citrus clones. HortScience**:** 118-119

Musetti R, di Toppi LS, Ermacora P, Favali MA (2004) Recovery in Apple Trees Infected with the Apple Proliferation Phytoplasma: An Ultrastructural and Biochemical Study. Phytopathology **94:** 203-208

Musetti R, Favali MA, Pressacco L (2000) Histopathology and polyphenol content in plants infected by phytoplasmas. Cytobios **102:** 133-147

Musetti R, Sanità di Toppi L, Martini M, Ferrini F, Loschi A, Favali MA, Osler R (2005) Hydrogen peroxide localization and antioxidant status in the recovery of apricot plants from European ston fruit yellows. European Journal of Plant Pathology **112:** 53-61

Neill S, Desikan R, Hancock J (2002) Hydrogen peroxide signalling. Curr Opin Plant Biol **5:** 388-395

Neill SJ, Desikan R, Clarke A, Hurst RD, Hancock JT (2002) Hydrogen peroxide and nitric oxide as signalling molecules in plants. J Exp Bot **53:** 1237-1247

Nicot N, Hausman JF, Hoffmann L, Evers D (2005) Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. J Exp Bot **56:** 2907-2914

Nolan T, Hands RE, Bustin SA (2006) Quantification of mRNA using real-time RT-PCR. Nat Protoc **1:** 1559-1582

Nybom H, Esselink GD, Werlemark G, Leus L, Vosman B (2006) Unique genomic configuration revealed by microsatellite DNA in polyploid dogroses, Rosa sect. Caninae. J Evol Biol **19:** 635-648

Ortiz E, Estrada G, Lizardi PM (1998) PNA molecular beacons for rapid detection of PCR amplicons. Mol Cell Probes **12:** 219-226

Oshima K, Kakizawa S, Nishigawa H, Jung HY, Wei W, Suzuki S, Arashida R, Nakata D, Miyata S, Ugaki M, Namba S (2004) Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma. Nat Genet **36:** 27-29

Palma JM, Corpas FJ, del Rio LA (2009) Proteome of plant peroxisomes: new perspectives on the role of these organelles in cell biology. Proteomics **9:** 2301-2312

Pedrazzoli F, Ciccotti AM, Bianchedi PL, Salvadori A, Zorer R (2008) Seasonal colonisation behaviour of *Candidatus* phytoplasma mali in apple trees in Trentino. Acta Horticulturae **781:** 483-489

Peirson SN, Butler JN, Foster RG (2003) Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis. Nucl. Acids Res. **31:** e73-

Perica MC, Lepedus H, Music MS (2007) Effect of indole-3-butyric acid on phytoplasmas in infected Catharanthus roseus shoots grown in vitro. FEMS Microbiol Lett **268:** 171-177

Petzold H, Marwitz R (1976) Versuche zur Infektion von Apfelbäumen mit dem möglichen Erreger der triebsucht des Apfels. Phytopath. Z.**:** 365-369

Phillips DR, Rasbery JM, Bartel B, Matsuda SP (2006) Biosynthetic diversity in plant triterpene cyclization. Curr Opin Plant Biol **9:** 305-314

Polesani M, Desario F, Ferrarini A, Zamboni A, Pezzotti M, Kortekamp A, Polverari A (2008) cDNA-AFLP analysis of plant and pathogen genes expressed in grapevine infected with Plasmopara viticola. BMC Genomics **9:** 142

Pose D, Castanedo I, Borsani O, Nieto B, Rosado A, Taconnat L, Ferrer A, Dolan L, Valpuesta V, Botella MA (2009) Identification of the Arabidopsis dry2/sqe1-5 mutant reveals a central role for sterols in drought tolerance and regulation of reactive oxygen species. Plant J **59:** 63-76

Pracros P, Renaudin J, Eveillard S, Mouras A, Hernould M (2006) Tomato flower abnormalities induced by stolbur phytoplasma infection are associated with changes of expression of floral development genes. Mol Plant Microbe Interact **19:** 62-68

Ramakers C, Ruijter JM, Deprez RH, Moorman AF (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci Lett **339:** 62-66

Rasbery JM, Shan H, LeClair RJ, Norman M, Matsuda SP, Bartel B (2007) Arabidopsis thaliana squalene epoxidase 1 is essential for root and seed development. J Biol Chem **282:** 17002-17013

Rui D, Ciferri R, e Refatti E (1950) La virosi degli "scopazzi del melo" nel Veronese. Not. Mal. Piante**:** 7-11

Ruijter JM, Ramakers C, Hoogaars WM, Karlen Y, Bakker O, van den Hoff MJ, Moorman AF (2009) Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Res **37:** e45

Saglio P, L'hospital M, Laflèche D, Dupont G, Bove JM (1973) *Spiroplasma citri* gen. and sp.n.: a mycoplasma-like organism associated with "stubborn" disease of citrus. International Journal of Systematic Bacteriology **23:** 191-204

Saglio P, Laflèche D, Bonissol C, Bove JM (1971) Isolement et culture *in vitro* des mycoplasmes associés au stubborn des agrumes et leur observation au microscope électronique. Comptes rendus de l'Academie des sciences Paris **272:** 1387-1390

Schaper U, and Seemüller E (1984) Recolonization of the stem of apple proliferation and pear decline-diseased trees by the casual organisms in spring. Journal of Plant Disease and Protection **91:** 608-613

Schaper U, Seemüller E (1982) Conditions of the Phloem and the Persistence of Mycoplasmalike Organisms Associated with Apple Proliferation and Pear Decline. Phytopathology **72:** 736-742

Schindelman G, Morikami A, Jung J, Baskin TI, Carpita NC, Derbyshire P, McCann MC, Benfey PN (2001) COBRA encodes a putative GPI-anchored protein, which is polarly localized and necessary for oriented cell expansion in Arabidopsis. Genes Dev **15:** 1115-1127

Schmid G (1965) Five and more years of observations on the proliferation virus of apples in the field. Zast.Bilja **85-88:** 285-289

Schmidt H (1988) Criteria and procedures for evaluating apomictic rootstocks for apple. Hort.Science **23:** 104-107

Schneider B, Ahrens U, Kirkpatrick BC, Seemüller E (1993) Classification of plant-pathogenic mycoplasma-like organisms using restriction-site analysis of PCR-amplified 16S rDNA. Journal of General Microbiology **139:** 519-527

Schneider B, Gibb KS, Seemuller E (1997) Sequence and RFLP analysis of the elongation factor Tu gene used in differentiation and classification of phytoplasmas. Microbiology **143 (Pt 10):** 3381-3389

Schneider B, Seemuller E (1994) Presence of Two Sets of Ribosomal Genes in Phytopathogenic Mollicutes. Appl Environ Microbiol **60:** 3409-3412

Schneider B, Seemüller E (2009) Strain differentiation of *Candidatus* phytoplasma mali by SSCP- and sequence analysis of the *hfl*B gene. Journal of Plant Pathology **91:** 103-112

Schwachtje J, Minchin PE, Jahnke S, van Dongen JT, Schittko U, Baldwin IT (2006) SNF1-related kinases allow plants to tolerate herbivory by allocating carbon to roots. Proc Natl Acad Sci U S A **103:** 12935-12940

Seemüller E (1976) Investigations to demonstrate mycoplasmalike organisms in diseased plants by fluorescence microscopy. Acta Horticulturae **67:** 109-112

Seemüller E, Garnier M, and Schneider B (2002) Mycoplasmas of plants and insects. *In* SRR Herrmann, ed, Molecular Biology and Pathogenicity of Mycoplasmas. Kluwer Academic/Plenum Publishers, New York, pp 91-115

Seemüller E, Kunze L, Schaper U (1984) Colonization behaviour of MLO, and symptom expression of proliferation-diseased apple trees and decline-diseased pear trees over a period of several years. Journal of Plant Disease and Protection **91:** 525-532

Seemüller E, Marcone C, Lauer U, Ragozzino A, Goschl M (1998) Current status of molecular classification of the phytoplasmas. Journal of Plant Pathology **80:** 3-26

Seemüller E, Moll E, Schneider B (2008) Apple proliferation resistance of Malus sieboldii -based rootstocks in comparison to rootstocks derived from other Malus species. European Journal of Plant Pathology **121:** 109-119

Seemüller E, Moll E, Schneider B (2009) Pear decline resistance in progenies of *Pyrus* taxa used as rootstocks. Eur J Plant Path **123:** 217-223

Seemüller E, Schneider B (2004) 'Candidatus Phytoplasma mali', 'Candidatus Phytoplasma pyri' and 'Candidatus Phytoplasma prunorum', the causal agents of apple proliferation, pear decline and European stone fruit yellows, respectively. Int J Syst Evol Microbiol **54:** 1217-1226

Seemüller E, Schneider B (2007) Differences in virulence and genomic features of strains of 'Candidatus Phytoplasma mali', the apple proliferation agent. Phytopathology **97:** 964-970

Seemüller E, Schneider B, Mäurer R, Ahrens U, Daire X, Kison H, Lorenz K-H, Firrao G, Avinent L, Sears BB, E. aS (1994) Phylogenetic classification of phytopathogenic mollicutes by sequence analysis of 16S ribosomal DNA. International Journal of Systematic Bacteriology **44:** 440-446

Sehnke PC, Rosenquist M, Alsterfjord M, DeLille J, Sommarin M, Larsson C, Ferl RJ (2002) Evolution and isoform specificity of plant 14-3-3 proteins. Plant Mol Biol **50:** 1011-1018

Singh R, Maganti RJ, Jabba SV, Wang M, Deng G, Heath JD, Kurn N, Wangemann P (2005) Microarray-based comparison of three amplification methods for nanogram amounts of total RNA. Am J Physiol Cell Physiol **288:** C1179-1189

Smart CD, Schneider B, Blomquist CL, Guerra LJ, Harrison NA, Ahrens U, Lorenz KH, Seemuller E, Kirkpatrick BC (1996) Phytoplasma-specific PCR primers based on sequences of the 16S-23S rRNA spacer region. Appl Environ Microbiol **62:** 2988-2993

Staden R, Judge P. D, Bonfield K. J (2003) Managing Sequencing Projects in the GAP4 Environment. Introduction to Bioinformatics. A Theoretical and Practical Approach. Human Press Inc., Totawa

Stamburski C, Renaudin J, Bove JM (1991) First step toward a virus-derived vector for gene cloning and expression in spiroplasmas, organisms which read UGA as a tryptophan codon: synthesis of chloramphenicol acetyltransferase in Spiroplasma citri. J Bacteriol **173:** 2225-2230

Stankiewicz-Kosyl M, Pitera E, Gawronski SW (2005) Mapping QTL involved in powdery mildew resistance of the apple clone U 211. Plant breeding **124:** 63-66

Stolting KN, Gort G, Wust C, Wilson AB (2009) Eukaryotic transcriptomics in silico: optimizing cDNA-AFLP efficiency. BMC Genomics **10:** 565

Sung D, Kaplan F, and Charles L. Guy CL (2001) Plant Hsp70 molecular chaperones: Protein structure, gene family, expression and function. Physiologia Plantarum **113:** 443-451

Tedeschi R, Alma A (2004) Transmission of apple proliferation phytoplasma by Cacopsylla melanoneura (Homoptera: Psyllidae). J Econ Entomol **97:** 8-13

Tedeschi R, Alma A (2006) Fieberiella florii (Homoptera:Auchenorrhyncha) as a Vector of "Candidatus Phytoplasma mali". Plant disease **90:** 284-290

Tedeschi R, Bosco D, and Alma A (2002) Population dynamics of *Cacopsylla melanoneura* (Homoptera:Psyllidae), a vector of apple proliferation phytoplasma in northwestern Italy. J. Econ. Entomol **95:** 544-551

Tran-Nguyen LT, Kube M, Schneider B, Reinhardt R, Gibb KS (2008) Comparative genome analysis of "Candidatus Phytoplasma australiense" (subgroup tuf-Australia I; rp-A) and "Ca. Phytoplasma asteris" Strains OY-M and AY-WB. J Bacteriol **190:** 3979-3991

Tyagi S, Kramer FR (1996) Molecular beacons: probes that fluoresce upon hybridization. Nat Biotechnol **14:** 303-308

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol **3:** RESEARCH0034

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, et al. (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res **23:** 4407-4414

Wan H, Zhao Z, Qian C, Sui Y, Malik AA, Chen J (2010) Selection of appropriate reference genes for gene expression studies by quantitative real-time polymerase chain reaction in cucumber. Anal Biochem **399:** 257-261

Wang E, Miller LD, Ohnmacht GA, Liu ET, Marincola FM (2000) High-fidelity mRNA amplification for gene profiling. Nat Biotechnol **18:** 457-459

Wang X, Liu W, Chen X, Tang C, Dong Y, Ma J, Huang X, Wei G, Han Q, Huang L, Kang Z (2010) Differential gene expression in incompatible interaction between wheat and stripe rust fungus revealed by cDNA-AFLP and comparison to compatible interaction. BMC Plant Biol **10:** 9

Wang X, Tang C, Zhang G, Li Y, Wang C, Liu B, Qu Z, Zhao J, Han Q, Huang L, Chen X, Kang Z (2009) cDNA-AFLP analysis reveals differential gene expression in compatible interaction of wheat challenged with Puccinia striiformis f. sp. tritici. BMC Genomics **10:** 289

Weintraub PG, Beanland L (2006) Insect vectors of phytoplasmas. Annu Rev Entomol **51:** 91-111

Weisburg WG, Tully JG, Rose DL, Petzel JP, Oyaizu H, Yang D, Mandelco L, Sechrest J, Lawrence TG, Van Etten J, et al. (1989) A phylogenetic analysis of the mycoplasmas: basis for their classification. J Bacteriol **171:** 6455-6467

Whitcombe D, Theaker J, Guy SP, Brown T, Little S (1999) Detection of PCR products using self-probing amplicons and fluorescence. Nat Biotechnol **17:** 804-807

Wittwer CT, Herrmann MG, Gundry CN, Elenitoba-Johnson KS (2001) Real-time multiplex PCR assays. Methods **25:** 430-442

Wittwer CT, Herrmann MG, Moss AA, Rasmussen RP (1997) Continuous fluorescence monitoring of rapid cycle DNA amplification. Biotechniques **22:** 130-131, 134-138

Wong HL, Sakamoto T, Kawasaki T, Umemura K, Shimamoto K (2004) Down-regulation of metallothionein, a reactive oxygen scavenger, by the small GTPase OsRac1 in rice. Plant Physiol **135:** 1447-1456

Woo TH, Patel BK, Cinco M, Smythe LD, Symonds ML, Norris MA, Dohnt MF (1998) Real-time homogeneous assay of rapid cycle polymerase chain reaction product for identification of Leptonema illini. Anal Biochem **259:** 112-117

Xu R, Fazio GC, Matsuda SP (2004) On the origins of triterpenoid skeletal diversity. Phytochemistry **65:** 261-291

Yang S, Yang H, Grisafi P, Sanchatjate S, Fink GR, Sun Q, Hua J (2006) The BON/CPN gene family represses cell death and promotes cell growth in Arabidopsis. Plant J **45:** 166-179

Zhang J, Byrne CD (1999) Differential priming of RNA templates during cDNA synthesis markedly affects both accuracy and reproducibility of quantitative competitive reverse-transcriptase PCR. Biochem J **337 (Pt 2):** 231-241

Aknowledgements

This work is the result of a “long” scientific and personal path. It is difficult to find the right words to express my gratitude to the people that allowed me to reach this goal: Dr. R.V. for giving me years ago the possibility to start this project supporting me and my freedom in experimenting ideas and approaches and preciously “imprinting” my first steps in the research; Dr. W.J for the invaluable scientific support, the help at the beginning of my stay in Germany, the efforts and the patience during the last years and in the final rush for reaching this objective.

I have also to thank Prof. Dr. G.K. for giving me the opportunity to work at the A. Institute.

I am grateful to Dr. K.B. for the valuable scientific discussions and more for his precious support and friendship during all these years

A special thought goes to my brother D. and the “adoptive” brother F. for the amazingly funny years and all our adventures and to Thanassis for the very “southern thinking” period.

Here I want to say how much I am grateful to three very special women: A. to have disclosed my view and making me understand that life is not only science, I will never forget that; E. for having always one ear and a word especially in the dark moments; I. for bringing me back a hidden way of thinking

Finally, no words can explain my gratitude to my parents. For all what I am and did I have to thank you, always being at my side.

General statement

I declare that I am the sole author of this submitted dissertation and that I did not make use of any sources or help apart from those specifically referred to. Experimental data or material collected from or produced by other persons is made easily identifiable.

I also declare that I did not apply for permission to enter the examination procedure at another institution and that the dissertation is neither presented to any other faculty nor used in its current or any other form in another examination

Curriculum Vitae

Personal Information

Name: Moser
First name: Mirko
Marital status: not married
Nationality: Italian
Birth date: 29/01/1976
Birth city: Bolzano, Italy

Education

Primary school: 1982-1987

Middle school: 1987-1990

Secondary school: 1990-1995; Technical-Industrial Institute “M.Buonarroti”, Trento

Title of qualified industrial chemist – Degree: July 1995 Note: 50/60

Studies: 1995-2001 University of Verona

Degree in Agroindustrial Biotechnology

- Duration: 5 years
- Final evaluation : 102/110
- Experimental thesis title: Study in *Oenococcus oeni* strains of growth kinetic, protein synthesis rate and enzymatic activity, in relation to the maleic acid metabolism

Januar 2004 – December 2009: Ph. D. work at AlPlanta Institute For Plant research – Neustadt an der Weinstrasse

Languages

English: written and spoken ; German written and spoken ; Italian (Mother Language)

Publications

Moser C, Gatto P, Moser M, Pindo M, Velasco R (2004) Isolation of functional RNA from small amounts of different grape and apple tissues. Mol Biotechnol 26: 95-100

Moser M, Jarausch W, Seemüller E, Velasco R (2005) Individuazione di marker cDNA-AFLP in portainnesti di melo resistenti ad apple proliferation (AP). Petria 15: 127-128

Jarausch W, Bisognin C, Ciccotti A M, Moser M & Grando S (2006) Etablierung eines in vitro-screening-Systems für Apfeltriebsucht-resistente Unterlagen-Genotypen. Mitt. Biol. Bundesanst. Land- und Forstwirtschaft. 400: 205

Moser M, Sprenger C, Bisognin C, Velasco R, Jarausch W (2007) Gene expression study in different 'Ca. Phytoplasma mali'-infected micropropagated Malus genotypes. Department of Agroenvironmental Sciences and Technologies, Bologna, Italy, Bulletin of Insectology 60, 2, pp 207-208

Bisognin C, Ciccotti A M, Salvadori A, Moser M, Grando M S, Jarausch W (2007) Effect of different strains of 'Candidatus Phytoplasma mali' in susceptible and resistant Malus genotypes inoculated by in vitro grafting. Department of Agroenvironmental Sciences and Technologies, Bologna, Italy, Bulletin of Insectology, 60, 2, pp 205-206

Bisognin C, Ciccotti A M, Moser M, Grando M S and Jarausch W (2008) Establishment of an in vitro screening system of apple proliferation-resistant rootstock genotypes based on micrografting. Acta Hort. (ISHS) 781:375-380

Moser M, Sprenger C, Bisognin C, Velasco R, Jarausch W (2008) Analyse der Genexpression in Ca. Phytoplasma mali-infizierten in vitro vermehrten Malus-Genotypen. Mitt. Julius-Kühn-Institut 417: 324

Bisognin B, Ciccotti A, Salvadori A, Moser M, Grando MS, and Jarausch W (2008) In vitro screening for resistance to apple proliferation in Malus ssp. Plant Pathology **57:** 1163-1171