Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN
SCIENZE E TECNOLOGIE AGRARIE, AMBIENTALI E ALIMENTARI

Ciclo XXX

Settore Concorsuale: 07/B2

Settore Scientifico Disciplinare: AGR/03

TOWARDS THE IDENTIFICATION OF THE GENE CONFERRING RESISTANCE AGAINST THE ROSY APPLE APHID (*Dysaphis plantaginea*) IN APPLE

Presentata da: Michela Dall'Agata

Coordinatore Dottorato

Prof. Giovanni Dinelli

Supervisore

Prof. Stefano Tartarini

Co-Supervisore

Dr. Charles-Eric Durel

Esame finale anno 2018

Table of contents

Abstract

Most of the apple cultivars are susceptible to rosy apple aphid (RAA, *Dysaphis plantaginea*) but resistance have also been described in apple germplasm laying the basis for the development of new resistant cultivars by breeding. The cultivar Florina, a *Malus floribunda* #821 derivative, is resistant to RAA and a single resistance gene (*Dp-fl*) has been mapped in a 330 Kb region on linkage group 8. In this work, a chromosome walking was performed by using a Florina Bacterial Artificial Chromosome (BAC) library to identify candidate resistance genes. A minimum tiling path of BACs covering regions from both 'resistant' and 'susceptible' chromosomes were identified and a 279 Kb resistance locus was fully sequenced. Through the development of new polymorphic markers, the resistance locus-mapping interval was narrowed down to a physical region of 56 Kb. During the fine-mapping process, two genotype-phenotype incongruences were identified. A single candidate gene, predicted to code for a protein similar to the Quirky gene of *Arabidopsis*, was identified. To understand the role of this gene, a gene expression analysis was performed on both Florina and Golden Delicious and the Quirky gene was found to be more expressed at 72 hours after the infestation only in Golden while in Florina the expression was generally very low. To validate the gene function a genetic transformation of Gala and Florina was started. Finally, to confirm the identification of the resistance locus, large progenies derived from *Malus floribunda* were screened to identify more recombinants. This analysis extended the resistance region at the bottom of chromosome 8. Various genes putatively involved in defense response were also identified in the GDDH13 genome sequence. Thus, the presence of additional genes involved in RAA resistance cannot be excluded and a further step of chromosome walking would be necessary for the identification of new candidate genes.

Chapter 1 – General introduction

1.1 Economic importance and origin of cultivated apple

The apple tree belongs to the *Rosaceae,* a family including one-third of all flowering plants. The *Rosaceae* family includes commercially edible genera (i.e. *Prunus, Malus* and *Pyrus*) but also ornamental and invasive genera (i.e. *Chaenomeles*, *Crateagus*, *Pyracantha* and *Sorbus*). Apple is a deciduous tree cultivated worldwide as a fruit tree and is the most widely grown species in the *Malus* genus (Challice 1974). The cultivated apple belongs to the *Malus* × *domestica* (Borkh.)(Hummer and Janick, 2009) species and is the most economically important rosaceous species with annual world fruit production of about 89 million metric tons in 2016 (FAO Statistics). In last ten years the harvested area is rather stable but the total apple production is increasing (FAO Statistics, Figure 1.1).

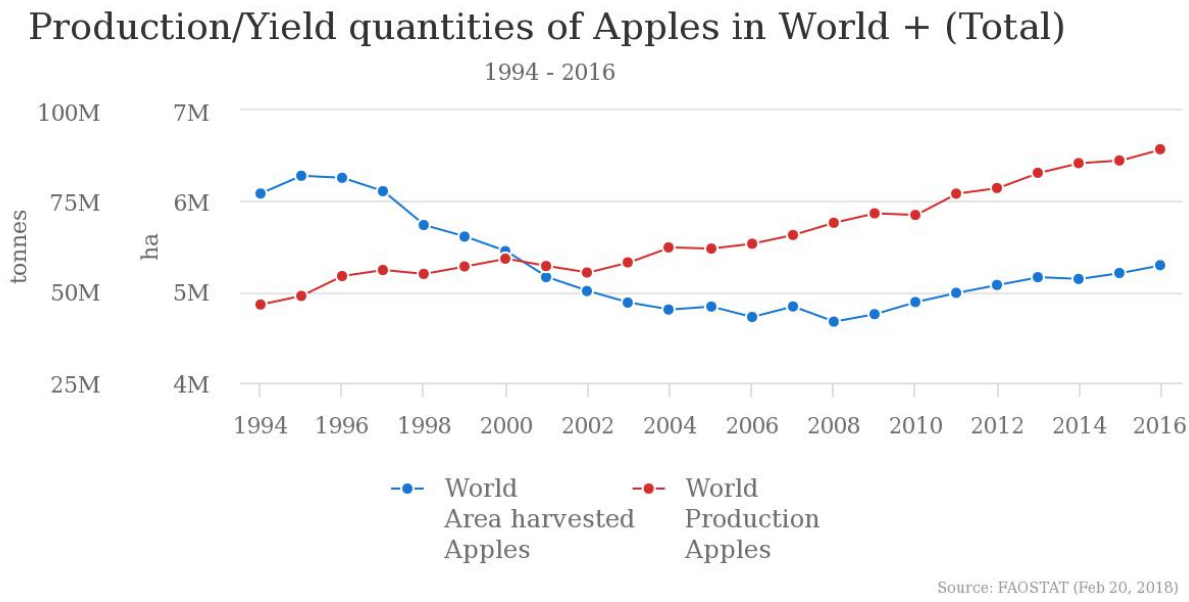

Figure 1.1 Production of apple (million metric tons) and harvested area (ha) in the world from 1994 till 2016 (FAO Statistics)

China is the largest apple producer which production is accounting for about 50% of the total world production, followed by USA, Poland and India. In 2016, Italy was the third largest apple producer in Europe with an amount of 2,455,616 tons after Poland and Turkey (FAO Statistics 2016, Table 1.1).

Table 1.1: Top ten apple producer in the world in 2016 (FAO statistics)

Country	Production (tons)
China	44,447,793
United States of America	4,649,323
Poland	3,604,271
Turkey	2,925,828
India	2,872,000
Iran	2,799,197
Italy	2,455,616
Russian Federation	1,843,544
France	1,819,762
Chile	1,759,421

The origin of apples was recently dated 21 million years ago in the Tian Shan mountain range (Daccord et al. 2017). From antiquity apples are cultivated in the temperate and subtropical climatic regions and in some mountain areas of tropical regions. Through its history of cultivation at least 10,000 apple cultivars were developed, many of which now are lost (Janick and Moore, 1996). Most of commercial cultivars have an ancient origin and a long propagation history, such as McIntosh (1800s), Jonathan (1820s), Cox's Orange Pippin (1830), Granny Smith (1860s), Delicious (1870s) and Golden Delicious (1890s) (Noiton and Alspach, 1996).

The haploid (x) chromosome number of apple is 17; while for most *Rosaceae* is 7, 8, or 9. The origin of this chromosome assessment could be explained with two hypotheses: i) an allopolyploidization between species related the existing *Spiraeoideae* (x = 9) and *Amygdaleoideae* (x = 8) subfamilies or ii) a within-lineage polyploidization event followed by chromosome rearrangements (Evans and Campbell, 2002).

This second hypothesis was first confirmed by the sequencing of the *Malus domestica* genome, of Golden Delicious (Velasco et al. 2010) and further supported by the high-quality *de novo* assembly of a Golden Delicious doubled-haploid tree (GDDH13) genome (Daccord et al. 2017). A large colinearity has been identified between pairs of apple chromosomes (i.e. 5/10, 13/16, 9/17, 8/15 and 11/3) or between large portions of chromosome (i.e. 1-2 and 7, 4-14 and 12, 7-15 and 1; 7-15 and 2). All these findings point out an ancient polyploid origin of apple with chromosomal rearrangements (Figure 1.2). Furthermore, approximately 2 million SNPs were detected in the Golden Delicious genome, confirming the highly heterozygous status of apple.

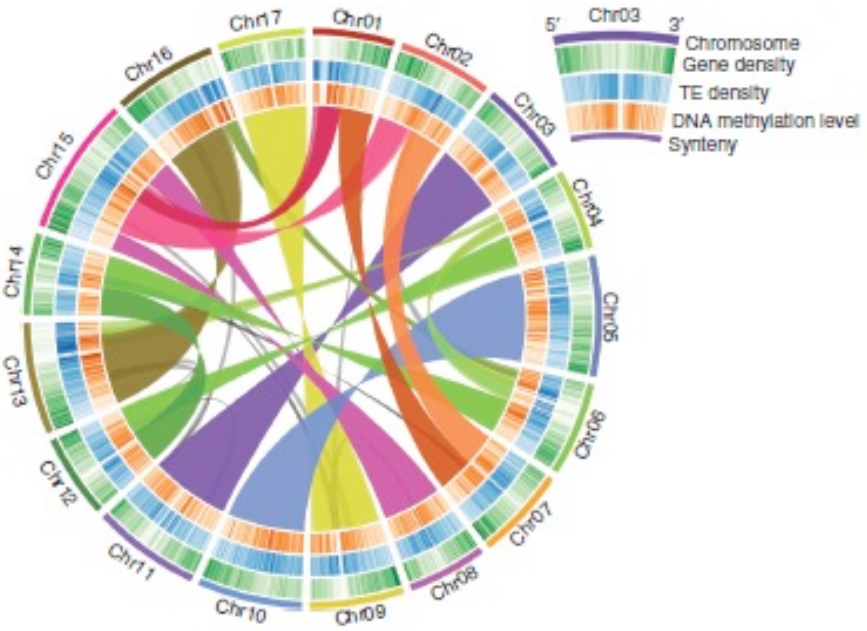

Figure 1.2 Representation of Golden Delicious double haploid apple genome (Daccorrd et al. 2017). The rings indicate chromosomes (Chr), in green is represented gene density, in blue transposable elements (TE) and in orange level of DNA methylation. Inside the figure is represented a map connecting homologous regions of the apple genome.

In the apple genome about 57,000 genes were firstly identified by Velasco et al. 2010, but the sequencing of the double haploid (Daccord et al. 2017) reduced the total number of genes to about 42,000 genes. Nevertheless, the number of genes is one of the highest of any plant genome studied to date. Different classes of apple genes were identified, some of which were greatly different in their degree of duplication, for example genes involved in metabolism of anthocyanins, flavonoids and terpenes.

The availability of the whole genome apple sequence anchored with more than 1,700 molecular markers to the apple genetic map already speed up the identification of genes involved in controlling traits of interest (i.e. resistances and fruit quality traits). Other genetic tools such as bacterial artificial chromosome (BAC) libraries can be successfully used for cloning genes of interest (Vinatzer et al. 2001; Xu and Korban 2002; Han et al. 2007, Cova et al. 2015, Padmarasu et al. 2014). Recently, other tools for genetic analysis have been developed, including a 20 K apple Infinium® SNP chip Illumina chip (Bianco et al. 2014) and Axiom®Apple480K SNP genotyping array from Affymetrics (Bianco et al. 2016). These high-density chips have been already used for genetic mapping in various segregating progenies (Falginella et al. 2015; Di Pierro et al. 2016) and for association mapping of the flowering and ripening dates in apple (Urrestarazu et al. 2017).

All these developments in molecular genetics made it possible to really implement new breeding approaches devoted to the application of the Marker Assisted Selection (MAS) in the future breeding programs (Baumgartner et al. 2016).

Apple trees are affected by numerous pathogens and pests which every year are causing serious yield losses in apple production (McVay et al. 1993). The main diseases affecting apple in Europe are scab (*Venturia inaequalis*), powdery mildew (*Podosphaera leucotricha*), stem canker (*Neonectria ditissima*) and fire blight (*Erwinia amylovora*). Various insects and mites can also affect apple productions such as some sap-sucking insect (various aphid spp.) and moths, whose larvae can damage both leaves and fruits, as for example codling moth, *Cydia pomonella* (Conceiçao et al. 1999).

1.2 Apple aphids and control of aphid populations

Aphids are small sap-sucking insects and members of the superfamily *Aphidoidea*, *Sternorrhyncha* division of the order *Hemiptera*. They can migrate great distances, mainly through passive dispersal by riding on winds. Aphids are among the most destructive insect pests on cultivated plants in temperate regions (McGavin, 1993). They are capable of an extreme rapid increase in number by asexual reproduction. About 4,400 species are known, all included in the *Aphididae* family. Around 250 species cause serious damages for agriculture. Their characteristic is that they are specialized in feeding on the phloem of vascular plants which contains an abundance of simple sugars (produced in 'source' leaves by photosynthesis and transported through the phloem to the 'sinks' of plant growth), as well as nutrients and plant secondary metabolites, but also contains few essential amino acids. Aphids during their probing are known to transmit plant pathogenic viruses causing the detrimental effects due to the virus transmission that often exceed the direct effects of aphid feeding. Aphids extensively probe and salivate into host-plants and frequently inject saliva after initial sampling of the epidermal cell contents, or subsequent sampling of mesophyll cell contents (Figure 1.3, Zust and Agrawal 2016).

Among aphids that attack apple, the more destructive are green apple aphid (*Aphis pomi*), woolly apple aphid (*Eriosoma lanigerum*), rosy leaf curling aphid (*Dysaphis devecta*) and rosy apple aphid (*Dysaphis plantaginea*).

Green apple aphid has a widespread distribution in Europe, western Asia as far east Asia and Pakistan, North Africa and North America. *Aphis pomi* infests apple trees but can attack also other plants in the *Rosaceae* family including pear, quince and roses. It completes its life cycle on a single host species. Leaves carrying colonies may roll and curl, but will not discolor. The aphid occasionally feeds on immature apples, which then become malformed. Heavy infestation, especially of young leaves, can lead to stunting, and in extreme cases to permanently deformations (Hull and Grimm, 1983).

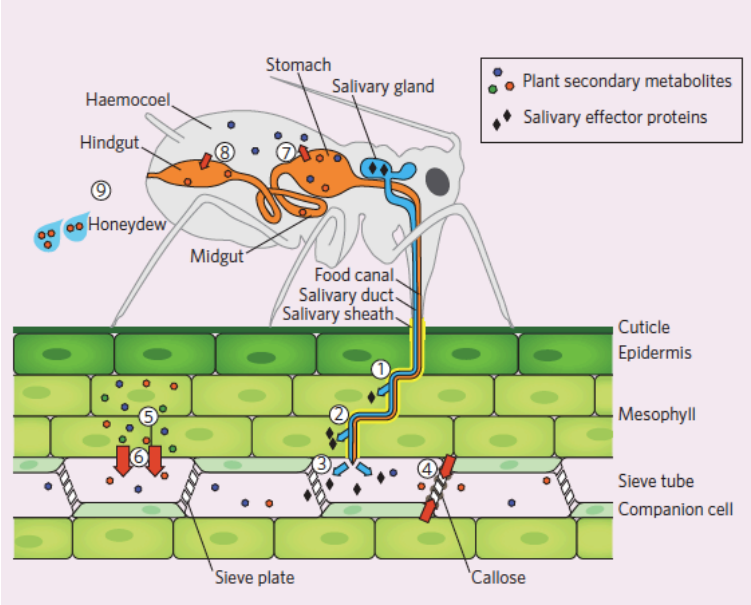

Figure 1.3: Schematic representation of aphid feeding and plant responses. Red arrows indicate key processes for the plant–aphid interaction. Aphids penetrate the apoplast with their stylet and move it between individual cells while exuding gelling saliva into the intercellular space (1), encasing the stylet in a salivary sheath and sealing off any cell leaks caused by the insertion process. During insertion, aphids puncture mesophyll cells and inject small amounts of watery saliva containing effector proteins (2) before sucking back some liquid to assess plant quality. After the phloem is reached, aphids alternate between sap ingestion and secretion of watery saliva containing effector proteins into the phloem (3) to prevent callose deposition at sieve plates leading to phloem sealing (4). Plant cells synthesize defensive secondary metabolites (5), of which a subset is transported into the phloem (6). Secondary metabolites are ingested during feeding, and may be taken up into the haemocoel by passive or active transport across the gut membrane (7). Secondary metabolites either accumulate in the haemocoel or are excreted back into the hindgut (8) and exuded with the aphid's honeydew together with all remaining metabolites (9) (Zust and Agraval, 2016).

Woolly apple aphid is a relatively small medium-sized aphid, is characterized by a reddish-brown body, a blood-red stain when crushed and fluffy, flocculent wax covering. Specialized dermal glands produce the characteristic fluffy or powdery wax, which gives *E. lanigerum* its characteristic 'woolly' appearance. Hibernating apterous virginoparae occurring on roots of apple are very dark green (Palmer, 1952; Blackman and Eastop 1994). *E. lanigerum* probably originated in eastern North America, but now it has a worldwide distribution, having been distributed mainly via apple rootstock (CIE, 1975). Woolly apple aphid is considered a phytosanitary risk in many regions, due to its root-dwelling habitat and its possible presence on imported apple rootstocks. *E. lanigerum* is found on apple, on which it can be a severe pest, and occasionally on certain other woody host plants in the *Rosaceae* family. It is restricted to apple in some areas where it has been introduced, for example in Australia (Asante, 1997). Is elsewhere found on species of *Crataegus*, *Sorbus* and *Cotoneaster* and also rarely on pear and *Cydonia* (Blackman and Eastop, 1994).

Dysaphis devecta (Wlk.) affects cultivated and ornamental apples causing severe leaf curl with conspicuous red galls. *Dysaphis devecta* remains all year on apple and has no alternative hosts. Sexual morphs appear before mid-summer, after only three parthenogenetic generations, and overwintering eggs are laid on apple. The occurrence of the aphid has been reported in Europe: Germany, United Kindom, Hungary and Italy. This aphid cause economic damages to apple crops if not controlled by routine aphicide applications. Sources of genetic host-plant resistance are readily available in apple varieties as for example Cox's Orange Pippin and Fiesta (Roche et al. 1997).

Rosy apple aphid (RAA) has been identified by different botanical names, in 1960s was *Sappaphis pomi*, then *Dysaphis plantaginea* (Passerini) became the definitive name. RAA is a medium-sized globe-shaped aphid; adults are pink-grey with a grey-white wax bloom. The antennae of apterae are at least as long as distance from the frons to the base of the siphunculi. The siphunculi of *D. plataginea* are quite long compared to the other *Dysaphis* species (Figure 1.4) The RAA is a dioecious species whose primary host until mid-summer is apple (sometimes quince, *Cydonia oblonga* Mill.) and its secondary host are herbaceous plants of the genus *Plantago* as *P. lanceolata* L. the preferred secondary host, but also *P. media* L. and *P. major* L. (Blommers et al. 2004).

Geographically, the rosy apple aphid is spread all over Asia, North Africa, North America and Europe including the whole Italian territory (Barbagallo et al. 1996). As in other apple growing areas, *D. plantaginea* is an economically important early-season pest in the apple orchards of Italy (Barbagallo et al, 1996). During their cycle, the apterous virginoparae settle usually locate on the adaxial side of leaves, causing severe damages as petal fall, abscission and deformation of growing shoots, but the most relevant damage is the deformation of the developing fruits that loss their economic value (Faccioli et al. 1985; Pasqualini et al. 1996). Because of the significant economic losses that it cause, RAA is considered one of the major insect pest of European apple orchards. Furthermore, large aphid populations may produce large amounts of honeydew that constitutes a growing ground for mold fungi, affecting final apple product. The damages are due to the salivary secretion released while probing intercellularly during the food-plant selection process and while feeding in the phloem. The saliva contains peroxidases, β-glucosidase and other potential signal-generation enzymes (Miles 1999). Signals arising from the phloem feeding are able to alter the expression of inducible plant physiological factors similar to those involved in defense against pathogens (Van Der Westhuizen et al. 1998; Fidantsef et al. 1999).

Figure 1.4 Golden Delicious leaf infested by *D. plantaginea* (RAA)

Many commercial orchards pursue a program of chemical sprays to maintain high fruit quality, tree health and high yields. The most common strategy to control *D. plantaginea* in conventional apple production, is the application of an aphicide during early spring as soon as fundatrices appear (Wyss et al. 1999). Usually, a treatment is applied just before blooming, then, a second one is applied after blooming when the first resulted not enough effective. Systemic insecticides belonging to different chemical groups (e.g. neonicotinoids and pyrethroids) are utilized (Cross et al, 2007). Control of aphid populations with repeated chemical applications could be difficult because of the reduction of beneficial organisms and the occurrence of reported resistance to insecticides (Angeli and Simoni, 2006; Delorme et al. 1998).

A new trend in orchard management is the reduction of pesticides in favor of the use of bio-control methods, including, for instance, the introduction of aphid natural predators to reduce their populations. Bio-control methods requires a good knowledge of the interaction between natural enemies and *D. plantaginea* populations. Aphids have natural enemies belonging to *Coccinellidae* family (i.e. *Adalia bipunctata* L. and *Coccinella decempunctata* L.), but other enemies are hoverfly larvae (*Syrphidae*) and aphid midgae larvae (*Aphidoletes aphidimyza* Rond.). Some parasitoids as wasps of *Ephedrus* species have also been reported. Among the RAA natural enemies in orchards appear in a chronological order, syrphids followed by coccinellids and earwings. Because of its early arrival in orchards, *Syrphids* seems to be a very efficient group of natural enemies against RAA, but its predatory effect is not sufficient to prevent damages (Miñarro et al. 2005; Dib et al. 2010; Dib et al. 2017). Therefore, the use of natural enemies is not sufficiently effective to prevent pest damage fully and reliably (Solomon et al. 2000).

Contrariwise, ant populations have a negative effect in the abundance of natural enemies and positively affect *D. plantaginea* abundance (Miñarro et al. 2010). Indeed, benefits in the relationship

between *D. plantaginea* and ants is reported in Stewart-Jones et al. work (2008) where is showed that in presence of ants a successful growing aphid population is attended, which is directly correlated with higher levels of apple damages at harvest.

An alternative aphid control method is the application of kaolin. This compound, generally used for prevention against solar injury, is applied repeatedly and it reduce significantly RAA populations (Wyss and Daniel 2004).

1.3 Plants resistance genes and aphid resistance

During their life plants need to defend themselves against different pathogens and pests (fungi, bacteria, viruses, invertebrates, and insects). Resistance (R) genes, that confer resistance to plants, are selected by breeders for the control of different diseases. In the "gene-for-gene" theory, for the interaction between plants and their pathogens two genes are needed: *R* genes in the plant and a corresponding avirulence (*Avr*) gene of the pathogen which express for proteins that are delivered directly into the plant cells during initial stage of infection. In the gene-for-gene relationship, a plant carrying a resistance gene can resist to a specific pathogen race carrying the corresponding effector (Figure 1.5).

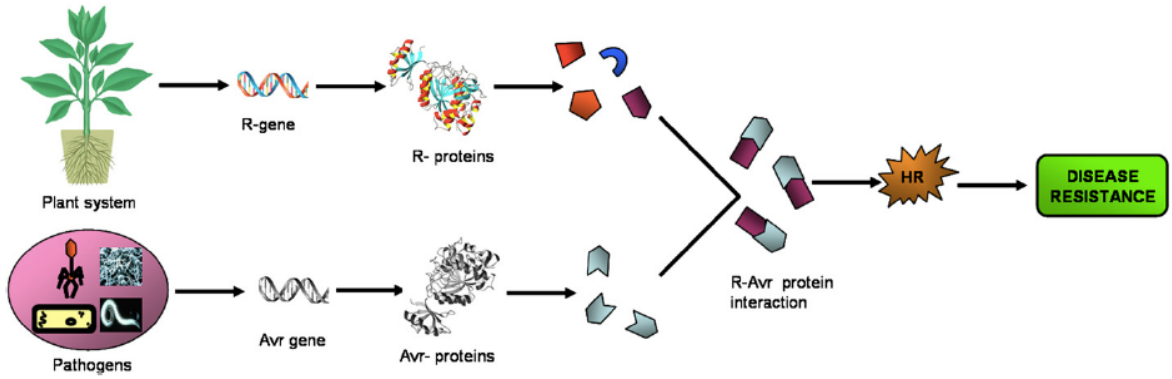

Figure 1.5 Plant pathogen interaction and development of disease resistance (Gururani et al. 2012)

The first step of this interaction is the recognition of the pathogen signal which often lead to the cell death phenotype that is typical of the hypersensitive response (HR) which is a form of programmed cell death. The signaling cascade behind the HR is triggered either when an appropriate disease resistance gene recognize a specific pathogen effector or by an elicitor of calcium ions from extracellular space and/or anion flux. This recognition determines an oxidative burst production by reactive oxygen intermediates (ROIs) and a defense gene activation, that finally results in the development of local and systemic disease resistances. Eight main classes of R genes differently divided on the base of their amino acid motif organization and their membrane spanning domains have been described (Figure 1.6). Leucine rich repeats (LRRs) represent one of the main components having an important role for the recognition specificity, thus these domains are present in the majority of the R proteins identified so far (Gururani et al. 2012).

Plants resistant to aphids recognize when an aphid feed and mount rapid defenses that are similar to those involved in response to mechanical damages (Zust and Agrawal 2016). In these plants, aphid salivary compounds may be recognized by plants and activate targeted defenses, including the induction of secondary metabolites and other mechanisms of resistance (Peck and Merwin 2010). Anatomical and chemical characters that avoid the availability for the aphids of the feeding sites and

the access to the plant phloem of these plants have been studied. Different phases have been studied to be important in the plant resistance against aphids. During the first stages of selection, plant volatiles can guide or disrupt aphid olfactory orientation to the plant. The plant acceptance depends also upon the aphid's ability to probe the leaf surface, penetrate some cells to 'taste' a leaf and reach the phloem. Then, at the phloem feeding sites, aphids secrete calcium-binding proteins which prevent stylet probing from clogging sieve elements. Depending on the presence/absence of specific chemical compounds the aphid growth, development, survival and fecundity may be affected (Smith and Chuang, 2014).

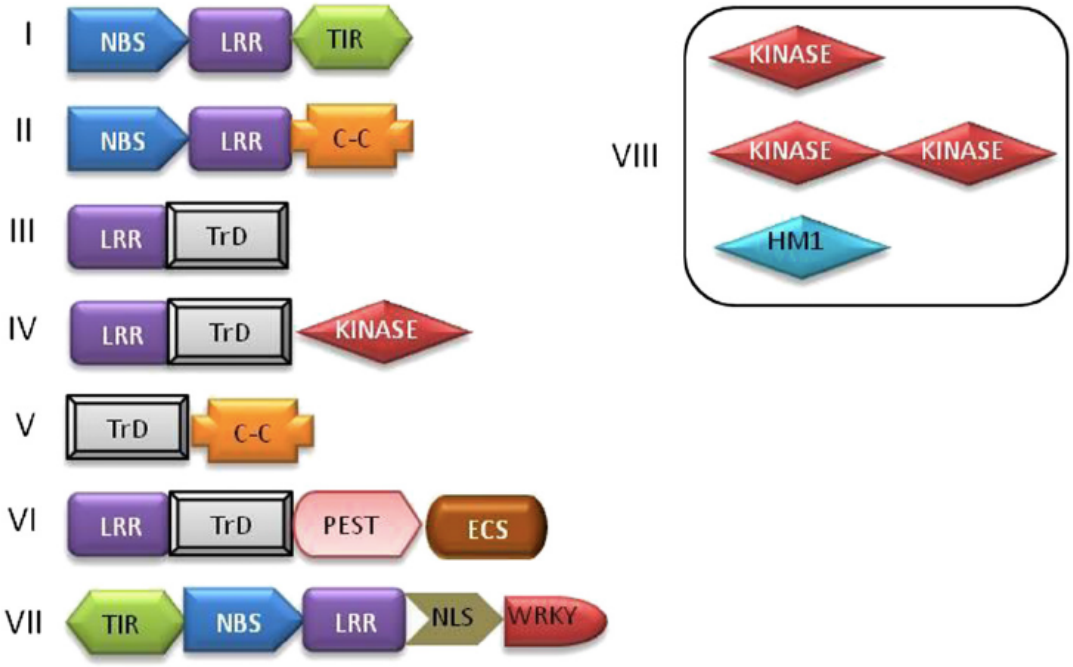

Figure 1.6 Major classes of resistance (R) genes and relarive protein domain in components. LRR – Leucine rich repeats; NBS – Nucleotide-binding site; TIRToll/Interleukin-1-receptors; C-C – Coiled coil; TrD – Trasmembrane domain; PEST – Protein degradation domain (proline-glycine-serine-threonine); ECS – Endocytosis cell signaling domain; NLS – Nuclear localization signal; WRKY – Amino acid domain; MH1 – *Helminthosporium carbonum* toxin reductase enzyme (Gururani et al. 2012).

In other words, insects-resistant plants can alter the relationship that an insect pest has with its host-plant. How the relationship between the insect and plant is affected depends on the kind of resistance. Three different resistance mechanisms to pests have been described: antibiosis, antixenosis (non-preference) and tolerance. Antixenosis is a non-preference plant that affect the way an insect pest perceives the desirability of the host plant. Non-preference plants either provide stimuli that are unattractive to the pest (e.g. color, odor, repellents, texture such as hairs or thick) or fail to provide stimuli that are attractive to the pest. In this way, non-preference plants affect the behavior of pests. Antibiosis is another type of resistance in which the host plant cause injury, death, change in

term of longevity, or reduced reproduction of the pest. Antibiosis plants, for example, may produce defensive compounds (called allelochemicals) that protect them from insects. These compounds may reduce growth, inhibit reproduction, alter physiology, delay maturation, or induce various physical or behavioral abnormalities in insects. Finally, some genotypes are simply tolerant to the injurious insects better than others. Tolerant cultivars may be exposed to the same pest populations as susceptible ones, but they do not suffer as much injury (Painter 1951). Resistance to aphids raises the level at which the economic injury level occurs, delay the need for insecticidal control and rarely need to elicit the occurrence of virulence. Resistance to many aphid species were studied in alfalfa, barley, maize, rice, rye, sorghum and wheat (Smith and Chuang, 2014). Antixenosis, antibiosis and tolerance are in relation each other and can be combined in plants, the interrelation of their effects results in different level of resistance. Another aspect that influence the interaction between plant and aphids is the asynchrony between host and pest. For example, eggs of *D. plantaginea* hatch in early spring (around 4,5^{o}C) simultaneously with the burst of the apple buds. However, apple cultivars show wide differences in flowering phenology, thus different pest fitness are reached in relation to synchronization between pest and host (Briggs and Alston, 1967; Knight and Alston, 1974; Miñarro and Dapena, 2014).

Knowledge of the aphid resistant plants interactions has proved to be extremely useful in both applied and basic studies, but determining the inheritance, the genetic location and the resistance gene products are essential in breeding for aphid resistance.

Studies using the model plant of *Arabidopsis* have contributed to our understanding of R-gene mediated plant defense, especially against pathogens, as well as the basal defense mechanisms against aphid feeding (Gururani et al. 2012). Plant populations developed from crosses of aphid-resistant and aphid-susceptible parents have been successfully used to map and link the loci of aphid resistance by using different types of molecular markers and developing chromosome maps including the resistance trait. In many cases, aphid resistance is monogenic and inherited as a dominant trait. In other cases, aphid resistance has been created using a natural mutualism, as for example the symbiosis between the fungus *Neotyphodium uninatum* and the forage grass *Lolium arundinaceum* (Schreb.) where a constitutive resistance is induced by the infection of the fungus that producing loline alkaloids (Smith and Chuang, 2014). Resistance also occurs as a result of polygenic quantitative traits. The use of molecular-marker-based linkage analysis and advanced computational software allow to map easily quantitative trait loci (QTL) in plants. QTLs related to aphid antixenosis and/or antibiosis and/or plant tolerance have been mapped in apple, barley, barrel medic, melon, potato, sorghum and wheat, which comprise approximately half of all mapped resistance genes (Table 1.2).

In Table 1.2 are pointed out a list of genes identified in some of the plants studied for their resistance to different aphid species. For each crop plant is indicated the aphid species studied, the

relative identified gene/s and the category of resistance (antibiosis, antixenosis or tolerance) (Smith and Chuang, 2014). Most of the genes studied have antibiosis as resistance mechanism. The majority of these genes showed have been studied in staple foods like wheat and rice. For example, in wheat, barley and rye resistance to green bug *Shizaphis graminum* is conferred by the single dominant *gb1* gene from *Triticum durum* (Boyko et al. 2004; Berzonsky et al. 2003). An example of a mapped dominant gene is also the *Dp-1* that provides resistance to *Dysaphis pyri* in pear (*Pyrus* spp.) (Evans et al. 2008).

Several aphid resistances are reported with a recessive inheritance but researches on these resistances are poor. Recessive gene *dn3* confers resistance to the Russian wheat aphid, *D. noxia* in *Triticum tauschii* (Nkongolo et al. 1991). Other example of monogenic recessive resistance are in peanut to the *Aphis craccivora* (Herselman et al. 2004), in maize to *Rhopalosiphum maidis* (Carena and Glogoza, 2004), but also in soybean against the *A. glycines* (Kim et al. 2010).

Table 1.2: Genes, genetics and categories of plant resistance to different aphid species (modified from Smith and Chuang 2014)

Crop plant	Aphid	Gene(s)	Category	References
Alfalfa, *Medicago sativa*	*Acyrthosiphon pisum* Harris	Polygenic	Antibiosis	Julier et al. 2004
Apple, *Malus domestica* and spp.	*Aphis pomi* De Geer	Polygenic, one QTL mapped	Antibiosis	Stoeckli et al. 2008
	Dysaphis devecta Walker	*Sd-1, Sd-2* closely linked or alleles*; Sd-3* from *M. robusta;* QTLs mapped	Antibiosis	Stoeckli et al. 2008; Cevik and King 2002
	Dysaphis plantaginea Passerini	*Sm-h* from *Malus robusta; Dp-fl* from *Malus floribunda* #821	Antibiosis, antixenosis, tolerance	Stoeckli et al. 2008, Pagliarani et al. 2016, Alston et al 1970
	Eriosoma lanigerum Hausmann	*Er1, Er2, Er3, Er4* (*Er1* and *Er2* closely linked or alleles)	Antibiosis, antixenosis	Bus et al. 2008; Bus et al. 2010
Barley, *Hordeum vulgare*	*Diuraphis noxia Mordvilko*	*Rdn1, Rdn2; QTLs mapped*	Antibiosis, tolerance	Mornhinweg et al. 2002; Mittal et al. 2008; Murugan et al 2010; Porter et al. 2007; Carena et al. 2004
	Schizaphis graminum Rondani	*Rsg1, Rsg2*	Antibiosis, tolerance	Porter et al. 2007
Barrel medic, *Medicago truncatula*	*Acryrthosiphon kondoi* Shinji, *Acryrthosiphon pisum* Harris, *Therioaphis trifolii* Monell *f. maculata*	*AKR, AIN, RAP1, TTR*	Antibiosis, antixenosis, tolerance	Klingler et al. 2005; Gao et al. 2008; Klingler et al. 2007; Klingler et al. 2009; Steward et al 2009
Cowpea, *Vigna unguiculata*	*Aphis craccivora* Koch	*Rac1, Rac2*	Antibiosis, antixenosis	Githiri et al. 1996
Lettuce, *Lactuca sativa*	*Nasonovia ribisnigri* Mosely	*Nr* from *L. virosa*, mapped in *L. seriola*	Antibiosis	Lebeda et al. 2014
	Pemphigus bursarius L.	*Ra* or *Lra*	Antibiosis	Wroblewski et al. 2007
Maize, *Zea mays*	*Rhopalosiphum maidis* Fitch	*aph, aph2* two recessive genes (*aph2* mapped)	Antibiosis	Carena et al. 2004
Melon,	*Aphis gossypii* Glover	*Vat* gene cloned, polygenic, four	Antibiosis,	Dogimont et al. 2008;

Cucumis melo		QTLs + two pairs of epistatic QTLs mapped	antixenosis	Boissot et al. 2010
Peach, *Prunus persicae* and spp.	*Myzus persicae* Sulzer	*Rm1, Rm2,* polygenic in *P. davidiana,* 8 QTLs	Antibiosis, antixenosis	Sauge et al. 2002
Peanut, *Arachis hypogea*	*Aphis craccivora Koch*	*Recessive gene (unnamed)*	Antibiosis	Herselman et al. 2004
Pear, *Pyrus* spp.	*Dysaphis pyri* Boyer de Fonscolombe	*Dp-1* from *P. nivalis*	Antibiosis	Evans et al. 2008
Red raspberry, *Rubus ideaus*	*Amphorophora idaei* Börn	13 dominant genes *A1-A10, AK4a, Acor1, Acor2* (only A1 mapped)	Antibiosis, antixenosis	Sargent et al. 2007; McMenemy et al. 2009
Black raspberry, *Rubus ideaus*	*Amphorophora agathonica*	*Ag1, two complementary genes Ag2 and Ag3*	Antibiosis	Dossett et al. 2010
Sorgum, *Sorghum bicolor*	*Schizaphis graminum* Rondani	*Ssa1, Ssg2, Ssg3, Ssg4, Ssg5, Ssg6, Ssg7, Ssg8, Ssg9;* polygenic, QTLs mapped	Antibiosis, tolerance	Agrama et al. 2002; Wu and Huang 2008
Soybean, *Glycine max*	*Aphis glycines* Matsumura	*Rag1, Rag2, Rag3, rag1, rag4;* two QTLs mapped	Antibiosis, antixenosis	Wilkinson et al. 2000; Hill et al. 2012; McCarville et al. 2012; Kim et al. 2010
Tall Fescue, *Festuca arundinacea*	*Rhopalosiphum padi L., Schizaphis graminum* Rondani	*LOL-1, LOL-2* (loline fungal symbiont genes)	Antibiosis, antixenosis	Wilkinson et al. 2000
Tomato, *Solanum lycopersicum*	*Macrosiphum euphorbiae* Thomas	*Mi-1.2 (*or *Meu)* cloned	Antibiosis, antixenosis	Rossi et al. 1998; Goggin et al. 2006
Wheat, *Triticum aestivum* and spp.	*Diuraphis noxia* Mordvilko	*Dn1, Dn2, dn3,* from *T. tauschii Dn4, Dm5, Dn6, Dn7* from *rye Dn8, Dn9, Dnx;* *Dn2414,* all QTLs mappedexcept dn3	Antibiosis, antixenosis, tolerance	Liu et al. 2001 and 2002; Salehi et al. 2012; Castro et al. 2005; Ricciardi et al. 2011; Berzonsky et al. 2003
	Schizaphis graminum Rondani	*gb1* from *T. durum, Gb2* and *Gb6* from *rye, Gb3 and Gb4* from *Aegilops tauschii,* *Gb5* from *A. speltoides, Gby; Gb2, Gb3, Gb5, Gb6* mapped	Antibiosis, antixenosis, tolerance	Castro et al. 2005; Berzonsky et al. 2003

In several cases, aphid resistance is quantitative and polygenic. Genome locations of QTLs involved in aphid resistance have thus far been reported in very few instances. Aphid resistance was measured as intensity of infestation in field conditions or after controlled infestation in most cases. QTL analysis performed in sorghum for resistance to the greenbug *S. graminum*, using different resistance sources and different aphid biotypes revealed three to nine genomic regions involved in the resistance that are likely distinct in the different resistance sources (Agrama et al. 2002; Wu and Huang, 2008).

Aphid resistance genes are often located within clusters of resistance genes in the same chromosomal region, as is the case of many pathogen resistance genes. These 'hot spots' of resistance genes combine genes that confer resistance to aphids and other insects and pathogens. For instance, the barrel medic, *M. truncatula AKR*, *TRR* and *RAP1* genes, each of which confers resistance to a distinct aphid species (*A. kondoi*, *Therioaphis trifolii f maculata* and *A. pisum*) are located within about 40 cM. The *Mi-1* gene, indeed, was located on the chromosome 6 of tomato, which carries an impressive collection of resistance genes effective against fungi, oomycetes and nematodes (Seah et al. 2007).

These cluster of resistance genes targeting taxonomically distinct pests and pathogens suggest that genes with a similar nature confer resistance against different organisms; duplication, recombination and multiple rearrangements events during evolution may have contributed to the development of new resistance specificities (Dogimont et al. 2010).

Aphid resistant cultivars decrease the spread of viruses and two examples elucidate that point. An example is the resistance in raspberry to *Amphorophora idaei* that greatly reduced the severity of virus infections in Europe (McMenemy et al. 2009) and the second example is the *Vat* gene discovered in melon that not only imparts antibiosis and antixenosis against *A. gossypii* but also controls the resistance to *A. gossypii* transmission of cucumber mosaic virus (Table 1.2) (Boissot et al. 2010).

In apple, single dominant genes have been reported to provide resistance to different biotypes of *Dysaphis devecta* (*Sd1*, *Sd2* and *Sd3*), and to the woolly apple aphid *Eriosoma lanigerum* (*Er1*, *Er2* and *Er3, Er4*) (Alston and Briggs 1968, 1977; Cevik and King 2002; Bus et al. 2008, 2010). In detail three genes for the woolly apple aphid resistance were mapped on the genetic map of apple: *Er1* and *Er2* genes are derived from Northern Spy and Robusta 5, respectively. They were used to improve the resistance of apple rootstocks; while the *Er3* donor is Aotea. The genes *Er1* and *Er3* were mapped on the top of the chromosome 8 while the *Er2* gene was located on the chromosome 17. Genes *Er1* and *Er3* map to the same genomic region with a major gene for powdery resistance and with various resistance gene analogues, confirming the clustering of resistance genes in the same genomic regions (Bus et al. 2008). The *Sd1* resistance gene against *D. devecta* was firstly mapped at the top of chromosome 7 of Fiesta. In this genomic region, sequences similar to the NBS-LRR genes, were retrieved during the development of the BAC contig spanning the *Dysaphis devecta* resistance locus. To date the involvement of these genes in *D. devecta* resistance has not been demonstrated yet (Cevik and King 2002). futhermore, *Er4* gene was also mapped in linkage group 7 but in a region of about 21 cM below the two genes *Sd1* and *Sd2* (Bus et al. 2010).

1.4 Apple plants resistant to *Dysaphis plantaginea*

To reduce insecticide applications, the selection of new cultivars resistant to RAA though marker assisted breeding programs is considered a good approach (Arnaoudov and Kurtinova 2006; Miñarro and Dapena 2008). Aphid resistances are mostly carried by wild species or old cultivars which, however, have poor fruit quality traits.

RAA resistance was fisrtly described in *Malus robusta*, where a derivative was shown to carry a single dominant gene for hypersensivity, the gene *Smh* (Alston end Briggs, 1970), but to date this gene has not been mapped yet. Then, a QTL for RAA resistance was reported in LG17 of Fiesta. This study was performed in field conditions in a segregating population of Fiesta. This putative QTL of resistance to RAA co-localize with the QTL of resistance to the green apple aphid, and these QTLs explain from 8 to 20% of the observed variation depending on the sites and years (Stoeckli et al. 2008).

Florina till now is one of the best well studied cultivar for its resistance to RAA. By analyzing the segregation pattern among the numerous progenies derived from Florina crossed different susceptible selections, a 1:1 (resistant : susceptible) segregation was found and this ratio was supporting the hypothesis of the presence of a major resistance gene in a heterozygous (r/s) allelic state in Florina (Rat-Morris 1994; Dapena et al. 2009). Therefore, other studies focused on the Florina cultivar demonstrated the presence of a single resistance gene, named *Dp-fl* (*Dysaphis plantaginea* Florina) located at the bottom of the chromosome 8, precisely at about 6 cM below the SSR maker Ch01h10 (Miñarro and Dapena, 2004; Durel, unpublished data). Recently, the *Dp-fl* locus was mapped within a genomic region of about 330 Kb flanked by two SNP markers identified with the 20K Illumina SNP chip (Bianco et al. 2014; Pagliarani et al. 2016). By performing an *in silico* analysis of the Golden Delicious genome sequence (Velasco et al. 2010), twelve candidate genes putatively involved in RAA resistance were identified in the *Dp-fl* interval (Pagliarani et al. 2016).

Some of RAA resistant cultivar originated from a breeding program that started in the beginning of the XX century at the University of Illinois were the program for the resistance to apple scab (*Venturia inaequalis*) started. The program started when the breeder Crandall crossed a high number of crab apples with commercial cultivars to study fruit size inheritance (Crandall et al. 1926). In detail, in 1914-15 he crossed the apple cultivar Rome Beauty and *Malus floribunda* #821. Twenty years later two siblings ($F_2$26829-2-2 and $F_2$26830-2) were recognized to be scab resistant (Hough, 1944) and were used for further crosses and the inheritance of scab resistance was deeply analyzed (Hough et al. 1953). This resistance was initially named *Vf* from *Venturia* and *floribunda* (Williams et al. 1966) and these two scab-resistant plants were the starting material for the whole PRI Cooperation Program (Bus et al. 2011, Gessler and Pertot 2012, Soriano et al. 2014).

Many years later, different accessions resistant to scab also resulted resistant to the rosy apple aphid. In particular, the scab-resistant cultivar Florina, Galarina, Golden Orange, GoldRush and Liberty are also resistant to the rosy apple aphid (Pagliarani et al. 2016). By the comparison of the pedigrees of these accessions (Figure 1.7) it was clear that the resistance is coming from the wild apple ancestor *Malus floribunda* 821 through its derivative $F_2$26829-2-2.

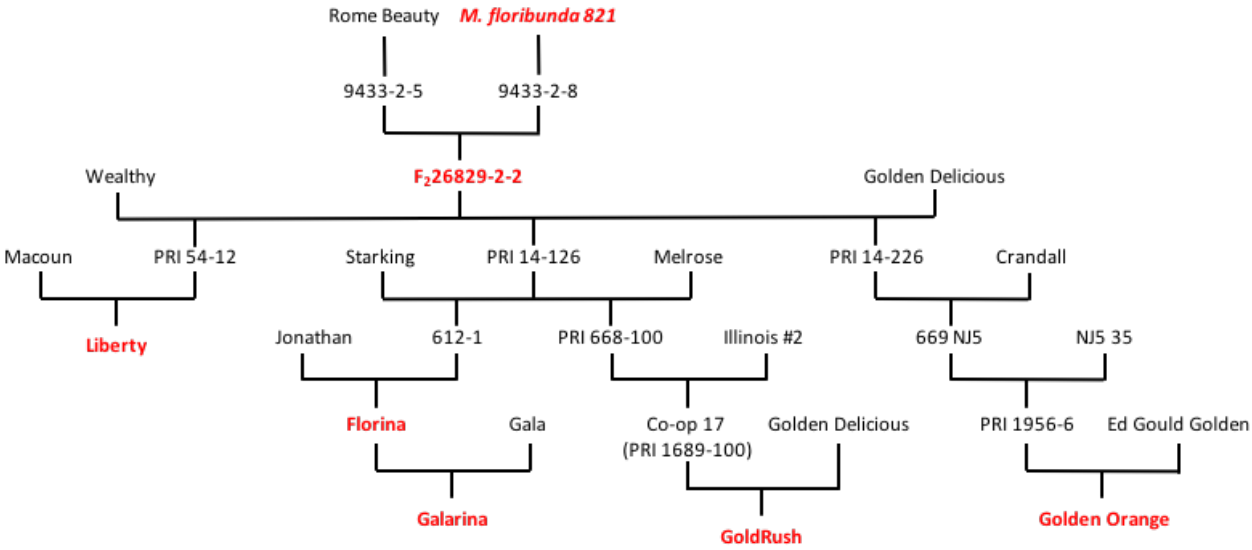

Figure 1.7: Pedigree of Florina starting from *Malus floribunda* clone 821. In red are highlighted the plants that are known to be resistant to *D. plantaginea*.

Among these, Florina, also called Querina, was selected at INRA, Angers, France (Rat-Morris and Lespinasse, 1995) from $F_2$26829-2-2 after subsequent steps of cross and selection with Golden Delicious (PRI 14-126), Starking (PRI 612-1) and Jonathan (Figure 1.7). Florina is not only resistant to apple scab but it is resistant to *D. plantaginea*, fire blight (*Erwinia amylovora*) and red mite (*Panonychus ulmi*) (Lespinasse et al. 1985). Galarina is a direct descendant of the cross Florina x Gala that was also selected at INRA Angers and is resistant to scab, RAA, mildew and fire blight. To date, Florina is the most extensively studied cultivar for its resistance to RAA, characterized by tolerance, antibiosis and antixenosis. When attacked by *D. plantaginea*, this cultivar do not show the typical leaf and shoot deformations and after feeding on Florina leaves RAA has also been shown to be less fecund and with a high mortality (Rat-Morris 1994). On Florina plants, RAA has been observed moving from the leaves to the stems, suggesting repellent compounds released by the leaves or a more difficult stylet penetration (Angeli and Simoni 2006). Consistently, electrical penetration graphs demonstrated that RAA stays on Florina leaves without stylet penetration for a longer period before the first probe and with a reduced duration of sap ingestion, thus indicating mechanical, biochemical or vascular resistance (Marchetti et al. 2009).

GoldRush was also bred from one of the Florina ancestors (PRI 14-126) after three generations of crosses and selection with Melrose (PRI 688-100), Illinois #2 (PRI COOP 17) and Golden Delicious (Figure 1.7). GoldRush is also resistant to scab, RAA, powdery mildew and fire blight.

Golden Orange was selected in 1996 at CRA-FRU in Rome, Italy from the cross between a $F_2$26829-2-2 derivative (PRI 1956-6) and Ed Gould Golden.

Finally, Liberty was selected at the New York State Agricultural Experimental Station of Geneva (New York - U.S.A.) from the parents Macoun and Purdue 54-12, a $F_2$26829-2-2 derivative. Liberty was introduced in 1978 and is also resistant to scab, RAA, fire blight, russet, mildew and cider rust.

However, recent studies conducted on different aphid resistances highlighted the possible low-durability of the single-gene resistances (Bus et al. 2008; Costa et al. 2014; Smith and Chuang 2014). Thus, RAA genetic population variability is an important aspect to take into consideration to avoid possible overcome of the new selected resistant cultivars (Harvey et al. 2003).

1.5 Biotechnology approaches to confer resistance against pathogens and pests

Numerous individual plant resistance (R) genes have already been characterized and are being efficiently used in crop improvement. Benefits of using plant resistance genes in breeding programs include the efficient reduction of pathogens growth, minimal damage to the host plant, reduced input of pesticides and most important the beneficial effects on human health. However, in case of conventional breeding for resistance, the introgression of resistance genes from one species into the gene pool of another by repeated backcrossing is a long-term process which usually takes many generations to obtain the final product. This goal is hindered in highly heterozygous species, like apple where isogenic lines cannot be obtained by conventional breeding techniques. It is assumed that the complete functional studies, cloning characterization and genetic transformation of plant resistance genes could help the researcher to overcome these problems.

During the past 30 years, transgenic arthropod-resistant plants were obtained. In particular, genes encoding toxins from *Bacillus thuringensis* (*Bt*) have been widely adopted to protect cotton, maize and rice from larval herbivory of major coleopteran or lepidopteran pests (Gatehouse 2008). To date, *Bt*-derived toxins have not proved to be effective against aphids (Chougule et al. 2012), but significant progresses to enhance gut binding of the Bt Cy2Aa toxin for toxicity to *Myzus persicae* (Sulzer) and the pea aphid *Aphis gossypii* Gloverhave been described (Sattar and Maiti 2011). These results may suggest the possibility to develop new technologies for using the *Bt* transgenes in developing aphid-resistant crop plants soon (Smith and Chuang, 2014).

Transgenic plants containing genes encoding mannose-specific or N-acetylglucosammine specific lectines are resistant to the green peach aphid *Myzus persicae* (Gatehouse et al. 1996; Birch et al. 1999), the tobacco aphid *Myzus nicotianae* (Blackman) (Wu et al. 2012), the grain aphid *Sitobion avenae* F. (Stoger et al. 1999) and the mustard aphid *Lipaphis erysimi* Kalt. (Kanrar et al. 2002). For both economic and social reason, transgenic aphid resistant crop plants are yet to be developed and deployed.

Till now, two aphid resistance genes have been isolated and their resistance is mediated by a specific recognition of aphid-effector proteins triggering signaling cascades that rapidly activate plant defense against aphids. This scheme was widely described for most plant-pathogen interactions (Dogimont et al. 2010).

The gene *Mi-1.2* is conferring resistance to three species of the root knot nematode *Meloidogyne* (Rossi et al. 1998, Goggin et al 2006). This gene was isolated in wild tomato *Lycopersicum peruvianum* (L.) P. Mill, the same gene was shown to confer resistance to a biotype of the potato *Macrosiphum euphorbiae* (Thomas) as well as to other insects, psyllids and whiteflies (Rossi et al. 1998, Kaloshian et al. 1997; Milligan et al. 1998). Similarly, the melon *Vat* gene confers resistance at the melon-cotton *Aphis gossypii*. This gene was isolated by a map-based cloning strategy and was also demonstrated to possess a unique feature of conferring resistance to non-persistent viruses when vectored by *A. gossypii* (Dogimont et al. 2008). Both isolated aphid resistance genes *Vat* and *Mi-1*, belongs to the same NBS-LRR family of resistance proteins, to which belongs the majority of the genes, isolated to date, conferring resistance to bacteria, viruses, fungi and nematodes (Dangl and Jones 2001). Both genes are constitutively expressed at low levels and encode proteins predicted to be located in the cytoplasm.

To date is still not well known how the aphid resistance genes are involved in the signaling cascades, but the data available indicate that their function partially overlap with those activated by pathogens (Kaloshian and Walling 2005). Both in dicotyledons and monocotyledons, tissues respond to aphid feeding in one hour resulting in the generation of Reactive oxygen species (ROS) elicitors of defense signaling pathways involved in response to aphid attack. ROS elicit the production of defense response signal cascade involving jasmonic acid (JA), salicylic acid (SA) ethylene (ET), abscisic acid (ABA) and gibberellic acid (GA), which are commonly involved in aphid-resistant plants. JA and SA signals modulate plant response to aphid herbivory and function in aphid resistance. Although comparatively less is known about ethylene defense response signaling, aphid feeding also causes a significant upregulation of ET production or ET-signaling genes in aphid-resistant barley and wheat when compared with susceptible plants (Smith and Chuang 2014).

References

Agrama H, Widle G, Reese J, Campbell L & Tuinstra M (2002). Genetic mapping of QTLs associated with greenbug resistance and tolerance in *Sorghum* bicolor. Theor Appl Genet, 104:1373-1378.

Alston FH & Briggs JB (1968). Resistance to *Sappaphis devecta* (Wlk.) in apple. Euphytica, 17:468-472.

Alston FH & Briggs JB (1970). Inheritance of hypersensitivity to rosy apple aphid *Dysaphis plantaginea* in apple. Can J Genet Cytol 12:257-258.

Alston FH & Briggs JB (1977). Resistance genes in apple and biotypes of *Dysaphis devecta*. Ann Appl Biol 87:75-81.

Angeli G & Simoni S (2006). Apple cultivars acceptance by *Dysaphis plantaginea* Passerini (Homoptera: Aphididae). J Pest Sci 79:175-179.

Arnaoudov V, Kutinkova H (2006). Susceptibility of some apple cultivars to infestation by the rosy apple aphid (*Dysaphis plantaginea* Pass., Homoptera: Aphididae). J Fruit Ornam Plant Res 14:137-142.

Asante SK (1997). Natural enemies of the woolly apple aphid, *Eriosoma lanigerum* (Hausmann) (Hemiptera: Aphididae): a review of the world literature. Plant Prot Quarterly, 12:166-172.

Barbagallo S (1996). Afidi delle principali colture fruttifere. Bayer.

Baumgartner IO, Kellerhals M, Costa F, Dondini L, Pagliarani G, Gregori R, Tartarini S, Leumann L, Laurens F, Patocchi A. (2016). Development of SNP-based assays for disease resistance and fruit quality traits in apple (*Malus* × *domestica* Borkh.) and validation in breeding pilot studies. Tree Genet Genom 12:1-21.

Berzonsky WA, Ding H, Haley SD, Harris MO, Lamb RJ, McKenzie RIH, … & Ratcliffe RH (2003). Breeding wheat for resistance to insects. Plant Breed Rev 22:221-296.

Bianco L, Cestaro A, Linsmith G, Muranty H, Denancé C, Théron A, ... & Larger S (2016). Development and validation of the Axiom® Apple480K SNP genotyping array. Plant J 86:62-74.

Bianco L, Cestaro A, Sargent DJ, Banchi E, Derdak S, Di Guardo M, Salvi S, Jansen J, Viola R, Gut I, Laurens F, Chagné D, Velasco R, van de Weg E, Troggio M (2014) Development and validation of a 20K single nucleotide polymorphism (SNP) whole genome genotyping array for apple (*Malus* x *domestica* Borkh). PLoS One 9(10) e110377. doi:10.1371/journal.pone.0110377

Birch ANE, Geoghegan IE, Majerus ME, McNicol JW, Hackett CA, Gatehouse AM & Gatehouse JA (1999). Tri-trophic interactions involving pest aphids, predatory 2-spot ladybirds and transgenic potatoes expressing snowdrop lectin for aphid resistance. Mol Breed, 5:75-83.

Blackman RL, & Eastop, V. F. (1994). Aphids on the world's trees: an identification and information guide. Cab International in Association with the Natural History Museum, Wallingford, 987 pp.

Blommers LHM, Helsen HHM & Vaal FWNM (2004). Life history data of the rosy apple aphid *Dysaphis plantaginea* (Pass.) (Homopt., Aphididae) on plantain and as migrant to apple. J Pest Sci, 77:155-163.

Boissot N, Thomas S, Sauvion N, Marchal C, Pavis C & Dogimont C (2010). Mapping and validation of QTLs for resistance to aphids and whiteflies in melon. Theor Appl Genet 121:9-20.

Boyko, E., Starkey, S., & Smith, M. (2004). Molecular genetic mapping of *Gby*, a new greenbug resistance gene in bread wheat. Theor Appl Genet 109:1230-1236.

Briggs JB & Alston FH (1967). Pest avoidance by late-flowering apple varieties. East Mailing Res Sta Annu Rep 1966:170–171.

Bus VGM, Chagné D, Bassett HCM, Bowatte D, Calenge F, Celton JM ... & Rikkerink EHA (2008). Genome mapping of three major resistance genes to woolly apple aphid (*Eriosoma lanigerum* Hausm.). Tree Genet Genom 4:223-236.

Bus V, Bassett H, Bowatte D, Chagné D, Ranatunga C, Ulluwishewa D, Wiedow C, Gardiner S (2010) Genome mapping of an apple scab, a powdery mildew and a woolly apple aphid resistance gene from open-pollinated Mildew Immune Selection. Tree Genet Genom 6:477–487

Bus VG, Rikkerink EH, Caffier V, Durel CE & Plummer KM (2011). Revision of the nomenclature of the differential host-pathogen interactions of *Venturia inaequalis* and *Malus*. Annual Rev Phytopathol 49:391-413.

Carena MJ & Glogoza P (2004). Resistance of maize to the corn leaf aphid: a review. Maydica, 49:241-254.

Castro AM, Vasicek A, Manifiesto M, Gimenez DO, Tacaliti MS, Dobrovolskaya O, ... & Börner A (2005). Mapping antixenosis genes on chromosome 6A of wheat to greenbug and to a new biotype of Russian wheat aphid. Plant Breed, 124:229-233.

Cevik V & King GJ (2002). Resolving the aphid resistance locus *Sd-1* on a BAC contig within a sub-telomeric region of *Malus* linkage group 7. Genome 45:939-945.

Challice JS (1974). Rosaceae chemotaxonomy and the origins of the Pomoideae. Botanic J Linn Soc 69:239-259.

Chougule NP & Bonning BC (2012). Toxins for transgenic resistance to hemipteran pests. Toxins, 4:405-429.

CIE (1975). Commonwealth Institute of Entomology. Distribution Maps of Pests. Series A (Agriculture). Map No. 17 Wallingford, UK: CAB International.

Conceiçao MB, Catherine MCB, Pervin ME, Joanne MF, Jean-François MG, Neil MG, ... & Secretariat EPPO Organisation Europeenne et Mediterraneenne pour la Protection des Plantes European and Mediterranean Plant Protection Organization (1999)

Costa A, Williams DG & Powell KS (2014). Discovery of three woolly apple aphid *Eriosoma lanigerum* (Hemiptera: Aphididae) biotypes in Australia: the role of antixenosis and antibiosis in apple tree resistance. Aust Entomol 53:280-287.

Crandall CS (1926). Apple breeding at the University of Illinois. Bulletin (University of Illinois (Urbana-Champaign campus). Agric Exp Stn no. 275.

Cross JV, Cubison S, Harris A & Harrington R (2007). Autumn control of rosy apple aphid, *Dysaphis plantaginea* (Passerini), with aphicides. Crop Prot 26:1140-1149.

Daccord N, Celton JM, Linsmith G, Becker C, Choisne N, Schijlen E, van de Geest H, Bianco L, Micheletti D, Velasco R, Di Pierro EA, Gouzy J, Rees DJG, Guérif P, Muranty H, Durel CE, Laurens F, Lespinasse Y, Gaillard S, Aubourg S, Quesneville H, Weigel D, van de Weg E, Troggio M, Bucher E (2017) High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. Nat Genet 49:1099-1106.

Dangl JL, Jones JD (2001). Plant pathogens and integrated defence responses to infection. Nat 411:826-833.

Dapena E, Miñarro M, Blasquez MD (2009) Evaluation of the resistance to the rosy apple aphid using a genetic marker. ISHS Acta Hortic 814:787-790.

Delorme R, Auge D, Touton P, Vergnet C & Villatte F. (1998). La résistance des pucerons aux insecticides. Enquêtes 1997. In Proceedings of the 1st Colloque Transnational sur les Luttes biologiques, integrées et raisonnées.

Di Pierro EA, Gianfranceschi L, Di Guardo M, Koehorst-van Putten HJ, Kruisselbrink JW, Longhi S, ... & Tartarini S (2016). A high-density, multi-parental SNP genetic map on apple validates a new mapping approach for outcrossing species. Hortic Res 3,16057.

Dib H, Issa RB, Sauphanor B & Capowiez Y (2017). Feasibility and efficacy of a new approach for controlling populations of the rosy apple aphid, *Dysaphis plantaginea* Passerini (Hemiptera: Aphididae) in south-eastern France. Intern J Pest Manag 63:128-137.

Dib H, Simon S, Sauphanor B, Capowiez Y (2010). The role of natural enemies on the population dynamics of the rosy apple aphid, *Dysaphis plantaginea* Passerini (Hemiptera: Aphididae) in organic apple orchards in south-eastern France. Biol Control 55:97-109.

Dogimont C, Bendahmane A, Chovelon V & Boissot N (2010). Host plant resistance to aphids in cultivated crops: genetic and molecular bases, and interactions with aphid populations. Comptes rendus biol, 333:566-573.

Dogimont C, Chovelon V, Tual S, Boissot N, Rittener-Rüff V, Giovinazzo N, Bendahmane A, Pitrat M (2008). Molecular diversity at the *Vat*/*Pm*-*W* resistance locus in melon. Cucurbitaceae 2008 219-228.

Dossett M & Finn CE (2010). Identification of resistance to the large raspberry aphid in black raspberry. J Am Soc Hortic Sci 135:438-444.

Evans RC & Campbell CS (2002). The origin of the apple subfamily (Maloideae; Rosaceae) is clarified by DNA sequence data from duplicated GBSSI genes. Am J Bot, 89:1478-1484.

Evans KM, Govan CL & Fernandez-Fernandez F (2008). A new gene for resistance to *Dysaphis pyri* in pear and identification of flanking microsatellite markers. Genom 51:1026-1031.

Faccioli G, Baronio P, Pasqualini E & Malavolta C (1985). *Dysaphis plantaginea*: ethology and injuriousness. Informatore Agrario, 41:61-63.

Falginella L, Cipriani G, Monte C, Gregori R, Testolin R, Velasco R, ... & Tartarini S (2015). A major QTL controlling apple skin russeting maps on the linkage group 12 of 'Renetta Grigia di Torriana'. BMC Plant Biol, 15:150.

FAOSTAT website http://www.fao.org/faostat/en/#home

Fidantsef AL, Stout MJ, Thaler JS, Duffey SS, Bostock RM (1999). Signal interactions in pathogen and insect attack: expression of lipoxygenase, proteinase inhibitor II, and pathogenesis-related protein P4 in the tomato, Lycopersicon esculentum. Physiol Mol Plant Pathol, 54:97-114.

Gao LL, Klingler JP, Anderson JP, Edwards OR & Singh KB (2008). Characterization of pea aphid resistance in *Medicago truncatula*. Plant Physiol, 146:996-1009.

Gatehouse, A. M., Down, R. E., Powell, K. S., Sauvion, N., Rahbé, Y., Newell, C. A., ... & Gatehouse, J. A. (1996). Transgenic potato plants with enhanced resistance to the peach-potato aphid *Myzus persicae*. Entomol Exp Appl 79:295-307.

Gatehouse JA (2008). Biotechnological prospects for engineering insect-resistant plants. Plant Physiology, 146:881-887.

Gessler C & Pertot I (2012). *Vf* scab resistance of *Malus*. Trees 26:95-108.

Githiri SM, Kimani PM & Pathak RS (1996). Genetic linkage of aphid resistance gene, *rac* in cowpea. Afr Crop Sci Soc 4:145-150.

Goggin FL, Jia L, Shah G, Hebert S, Williamson VM & Ullman DE (2006). Heterologous expression of the *Mi-1.2* gene from tomato confers resistance against nematodes but not aphids in eggplant. Molec Plant-Microbe Interact 19:383-388.

Gururani MA, Venkatesh J, Upadhyaya CP, Nookaraju A, Pandey SK & Park SW (2012). Plant disease resistance genes: current status and future directions. Physiol Mol Plant Pathol 78:51-65.

Han Y, Gasic K, Marron B, Beever JE & Korban SS (2007). A BAC-based physical map of the apple genome. Genom 89:630-637.

Harvey NG, Fitzgerald JD, James CM & Solomon MG (2003). Isolation of microsatellite markers from the rosy apple aphid *Dysaphis plantaginea*. Molec Ecol Resour 3:111-112.

Herselman L, Thwaites R, Kimmins FM, Courtois B, Van der Merwe PJA & Seal SE (2004). Identification and mapping of AFLP markers linked to peanut (*Arachis hypogaea* L.) resistance to the aphid vector of groundnut rosette disease. Theor Appl Genet 109:1426-1433.

Hill CB, Chirumamilla A & Hartman GL (2012). Resistance and virulence in the soybean-*Aphis glycines* interaction. Euphytica 186:635-646.

Hough LF (1944). A survey of the scab resistance of the foliage on seedlings in selected apple progenies. Proc Am Soc Hortic Sci 44:260-272.

Hough LF, Shay JR & Dayton DF (1953). Apple scab resistance from *Malus floribunda* Sieb. Proc Am Soc Hortic Sc. 62:805-820.

Hull LA & Grimm JW (1983). Sampling schemes for estimating populations of the apple aphid, *Aphis pomi* (Homoptera: Aphididae), on apple. Environ Entomol 12:1581-1586.

Hummer KE & Janick J (2009). Rosaceae: taxonomy, economic importance, genomics. In Genetics and genomics of Rosaceae (pp. 1-17). Springer New York.

Janick J & Moore JN (1996). Fruit breeding, tree and tropical fruits (Vol. 1). John Wiley & Sons.

Julier B, Bournoville R, Landré B, Ecalle C & Carré S (2004). Genetic analysis of lucerne (*Medicago sativa* L.) seedling resistance to pea aphid (*Acyrthosiphon pisum* Harris). Euphytica, 138:133-139.

Kaloshian I, Kinsey MG, Ullman DE & Williamson VM (1997). The impact of *Meu1* ‐mediated resistance in tomato on longevity, fecundity and behavior of the potato aphid, *Macrosiphum euphorbiae*. Entomol Exp Appl 83:181-187.

Kaloshian I & Walling LL (2005). Hemipterans as plant pathogens. Annu Rev Phytopathol 43:491-521.

Kanrar S, Venkateswari J, Kirti P & Chopra V (2002). Transgenic Indian mustard (*Brassica juncea*) with resistance to the mustard aphid (*Lipaphis erysimi* Kalt.). Plant Cell Rep 20:976-981.

Kim KS, Bellendir S, Hudson KA, Hill CB, Hartman GL, Hyten DL, ... & Diers BW (2010). Fine mapping the soybean aphid resistance gene *Rag1* in soybean. Theor Appl Genet 120:1063-1071.

Klingler J, Creasy R, Gao L, Nair RM, Calix AS, Jacob HS, ... & Singh KB (2005). Aphid resistance in *Medicago truncatula* involves antixenosis and phloem-specific, inducible antibiosis, and maps to a single locus flanked by NBS-LRR resistance gene analogs. Plant Physiol 137:1445-1455.

Klingler JP, Edwards OR & Singh KB (2007). Independent action and contrasting phenotypes of resistance genes against spotted alfalfa aphid and bluegreen aphid in *Medicago truncatula*. New Phytol 173:630-640.

Klingler JP, Nair RM, Edwards OR & Singh KB (2009). A single gene, AIN, in *Medicago truncatula* mediates a hypersensitive response to both bluegreen aphid and pea aphid, but confers resistance only to bluegreen aphid. J Exp Bot 60:4115-4127.

Knight RL & Alston FH (1974). Pest resistance in fruit breeding. In The 13th Symposium of The British Ecological Society Oxford, 4-7 January 1972. (pp. 73-86). Blackwell Sci Publ.

Lebeda A, Křístková E, Kitner M, Mieslerová B, Jemelková M & Pink DA (2014). Wild Lactuca species, their genetic diversity, resistance to diseases and pests, and exploitation in lettuce breeding. European J Plant Pathol, 138:597-640.

Lespinasse Y, Olivier JM, Lespinasse JM, Le Lezec M (1985) Florina Querina®, la résistance du pommier à la tavelure.- Arboriculture Fruitière, 378, 43-46.

Liu XM, Smith CM, Gill BS & Tolmay V (2001). Microsatellite markers linked to six Russian wheat aphid resistance genes in wheat. Theor Appl Genet, 102:504-510.

Liu X, Smith C & Gill B. (2002). Identification of microsatellite markers linked to Russian wheat aphid resistance genes *Dn4* and *Dn6.* Theor Appl Genet, 104:1042-1048.

Marchetti E, Civolani S, Leis M, Chicca M, Tjallingii WF, Pasqualini E, Baronio P (2009). Tissue location of resistance in apple to the rosy apple aphid established by electrical penetration graphs. Bull Insectology, 62:203-208.

McCarville MT & O'Neal ME (2012). Measuring the benefit of biological control for single gene and pyramided host plant resistance for *Aphis glycines* (Hemiptera: Aphididae) management. J Econom Entomol 105:1835-1843.

McGavin GC (1993). Bugs of the world. Facts on File. Inc., Hong Kong.

McMenemy LS, Mitchell C & Johnson SN (2009). Biology of the European large raspberry aphid (*Amphorophora idaei*): its role in virus transmission and resistance breakdown in red raspberry. Agric For Entomol 11:61-71.

McVay JR, Walgenbach JF, Sikora EJ & Sutton TB. (1993). A grower's guide to apple insects and diseases in the southeast. Alabama Cooperative Extension Service, Auburn University, Circular ANR-838.

Miles PW (1999) Aphid saliva. Biol Rev Camb Philos Soc 74:41-85.

Milligan SB, Bodeau J, Yaghoobi J, Kaloshian I, Zabel P & Williamson VM (1998). The root knot nematode resistance gene *Mi* from tomato is a member of the leucine zipper, nucleotide binding, leucine-rich repeat family of plant genes. Plant Cell, 10:1307-1319.

Miñarro M, Dapena E (2004) Inheritance of the tolerance to the rosy apple aphid of the cv. 'Florina'. Acta Hortic 663:261–264. doi:10. 17660/ActaHortic.2004.663.42

Miñarro M, Dapena E. (2008) Tolerance of some scab-resistant apple cultivars to the rosy apple aphid, *Dysaphis plantaginea*. Crop Prot 27:391-395.

Miñarro M & Dapena E (2014). Resistance of apple cultivars to *Dysaphis plantaginea* (Hemiptera: Aphididae): role of tree phenology in infestation avoidance. Environ Entomol 36:1206-1211.

Miñarro M, Fernadez-Mata Gabriela & Medina P (2010). Role of ants in structuring the aphid community on apple. Ecol Entomol 35:206-215.

Miñarro M, Hemptinne JL & Dapena E (2005). Colonization of apple orchards by predators of *Dysaphis plantaginea*: sequential arrival, response to prey abundance and consequences for biological control. BioControl 50:403-414.

Mittal S, Dahleen LS & Mornhinweg D. (2008). Locations of quantitative trait loci conferring Russian wheat aphid resistance in barley germplasm STARS-9301B. Crop Sci 48:1452-1458.

Mornhinweg DW, Porter DR & Webster JA (2002). Inheritance of Russian wheat aphid resistance in spring barley germplasm line STARS-9577B. Crop Sci 42:1891-1893.

Murugan M, Khan SA, Cardona PS, Orozco GV, Viswanathan P, Reese J, ... & Smith CM (2010). Variation of resistance in barley against biotypes 1 and 2 of the Russian wheat aphid (Hemiptera: Aphididae). J Econom Entomol 103:938-948.

Noiton DA & Alspach PA (1996). Founding clones, inbreeding, coancestry, and status number of modern apple cultivars. J Americ Soc Hortic Sci 121:773-782.

Nkongolo KK, Quick JS, Limin AE & Fowler DB (1991). Sources and inheritance of resistance to Russian wheat aphid in Triticum species amphiploids and *Triticum tauschii*. Canadian J Plant Sci 71:703-708.

Padmarasu S, Sargent DJ, Jaensch M, Kellerhals M, Tartarini S, Velasco R, ... & Patocchi A. (2014). Fine-mapping of the apple scab resistance locus *Rvi12* (*Vb*) derived from 'Hansen's baccata# 2'. Molec breed 34:2119-2129.

Pagliarani G, Dapena E, Miñarro M, Denancé C, Lespinasse Y, Rat-Morris E et al. (2016). Fine mapping of the rosy apple aphid resistance locus *Dp-fl* on linkage group 8 of the apple cultivar 'Florina'. Tree Genet Genom, 12:1-12.

Painter RH (1951). Insect resistance in crop plants. The Macmillan Company New York.

Palmer MA. (1952). Aphids of the Rocky Mountain Region. Aphids of the Rocky Mountain Region.

Pasqualini E, Antropoli A & Zecchini G. (1996). Control perspectives against grey aphid (*Dysaphis plantaginea* Pass., Rhynchota Aphididae) on apple: fourth contribution. Informatore Fitopatologico, 46:50-55.

Peck G & Merwin IA (2010). A grower's guide to organic apples. NYS IPM Publication No. 223. Geneva, NY.

Porter DR, Burd JD & Mornhinweg DW (2007). Differentiating greenbug resistance genes in barley. Euphytica, 153:11-14.

Rat-Morris E. (1994) Analyse des relations entre *Dysaphis plantaginea* Pass. (Insecte Homoptère) et sa plante hôte *Malus* x *domestica* Borkh.: étude de la résistance du cultivar Florina. PhD thesis, Faculté des Sciences, Université de Tours, France (114 p).

Rat-Morris E, Lespinasse Y (1995). Pommier: la résistance au puceron cendré associée à la résistance à la tavelure. Une nouvelle source de résistance mise en évidence chez le cultivar Florina. Phytoma-La Défense des Végétaux, 471:15-17.

Ricciardi M, Tocho E, Tacaliti MS, Vasicek A, Giménez DO, Paglione A, ... & Castro AM. (2011). Mapping quantitative trait loci for resistance against Russian wheat aphid (*Diuraphis noxia*) in wheat (*Triticum aestivum* L.). Crop Pasture Sci 61 :970-977.

Roche P, Alston FH, Maliepaard C, Evans KM, Vrielink R, Dunemann F, ... & King, GJ (1997). RFLP and RAPD markers linked to the rosy leaf curling aphid resistance gene (*Sd1*) in apple. Theor Appl Genet 94 :528-533.

Rossi M, Goggin FL, Milligan SB, Kaloshian I, Ullman DE, Williamson VM (1998). The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. Proc Natl Acad Sci 95 :9750-9754.

Seah S, Telleen AC & Williamson VM (2007). Introgressed and endogenous *Mi-1* gene clusters in tomato differ by complex rearrangements in flanking sequences and show sequence exchange and diversifying selection among homologues. Theor Appl Genet 114:1289-1302.

Salehi T, Mirak TN, Bihamta MR & Heravan EM. (2012). Inheritance of Russian wheat aphid resistance in Iranian bread wheat cultivar 'Azadi'. Adv Environ Biol 2582-2586.

Sargent DJ, Fernández-Fernández F, Rys A, Knight VH, Simpson DW & Tobutt KR (2007). Mapping of *A1* conferring resistance to the aphid *Amphorophora idaei* and *dw* (dwarfing habit) in red raspberry (*Rubus idaeus* L.) using AFLP and microsatellite markers. BMC Plant Biol, 7:15.

Sattar S & Maiti MK (2011). Molecular characterization of a novel vegetative insecticidal protein from *Bacillus thuringiensis* effective against sap-sucking insect pest. J Microbiol Biotechnol 21 :937-946.

Sauge MH, Lacroze JP, Poëssel JL, Pascal T & Kervella J (2002). Induced resistance by *Myzus persicae* in the peach cultivar 'Rubira'. Entomol Exp Appl 102:29-37.

Smith CM & Chuang WP (2014). Plant resistance to aphid feeding: behavioral, physiological, genetic and molecular cues regulate aphid host selection and feeding. Pest Manag Sci 70:528-540.

Solomon MG, Cross JV, Fitzgerald JD, Campbell CAM, Jolly RL, Olszak RW, ... & Vogt H (2000). Biocontrol of pests of apples and pears in northern and central Europe-3. Predators. Biocontrol Sci Techn, 10:91-128.

Soriano JM, Madduri M, Schaart JG, van der Burgh A, van Kaauwen MP, Tomic L, ... & Schouten HJ (2014). Fine mapping of the gene *Rvi18* (V25) for broad-spectrum resistance to apple scab, and development of a linked SSR marker suitable for marker-assisted breeding. Molec Breed, 34:2021-2032.

Stewart SA, Hodge S, Ismail N, Mansfield JW, Feys BJ, Prospéri JM, ... & Powell G (2009). The *RAP1* gene confers effective, race-specific resistance to the pea aphid in *Medicago truncatula* independent of the hypersensitive reaction. Molec Plant-Microbe Interact 22:1645-1655.

Stewart-Jones A, Pope TW, Fitzgerald JD & Poppy GM (2008). The effect of ant attendance on the success of rosy apple aphid populations, natural enemy abundance and apple damage in orchards. Agric For Entomo, 10:37-43.

Stoeckli S, Mody K, Gessler C, Patocchi A, Jermini M & Dorn S (2008). QTL analysis for aphid resistance and growth traits in apple. Tree Genet Genom 4:833-847.

Stoger E, Williams S, Christou P, Down RE & Gatehouse JA (1999). Expression of the insecticidal lectin from snowdrop (*Galanthus nivalis* agglutinin; GNA) in transgenic wheat plants: effects on predation by the grain aphid *Sitobion avenae*. Molec Breed 5:65-73.

Urrestarazu J, Muranty H, Denancé C, Leforestier D, Ravon E, Guyader A, ... & Daccord N (2017). Genome-wide association mapping of flowering and ripening periods in apple. Front Plant Sci, 8, 1923.

Van der Westhuizen AJ, Qian XM, Botha AM (1998) Differential induction of apoplastic peroxidase and chitinase activities in susceptible and resistant wheat cultivars by Russian wheat aphid infestation. Plant Cell Rep 18:132-137.

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A et al. (2010). The genome of the domesticated apple (*Malus* [times] *domestica* Borkh.). Nat Genet 42:833-839.

Vinatzer BA, Patocchi A, Gianfranceschi L, Tartarini S, Zhang HB, Gessler C & Sansavini S. (2001). Apple contains receptor-like genes homologous to the *Cladosporium fulvum* resistance gene family of tomato with a cluster of genes cosegregating with *Vf* apple scab resistance. Molec Plant-Microbe Interact 14:508-515.

Wilkinson HH, Siegel MR, Blankenship JD, Mallory AC, Bush LP & Schardl CL (2000). Contribution of fungal loline alkaloids to protection from aphids in a grass-endophyte mutualism. Molec Plant-Microbe Interact, 13:1027-1033.

Williams, E. B., Dayton, D. F., & Shay, J. R. (1966). Allelic genes in *Malus* for resistance to *Venturia inaequalis*. Proc Amer Soc Hort Sci 88:52-56.

Wroblewski T, Piskurewicz U, Tomczak A, Ochoa O & Michelmore RW. (2007). Silencing of the major family of NBS–LRR-encoding genes in lettuce results in the loss of multiple resistance specificities. Plant J, 51:803-818.

Wu Y & Huang Y (2008). Molecular mapping of QTLs for resistance to the greenbug *Schizaphis graminum* (Rondani) in *Sorghum bicolor* (Moench). Theor Appl Genet 117:117.

Wu ZM, Yan HB, Pan WL, Jiang B, Liu JG, Geng BJ, ... & Dong WQ. (2012). Transform of an ectopically expressed bulb lectin gene from *'Pinellia pedatisecta*' into tobacco plants conferring resistance to aphids (*'Myzus nicotianae'*). Australian J Crop Sci 6:904.

Wyss E, Villiger M & Müller-Schärer H. (1999). The potential of three native insect predators to control the rosy apple aphid, *Dysaphis plantaginea*. BioControl, 44:171-182

Wyss E & Daniel C. (2004). Effects of autumn kaolin and pyrethrin treatments on the spring population of *Dysaphis plantaginea* in apple orchards. J Appl Entomol 128:147-149.

Züst T & Agrawal AA (2016). Mechanisms and evolution of plant resistance to aphids. Nature plants, 2, 15206.

Xu M & Korban SS (2002). A cluster of four receptor-like genes resides in the *Vf* locus that confers resistance to apple scab disease. Genet 162:1995-2006.

Chapter 2 – Aim of the thesis

Taking into consideration the consistent damages that the rosy apple aphid *D. plantaginea* (RAA, Passerini) infers every year to apple cultivations in temperate regions and considering that breeding new apple cultivars resistant to the aphid is considered a good approach to reduce the insecticide applications, the principal aim of this thesis is to identify the gene that confers resistance to *D. plantaginea* in apple. Detailed information about plant-aphid interaction will be searched to better understand the mechanisms that are at the base of the resistance mechanism. This thesis start from the knowledge of the mapping of the resistance locus *Dp-fl* obtained in Florina by Pagliarani et al. 2016.

The fist aim of this thesis was to identify a minimum tiling path of Bacterial Artificial Chromosome (BAC) (Vinatzer et al. 1998) clones covering the entire 330 Kb *Dp-fl* region of Florina. In order to obtain this objective, the development of new polymorphic markers was necessary to perform a chromosome walking within the *Dp-fl* region. The development of new markers tightly linked to the *Dp-fl* resistance locus was also an important step for the early selection of RAA-resistant plants in Marker Assisted Breeding programs.

Second aim of this project was to fine-map the resistance gene, objective that was carried on through the development of new markers and by the identification of further plants recombining within the *Dp-fl* locus. This goal was accomplished by the analysis of new progenies derived from Florina segregating for the resistance locus. The genotypic and phenotypic evaluation of the plants recombining at the *Dp-fl* locus made it possible to better define and reduce the *Dp-fl* region. To further confirm the location of the resistance gene, also large progenies directly derived from the wild species *Malus floribunda* #821, the donor of RAA resistance locus, were also analyzed.

The third aim was the identification of candidate genes within the *Dp-fl* region of Florina and verify their constitutive or induced expression after RAA infestation. This to better understand the role of the gene in the plant-aphid interaction. The final goal of this thesis was to start the validation of the function of a candidate gene by genetic transformation of Gala and Florina plantlets in order to prove the function of a candidate gene.

Chapter 3 - Identification of candidate genes at the *Dp-fl* locus conferring resistance against the rosy apple aphid *Dysaphis plantaginea*

3.1 ABSTRACT

The cultivated apple is susceptible to several pests including the rosy apple aphid (RAA, *Dysaphis plantaginea* Passerini), control of which is mainly based on chemical treatments. A few cases of resistance to aphids have been described in apple germplasm resources, laying the basis for the development of new resistant cultivars by breeding. The cultivar Florina is resistant to RAA and recently, the *Dp-fl* locus responsible for its resistance was mapped on linkage group 8 of the apple genome. In this paper, a chromosome walking approach was performed by using a Florina Bacterial Artificial Chromosome (BAC) library. The walking started from the available tightly-linked molecular markers flanking the resistance region. Various walking steps were performed in order to identify the minimum tiling path of BAC clones covering the *Dp-fl* region from both the 'resistant' and 'susceptible' chromosomes of Florina. A genomic region of about 279 Kb encompassing the *Dp-fl* resistance locus was fully sequenced by the PacBio technology. Through the development of new polymorphic markers, the mapping interval around the resistance locus was narrowed down to a physical region of 95 Kb. The annotation of this sequence resulted in the identification of four candidate genes putatively involved in the RAA resistance response.

3.2 INTRODUCTION

Rosy apple aphid (RAA, *Dysaphis plantaginea* Passerini) is one of the most damaging insects affecting cultivated apple *Malus* × *domestica* (Brown and Mathews 2007; Parisi et al. 2013). RAA is present in Europe, North America, North Africa and Asia (Aslan and Karaca 2005; Brown and Mathews 2007; Miñarro and Dapena 2007) and causes severe damage on shoots, leaves and fruits that remain small and deformed, leading to significant economic losses (De Berardinis et al. 1994). Aphids inject saliva into the sieve elements before sap ingestion. The saliva contains non-enzymatic/reducing compounds that, in the presence of oxidase, can combine and inactivate defensive phytochemicals of the plant, including those released in response to damage. The saliva containing these signals is transported in the phloem flow, but the mechanism of action is still unclear (Miles 1999). Moreover, these signals alter the expression of inducible plant physiological factors involved in defense against pathogens (Van Der Westhuizen et al. 1998; Fidantsef et al. 1999). Aphids may also transmit various plant viruses through the saliva causing irrevocable damage to the plant (Martín et al. 1997; Ng and Perry 2004).

Interest in biological control is increasing, but the use of natural enemies of RAA is not sufficiently effective in reducing the aphid population, so pesticide sprayings are still the main control strategy utilized (Miñarro et al. 2005; Brown and Mathews 2007; Dib et al. 2010; Andreev et al. 2012). To reduce insecticide applications, marker assisted breeding programs for the selection of new resistant cultivars is considered a good approach (Angeli and Simoni 2006; Arnaoudov and Kutinkova 2006; Miñarro and Dapena 2008). Determination of the inheritance, the genetic location and knowledge of the resistance gene products are essential for breeding aphid resistant apple trees. Two aphid resistance genes have been isolated and cloned in two different species: the tomato *Mi-1.2* gene and the melon *Vat* gene (Rossi et al. 1998; Dogimont et al. 2008). Both genes are constitutively expressed at low levels in the plant and encode for a protein located in the cytoplasm that belong to the NBS-LRR family, to which belongs the majority of plant resistance genes isolated so far (Dangl and Jones 2001).

The apple cultivar Florina was selected at INRA, Angers, France, from a derivative of the wild apple *Malus floribunda* clone #821 (Rat-Morris and Lespinasse 1995). This accession was initially selected for its resistance to apple scab (caused by the fungus *Venturia inaequalis*), but later it was also shown to be tolerant to the aphid *D. plantaginea*, fire blight *Erwinia amylovora* and red mite *Panonychus ulmi* (Lespinasse et al. 1985).

Florina RAA resistance is characterized by both tolerance and antibiosis (Rat-Morris 1993, 1994). Indeed, when infested by RAA, Florina does not show the typical leaf and shoot deformations caused by this aphid on susceptible cultivars. RAA appeared to be less fecund and showing higher mortality after feeding on Florina leaves. RAA has been observed moving from the leaves to the stems of Florina plants, which suggests repellent compounds released by the leaves or a more difficult stylet penetration (Angeli and Simoni 2006). Consistently, electrical penetration graphs demonstrated that RAA stays on Florina leaves without stylet penetration for a longer period before the first probe and with a reduced duration of sap ingestion (Marchetti et al. 2009), thus indicating mechanical, biochemical or vascular resistance. For the sake of simplicity, the general term 'resistance' has been used here for both the tolerance and antibiosis mechanisms of Florina.

The inheritance of Florina RAA resistance was explored in segregating progenies and a di-genic model was initially proposed (Rat-Morris 1994). A single dominant resistance (R) gene, named *Dp-fl*, was identified and located at the distal end of linkage group (LG) 8 of the apple genome (Durel, unpublished data) and confirmed by Dapena et al. (2009). Recently, the *Dp-fl* locus was mapped within a genomic region of about 330 Kb flanked by two SNP markers identified with the 20K Illumina SNP chip (Bianco et al. 2014; Pagliarani et al. 2016). By performing an *in silico* analysis of the Golden Delicious genome sequence (Velasco et al. 2010), twelve candidate genes putatively involved in RAA resistance were identified in the *Dp-fl* interval (Pagliarani et al. 2016). One sequence was homologous to a Defensin Ec-AMP-D2-like gene involved in reactions triggered in response to the presence of a

foreign body; two showed a Leucine-Rich Repeat (LRR) domain that is one of the classes of resistance (R) genes frequently observed in the Rosaceae family (Zhong et al. 2015); eight showed homology with a TMV resistance protein involved in defense responses; and one was homologous to a gene coding for a protein similar to a pectin acetylesterase family member that was found to be upregulated in Florina after RAA infestation (Qubbaj et al. 2005). However, Golden Delicious is susceptible to RAA, hence a detailed analysis of the *Dp-fl* 'resistant' chromosome of Florina is necessary. Fine mapping and map-based gene cloning together are a good way to approach this analysis and have already been successfully applied in apple to identify candidate genes in specific genomic regions (Patocchi et al. 1999; Galli et al. 2010; Cova et al. 2015).

The resistance cascade after RAA attack was investigated by Qubbaj et al. (2005), using a cDNA-AFLP method. They hypothesized that the Florina resistance against aphids is regulated by signal transduction mechanisms similar to those involved in the response to abiotic and biotic stresses. By the analysis of genes differentially expressed after infestation with *D. plantaginea* in Florina and in the RAA susceptible cultivar Topaz, three genes were assumed to be involved in different signal transduction pathways: a putative vacuolar type H(+)-ATPase, an ADP-ribosylation factor and a putative inositol phosphatase. Other genes that showed an interesting differential expression pattern in aphid infested and non-infested apple trees were a ribulose-1,5-biphosphate-carboxylase (found to be down-regulated), a pectin acetylesterase and a RNase-L-inhibitor (up-regulated in the resistant cultivar Florina upon RAA infestation).

In this paper, we report a detailed analysis of the rosy apple aphid *Dp-fl* resistance region that was sequenced from a minimum tiling path of BAC clones from the cultivar Florina. Through the development of new polymorphic markers, the 'resistant' *Dp-fl* locus was restricted to a physical region of 95 Kb. Structural and functional annotations were performed by comparing the Florina 'resistant' sequence with the homologous region of the Golden Delicious doubled haploid (GDDH13) genome sequence (Daccord et al. 2017). The identification of four candidate resistance genes within the *Dp-fl* window is discussed.

3.3 MATERIALS AND METHODS

3.3.1 Plant material and phenotypic evaluation

A total of nine individuals recombining between the two SNP markers TSP_104 and TSP_585 flanking the *Dp-fl* genomic region previously identified by Pagliarani et al. (2016) were considered (Table 3.1). Five of these recombinant individuals were already characterized by Pagliarani et al. (2016), including PF_X-9504-07 belonging to the Perico x Florina (PF) progeny, which was conserved despite its uncertain genotype (a-) for SNP_104 due to its <ab x ab> status. Three additional recombinant plants belonging to other progenies (Florina x Royal Gala: FR, Florina x Perleberg: FP, and GoldRush x Discovery: GD) were also considered after Baumgartner et al. (2016). Finally, a last recombinant individual (MF_7321) was identified in a Meana x Florina progeny (Dapena and Miñarro, unpublished results).

Table 3.1 List of progenies with abbreviations and particular recombinant individuals with corresponding reference.

Progeny	Abbreviation	Recombinants	Reference
Perico x Florina	PF	PF_P001; PF_X-9504-07; PF_X-9504-33	Pagliarani et al. 2016
Raxao x Florina	RF	RF_X-9104-8	
Florina x Melrose	FM	FM_F145	
Florina x Royal Gala	FR	FR_154	Baumgartner et al. 2016
Florina x Perleberg	FP	FP_21	
GoldRush x Discovery	GD	GD_4	
Meana x Florina	MF	MF_7321	Unpublished

Phenotypic evaluations of the recombinants of the PF, RF and FM populations were already reported by Miñarro and Dapena (2007) and Pagliarani et al. (2016). For the three individuals from FR, FP and GD progenies, phenotypic assessments were first conducted in the spring of 2015 and 2016. The evaluations were made after natural infestation in an unsprayed experimental orchard of Bologna University. The field data were then confirmed under controlled conditions in a greenhouse at INRA, Angers, France, as follows. From 3 to 7 replicates per individual were grafted on MM106 rootstocks and grown in pots at a temperature of about 22°C. Nine replicates of Florina and Gala were also used as resistant and susceptible controls, respectively. Plants were artificially infested by RAA in mid-June when the young shoots were about 20-30 cm. To avoid aphid movements between plants, pots were placed on petri dishes in a plate filled with water, paying attention to prevent shoots overlapping. Infestation was performed placing two young adult apterous virginiparous females on the youngest

and well-expanded leaf of each plant with a paint brush. All aphids were derived from a clonal aphid line of RAA obtained from one founder collected in the field and reared on seedlings raised from open-pollinated Golden Delicious seeds. Scoring of infested plants was done 21 days after infestation, using the following scale: 0 = no leaf distortion; 1= leaf very slightly curled; 2= leaf slightly curled and 3= typically rolled leaf (Rat-Morris & Lespinasse 1995). The average value of the scores assigned to each plant replicate was calculated and only individuals with an average value of 3 were considered as fully susceptible. Finally, the MF_7321 plant was phenotyped at SERIDA, Villaviciosa, Spain with the protocol reported by Pagliarani et al. (2016).

3.3.2 BAC library screening and BAC end sequencing

A bacterial artificial chromosome (BAC) library of Florina available at the University of Bologna (Vinatzer et al. 1998) was screened to identify clones spanning the *Dp-fl* locus. The screening was performed by polymerase chain reaction (PCR) analysis using the bi-dimensional pooling method developed by Cova. (2008). BAC clones carrying both the 'resistant' and 'susceptible' alleles at the *Dp-fl* locus were selected using the two flanking markers (TSP_104 and TSP_585) developed by Pagliarani et al. (2016). Validation of positive BAC clones was done by colony-PCR.

Plasmidic DNA of each selected BAC clone was purified using a Maxi prep - Alkaline lysis protocol adapted as described by Untergasser (2006). BAC-end sequencing was performed by an external service (BIOFAB Research, Rome, Italy) starting from 20 μg of purified plasmids and 24 pmol/μl of either the Sp6 (5'-AGGTGACACTATAGAATACTC-3') or T7 (5'-TAATACGACTCACTATAGGG-3') primers. The sequences were analyzed with Codon Code Aligner software (version 7.0.1) and alignment was performed both on the Golden Delicious v 1.0 genome (Velasco et al. 2010) with the BLASTN tool on the Genome Database of Rosaceae (GDR) website (https://www.rosaceae.org) and on the newly released sequence of the doubled haploid genome of Golden Delicious (GDDH13) available at https://iris.angers.inra.fr/gddh13 (Daccord et al. 2017).

For the chromosome walking, new SSR and SNP markers were developed; all primer sequences are listed in additional material (Additional Material 3.1). SSR_C, SSR_F and SCAR_1 were developed on the Golden Delicious v 1.0 sequence, and SSR_56 and TSP_57 on the sequenced BAC ends. SSR markers were developed in the specific target regions using the 'Tandem Repeats Finder' tool (https://tandem.bu.edu/trf/trf.html, setting the maximum period size at 3 and the minimum copy number at 12). Primers were designed with Primer3 (http://primer3.ut.ee, version 4.0.0,) using default parameters. The SSR genotyping was performed according to Gianfranceschi at al. (1998). Amplified fragments were separated in a 5% denaturing polyacrylamide gel using a 100 bp ladder (Biotium) as size standard. The gel was further stained with the silver staining method and images were acquired using the 440 CF Kodak Image System. SNP markers (SCAR_1 and TSP_57) were scored

using the Temperature Switch PCR (TSP) method (Hayden et al. 2008; Tabone et al. 2009) according to Pagliarani et al. (2016). PCR products were separated in 1.5% (*w/v*) agarose gel electrophoresis using a 100 bp ladder (Biotium) as standard, and images were acquired using the 440 CF Kodak Image System.

3.3.3 Fine mapping of the *Dp-fl* region

Genomic DNA of the nine recombinant individuals listed in Table 3.1 was isolated from young leaf tissues, using the CTAB protocol (Doyle 1987). DNA quantity and quality were measured spectrophotometrically with a Nanodrop ND-8000® (Thermo Scientific, USA). New SSR markers (SSR_377 and SSR_4) were developed from the sequences of the *Dp-fl* region using the procedure described above (Additional Material 3.1). PCR amplifications were performed on the available recombinant plants in a Biorad DNA Engine® thermal cycler in a volume of 11 µL containing 1.1× Qiagen Multiplex PCR Master Mix, 0.2 µM each of forward and reverse primers, and 10 ng DNA. Amplification was done with a touchdown program as follows: a first denaturation step at 94°C for 15 min, four cycles of denaturation (94°C for 30 s), annealing (57°C for 1 min, decreasing by one degree after each cycle), and extension (72 °C for 1 min). The program then continued with 29 cycles of denaturation (94°C for 30 s), annealing (50 °C for 1 min), and extension (72°C for 1 min). A final annealing (50°C for 15 min) and extension step (72°C for 15 min) were added. Electrophoresis of PCR products was performed using a 4-capillary sequencer (ABI 3130; ANAN platform, INRA-Angers). Amplification products were diluted 30 times and mixed (2.5 µL) with formamide (9.35 µL) and Gene Scan 500 LIZ (PE Applied Biosystems) as standard (0.15 µL). After marker scoring, the graphical genotypes (Young and Tanksley 1989) were drawn by combining genotypic and phenotypic data to delineate a reduced *Dp-fl* locus region and shed light on the recombination sites between flanking markers. Marker data on the recombination site were checked at least twice.

3.3.4 BAC clones digestions

Purified plasmid DNA (20 ng) was digested with 10 U EcoRI (Fermentas) for 6 hour at 37°C. Digested DNA fragments, were loaded into 1% agarose gels (Lonza, Basel, Switzerland) loading the 1 Kb DNA ladder (Biotium) as size standard. After electrophoresis in TAE buffer at 35 V overnight, the image was acquired by the Kodak Image System.

3.3.5 BAC clones sequencing and contig assembly

For selected BACs encompassing the 'resistant' *Dp-fl* allele of Florina, plasmidic DNA was extracted using the Qiagen® plasmid purification maxi kit protocol (Qiagen Company). A pool of plasmidic DNA

of three BACs was sent for sequencing. BAC clone sizes were initially estimated by alignment of the BAC ends on the Golden Delicious genome v1.0 and this information was considered for preparing the BAC pool. PacBio sequencing of these BACs was performed by The Genome Analysis Centre (Norwich Research Park, Norwich, UK).

Sequencing data from two PacBio SMRT cells was collected and processed with PacBio's proprietary SMRT Analysis Pipeline v.2.3.0. Contamination screening was performed following PacBio guidelines in the whitelisting tutorial (available at https://github.com/PacificBiosciences/Bioinformatics-Training/wiki/HGAP-Whitelisting-Tutorial). A custom python script (available upon request) was developed to remove the known plasmid sequence from the PacBio .h5 formatted reads. Trimmed reads were finally assembled using HGAP (Chin et al. 2013) with standard parameters and the resulting single contig was polished with quiver (Chin et al. 2013) to produce the final sequence. Quality of the final assembly was assessed by mapping available BAC-end sequences on the newly assembled region with Blast (Altschul et al. 1990).

3.3.6 Structural and functional annotation of *Dp-fl* region

Annotation at both structural and functional levels was performed in the *Dp-fl* region of Florina and these results were compared to that of the homologous region of the Golden Delicious double haploid GDDH13 sequence. The annotation was performed by Fgenesh (Solovyev et al. 2006) and Eugene (Foissac et al. 2003) prediction tools. Each predicted gene and its intron-exon structure was checked and corrected by combining data on detected similarities (blastx) and transcript mapping (EST, cDNA, RNAseq contigs). The sequences of the homologous genes of Florina and GDDH13 were compared through the full length alignment software blast2seq. The Artemis platform was used to combine evidence and produce the final annotation, and its ACT tool (Carver et al. 2008) allowed us to compare the Florina and GDDH13 genomic regions. The functional annotation of present genes was inferred from best homologous genes in the *Arabidopsis* genome (Berardini et al. 2015) and results of Interproscan for conserved domains and gene ontology (Jones et al. 2014; Finn et al. 2016, 2017).

3.4 RESULTS

3.4.1 Phenotypic evaluation of recombinant plants

At INRA-Angers, the Florina control plants were rated as tolerant with a score of 2, while three recombinants (FR_154, FP_21 and GD_4) clearly were susceptible to RAA with an average score of 3, as were Gala control plants. The susceptibility of these three recombinants was also confirmed by the phenotypic data obtained after natural infestation in the unsprayed Bologna orchard. The recombinant MF_7321 was also rated as susceptible to RAA after the resistance tests performed at SERIDA (Spain). In Figure 3.1 are showed Florina (A) and Gala (B) after 15 days from infestation, clearly resistant and susceptible respectively.

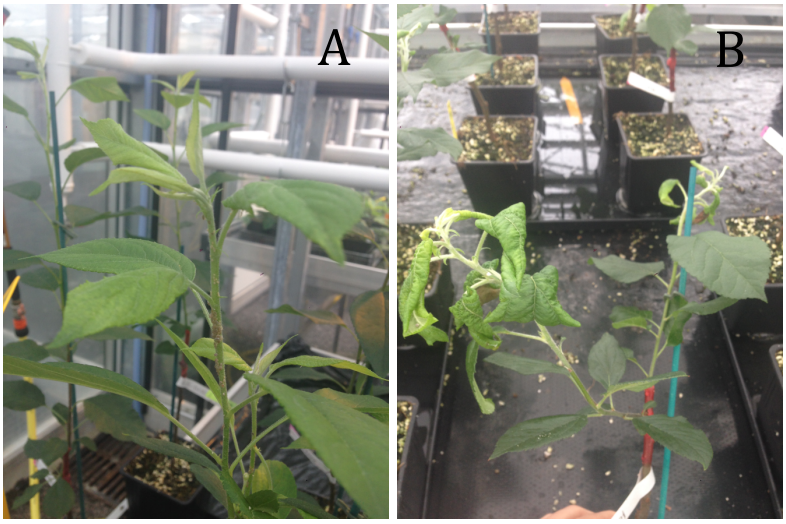

Fig. 3.1 In this figure are showed symptoms of two different genotypes 15 days after the infestation: in the first picture (A) is showed Florina; in the second picture (B) is showed Golden.

3.4.2 Identification of BAC clones encompassing the *Dp-fl* region

The Florina BAC library was screened to identify clones from both the 'resistant' and 'susceptible' chromosomes since this genotype is heterozygous at the *Dp-fl* locus. The results of the genome walking approach are shown in Figure 3.2. The two flanking markers TSP_104 and TSP_585 were used for the first step of BAC library screening. Three BAC clones were identified with TSP_104 on one side of the locus (60G23, 59G11 and 86D23) and another three with TSP_585 on the opposite side (47A15, 57P6 and 75L7). TSP marker polymorphisms made it possible to distinguish the BAC clones coming from the 'resistant' (60G23, 47A15 and 57P6) or 'susceptible' (59G11, 86D23 and 75L7) chromosomes of Florina. To continue the chromosome walking, a second screening step was performed with the newly-developed internal markers, SSR_C, SSR_F, SCAR_1, SSR_56 and TSP_57 (Fig. 3.2). Screening with these markers resulted in the identification of 63M14 and 88H21 BACs from the 'resistant' chromosome and 56D11 and 83P17 from the 'susceptible' one.

Order and orientation of the BACs were confirmed by aligning the BAC end sequences on both the Golden Delicious v 1.0 and GDDH13 genome sequences (Additional Material 3.2). The estimated BAC insert sizes ranged from 81 Kb (clone 86D23) to 149 Kb (clone 47A15). Each BAC end was correctly mapped on chromosome 8, as expected. Overlapping between the various clones was confirmed by sequencing PCR products with primers developed on the BAC ends and by BAC digestion with *Eco*RI, as showed in the following chapter. In summary, five BAC clones fully covering the 'resistant' locus were identified (60G23, 63M14, 88H21, 47A15 and 57P6) while two small gaps remained among the five clones from the 'susceptible' locus (59G11, 86D23, 56D11, 83P17 and 75L7). One gap of about 18 Kb occurs between BACs 86D23 and 56D11 with a second one of about 5 Kb between BACs 83P17 and 75L7 (Fig. 3.2).

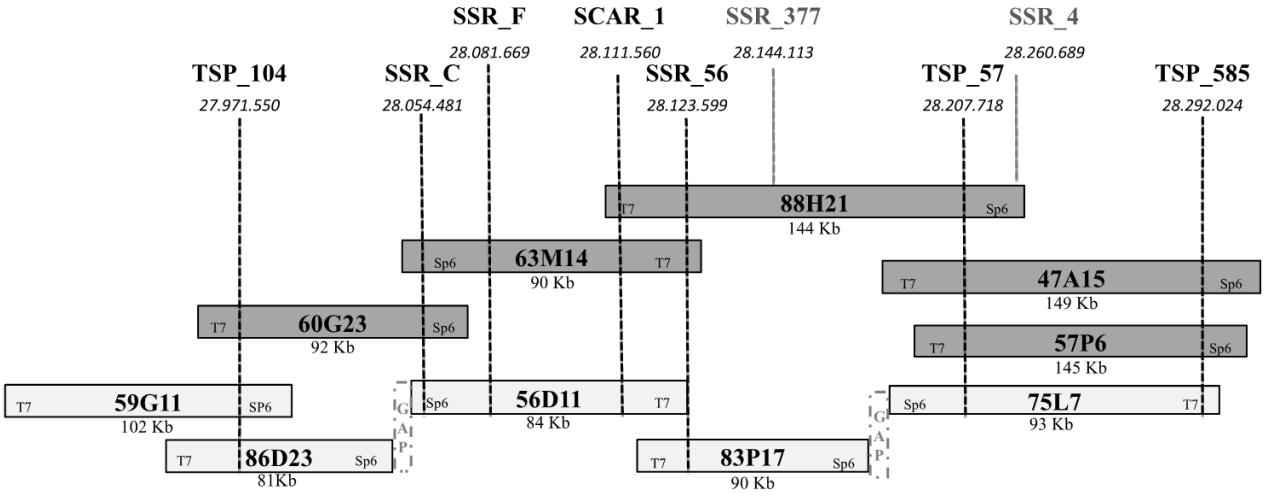

Fig. 3.2: The *Dp-fl* physical map. At the top of the image, markers are listed and ordered with positions on chromosome 8 of the GDDH13 genome. BAC clones covering the 'resistant' and 'susceptible' chromosomes of Florina are represented by dark grey and white rectangles respectively including their orientation with respect to the T7 and Sp6 ends of the cloning vector. The clone sizes as estimated by the alignment of the BAC ends on the GGDH13 genome are reported below the BAC rectangles. GAPs in the 'susceptible' chromosome are represented by two boxes in dashed lines. The region delimited by SSR_377 and SSR_4 was further analyzed to search for candidate genes.

3.4.3 BAC clones digestions

The banding pattern obtained with the *EcoRI* digestion for each BAC clone was useful to confirm the overlapping regions between adjacent clones (Fig. 3.3). In Figure 3.3 are showed both 'resistant' and 'susceptible' BAC clones. In this image, it's clearly visible that the pattern of the two clones 47A15 and 57P6 is very similar, mostly overlapping. Regarding the other BAC clones of the 'resistant' haplotype, some of the bands are shared between adjacent BAC clones (47A15, 57P6, 88H21 and 63M14). On the left side of the gel only 59G11 and 86D23 from the 'susceptible' haplotype show some common bands while the other two clones do not overlap (Fig. 3.3). In this figure, is missing the digestion of the susceptible BAC clone 83P17 that should appear between the BAC clone 56D11 and the 75L7, covering the central part of the *Dp-fl* region. The digestion patterns was also used to estimate the BAC sizes.

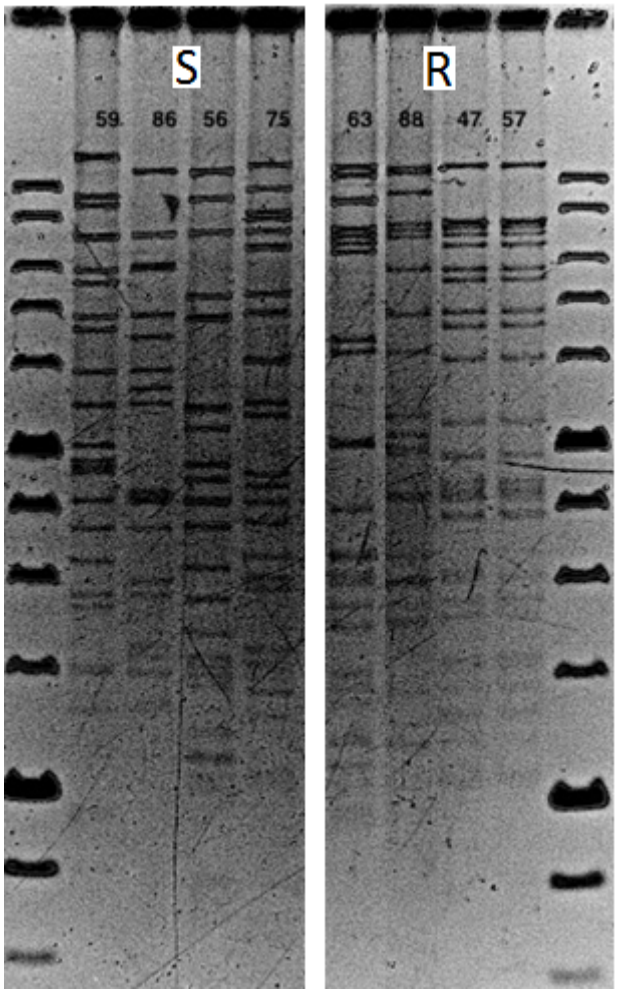

Fig. 3.3 BAC clones digestions. BAC clones from the susceptible chromosome of Florina: 59 – 59G11; 86 – 86D23; 56 – 56 D11; 75 – 75L7. BAC clones from the resistant chromosome of Florina: 88 – 88H21; 63 – 63M14; 47 – 47A15; 57 – 57P6. On both sides of the gel is loaded the 1 Kb DNA ladder (Biotium) as size standard.

3.4.4 Shrinkage of the *Dp-fl* window

After the identification of BAC clones that fully cover the 'resistant' phase of the *Dp-fl* locus, two other SSRs (SSR_377 and SSR_4) were developed. The polymorphism of both markers was confirmed by PCR on each BAC clone of the 'resistant' and 'susceptible' chromosomes. Genotyping of the nine recombinant individuals with these two markers allowed drawing of the graphical genotyping within the *Dp-fl* region (Fig 3.4).

	MARKERS							
RECOMBINANT	SNP_104	SSR_C	SSR_F	SCAR_1	SSR_377	***Dp-fl***	SSR_4	SNP_585
FR_154	ab	ab	ab	ab	ab	**S**	aa	aa
FP_21	ab	ab	ab	ab	aa	**S**	aa	aa
GD_4	ab	ab	ab	aa	aa	**S**	aa	aa
MF_7321	ab	ab	aa	aa	aa	**S**	aa	aa
FM_F145	ab	ab	aa	aa	aa	**S**	aa	aa
RF_X-9104-8	aa	aa	aa	aa	aa	**S**	ab	ab
PF_P001	aa	aa	aa	aa	aa	**R**	ab	ab
PF_X-9504-33	aa	aa	aa	-	aa	**R**	aa	aa
PF_X-9504-07	-	ab	ab	-	ab	**R**	ab	ab

	MARKERS							
PARENTAL CULTIVAR	SNP_104	SSR_C	SSR_F	SCAR_1	SSR_377	***Dp-fl***	SSR_4	SNP_585
Gala	aa	aa	aa	aa	aa	**S**	aa	aa
Golden Delicious	aa	aa	aa	aa	aa	**S**	aa	aa
GoldRush	ab	ab	ab	ab	ab	**R**	ab	ab
Florina	ab	ab	ab	ab	ab	**R**	ab	ab
Malus floribunda #821	ab	ab	ab	ab	ab	**R**	ab	ab
Melrose	aa	aa	aa	aa	aa	**S**	aa	aa
Raxao	aa	aa	aa	aa	aa	**S**	aa	aa
Perico	ab	aa	aa	ab	aa	**S**	aa	aa
Meana	aa	aa	aa	ab	aa	**S**	aa	aa
Perleberg 3	aa	aa	aa	aa	aa	**S**	aa	aa
Discovery	aa	aa	aa	aa	aa	**S**	aa	aa
Royal Gala	aa	aa	aa	aa	aa	**S**	aa	aa

Fig. 3.4 Graphical genotyping of *Dp-fl* region. Genotypic and phenotypic data of the recombinant and parental individuals of the different populations are reported. The allele 'aa' (red cell) is related to susceptibility and 'ab' (green cell) to resistance. 'R' is the resistant phenotype and 'S' the susceptible one. Dash in the white cell (-) is used for the markers segregating with 'abxab' pattern.

Four recombinant individuals (FR_154, FP_21, GD_4 and MF_7321) exhibit a recombination event on the left side of the postulated position of the *Dp-fl* locus, with FR_154 being the most crucial one since the recombination event was observed between SSR_337 and the locus of interest. Among the five recombinant individuals already characterized by Pagliarani et al. (2016), RF_X-9104-8 was also crucial for positioning the *Dp-fl* locus since a recombination event was observed between the locus and the next marker SSR_4. PF_P001 also showed a recombination event just next to the *Dp-fl* locus similarly to FR_154, but this plant is resistant to RAA and lack alleles in coupling with resistance. FM_F145 and MF_7321 showed the same marker pattern. PF_X-9504-33 was finally discarded from

the analysis since the genotype data for the SNP_585 indicated by Pagliarani et al. (2016) was not confirmed. A new phenotype test for this plant would have been necessary, but that was not possible as the plant was no longer available. PF_X-9504-07 did not recombine in the region flanked by SSR_C and SNP_585 (Fig. 3.4). Finally, after the analysis of the graphical genotyping, the *Dp-fl* gene was located between the markers SSR_377 and SSR_4.

The resistant parents (Florina, GoldRush and *Malus floribunda* #821) and the susceptible reference cultivars were consistently genotyped as 'ab' or 'aa', respectively, for all the markers, as expected. A couple of exceptions were detected in Perico ('ab' for SNP_104 and SCAR_1) and Meana ('ab' for SCAR_1). Therefore, the genotypes of the recombinants derived from crosses of the 'abxab' pattern cannot be clearly determined for the heterozygous genotypes (Fig. 3.4).

3.4.5 BAC clones sequencing and contig assembly of the Florina *Dp-fl* region

Three BAC clones (63M14, 88H21 and 47A15) were finally chosen for sequencing during the genome walking approach before having reduced the targeted genomic region to the interval flanked by markers SSR_377 and SSR_4. The pool of these BAC clones was successfully sequenced by two PacBio SMRT cells, resulting in the sequencing of 373 megabases that correspond to a 1000x coverage of the region.

The final assembly of the three BAC clones resulted in a single contig of 278,911 bp covering the *Dp-fl* locus of Florina, which is available from the EMBL database. By locating the BAC-end sequences on the assembled contig, the sizes of the three clones were calculated (about 109.7, 108.2 and 114.7 Kb for 63M14, 88H21 and 47A15, respectively), pointing out to possible structural differences with the BAC sizes estimated on the GDDH13 sequence.

3.4.6 Structural and functional annotation of *Dp-fl* locus

To perform a structural analysis, the sequences of the 'resistant' and 'susceptible' alleles of the *Dp-fl* locus were compared. The 95 Kb of the *Dp-fl* region of Florina was shorter than the corresponding sequence on the GDDH13 genome (116.8 Kb from 28,144,113 to 28,260,934 bp of chromosome 8) as reported in Fig. 3.5. There were five main structural changes: two regions (1 and 2) were present only in Florina, while three (3, 4 and 5) were present only in the GDDH13 sequence (see the green boxes in Fig. 3.5).

The annotation of the 95 Kb of the *Dp-fl* 'resistant' region of Florina resulted in the identification of a total of eleven genes and four vestiges of transposable elements (TE). Seven of the eleven predicted genes were putatively complete functional genes and four resulted as pseudogenes (i.e. coding

sequences disrupted by a frameshift and/or stop codon, or presence of only a portion of the complete coding sequence). The eleven predicted genes in the Florina 'resistant' locus (FLO-) are listed in Table 3.2, numbered starting from SSR_377 to SSR_4, specifying their position and their homolog gene in the GDDH13 genome. For seven genes (FLO-3, FLO-4, FLO-5, FLO-6, FLO-8, FLO-10 AND FLO-11) homologous genes in GDDH13 were identified and each corresponding code is indicated. No identical protein sequences were identified from the comparison of the homologous genes of Florina and GDDH13. The percentage of identity and similarity are above 98% for most of the genes (Additional Material 3.3).

Despite the structural changes, there were no substantial annotation differences between the Florina and Golden Delicious (GDDH13) *Dp-fl* sequences. In the two regions present only in Florina (1 and 2), a pseudogene (FLO-7) and a transposable element were predicted. Instead, in the three GDDH13 regions not present in Florina (3, 4, and 5), three pseudogenes and two transposable elements were predicted. An additional annotation difference was observed: the gene model MD08G1219700 of GDDH13 was not predicted by the Fgenesh software used for Florina, even if the sequence was conserved. Nevertheless, this gene sequence has no significant similarities with the Uniprot databank.

The functional annotations inferred from the best homologous gene in the *Arabidopsis* genome is reported in Table 3.2. In particular, the predicted gene FLO-3 belongs to the C2 calcium/lipid-binding plant phosphoribosyltransferase protein family. FLO-4 is predicted to encode for a Methylthiotransferase protein that is involved in tRNA modifications. FLO-5 encodes for a protein containing a CBS/octiosapeptide/Phox/Bemp1 (PB1) domain that is a structural component of the membrane. FLO-6 encodes for a Protease-associated (PA) RING/U-box zinc finger protein that belongs to a family of vacuolar sorting receptors involved in the secretory pathway of the Trans-Golgi network. FLO-8 is a SAC domain phosphoinositide (3,5) P2 phosphatase localized at Golgi apparatus, and is required for normal morphogenesis and cell wall synthesis and actin organization. FLO-9 is a partial duplication of protein FLO-8, but is not functional. FLO-10 encodes for a protein that belongs to the Fantastic Four meristem regulator (FAF) protein family, a group of proteins that regulate the size and shoot meristem. Lastly, the gene FLO-11 encodes for 5S rRNA, another structural component of the ribosome.

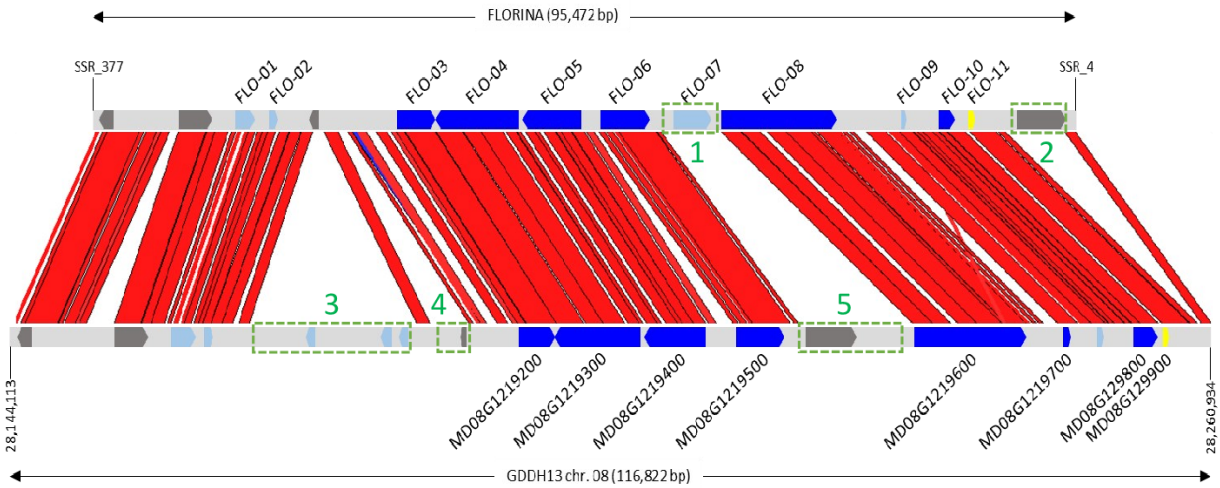

Fig. 3.5 Sequence comparison between the 95Kb region of Florina and the corresponding sequence of the GDDH13 genome. Red shapes represent highly conserved blocks detected with the blastn algorithm and displayed through the Artemis Comparison Tool (Carver et al. 2008). Colored arrows represent the results of the curated structural annotation: grey ones are vestiges of transposable elements, dark blue ones are complete coding genes, light blue ones are pseudogenes (coding sequences are partial or disrupted by frameshift and/or stop codon), and the yellow one is an rRNA gene. The name of the genes in GDDH13 are specified. Green boxes represent major differences between the Florina and GDDH13 sequences.

Table 3.2 Genes and pseudogenes detected in the 95 Kb of the *Dp-fl* region of Florina and their functional annotation. Their localization is relative to the targeted region of 95 Kb between SSR_377 and SSR_4, the corresponding gene name in GDDH13 is specified if possible. Biochemical functions are described through homologies and relative PFAM (Finn et al. 2016) and INTERPRO (Finn et al. 2017) accession numbers are listed.

Gene Name	Localization in 95 Kb region	GDDH13 gene	Functional Annotation	PFAM ID	INTERPRO ID
FLO-01	13952..15270 (+ strand)	nd	*unknown, pseudogene*	nd	nd
FLO-02	17257..17584 (+ strand)	nd	*agamous-like, pseudogene*	nd	nd
FLO-03	29663..32770 (+ strand)	MD08G1219200	C2 calcium/lipid-binding plant phosphoribosyltransferase	PF00168	IPR000008, IPR013583
FLO-04	33438..41294 (- strand)	MD08G1219300	Methylthiotransferase	PF00919, PF04055	IPR005839, IPR006466
FLO-05	42265..47511 (- strand)	MD08G1219400	CBS / octicosapeptide/Phox/Bemp1 (PB1) domains-containing protein	PF00571, PF00564	IPR034896
FLO-06	49440..53617 (+ strand)	MD08G1219500	Protease-associated (PA) RING/U-box zinc finger protein	PF02225, PF13639	IPR003137, IPR013083
FLO-07	56654..59549 (+ strand)	nd	*nucleolar protein, ribosome biogenesis co-factor, rRNA processing, pseudogene*	nd	nd
FLO-08	61144..71778 (+ strand)	MD08G1219600	Phosphoinositide phosphatase protein, cell wall synthesis, actin organization	PF02383	IPR030213
FLO-09	78706..78897 (+ strand)	nd	*Phosphoinositide phosphatase, pseudogene*	nd	nd
FLO-10	82291..83238 (+ strand)	MD08G1219800	unknown protein (FAF domain)	PF11250	IPR021410
FLO-11	85241..85347 (+ strand)	MD08G1219900	5S rRNA	nd	nd

3.5 DISCUSSION

In this study the genomic location of the *Dp-fl* locus of Florina conferring tolerance to RAA has been refined. A chromosome walking approach made it possible to narrow down the putative region harboring the locus to a segment of 95 Kb. Seven coding gene models have been predicted in this region, allowing a limited number of candidate genes to be highlighted with predicted functions putatively linked to aphid-apple interactions.

3.5.1 Graphical genotyping and new markers for the Dp-fl resistance locus

Phenotyping recombinant individuals is a key step in the map-based gene cloning approach. Evaluation of the RAA resistance was performed on the plants showing a recombination between TSP_104 and TSP_585, bracketing the *Dp-fl* locus. New recombinant individuals were added and successfully phenotyped in both field and controlled conditions. A genotype-phenotype incongruent (GPI) (Gygax et al. 2004) individual (PF_X-9504-33) was excluded from further analysis, and another (PF_X-9504-07) appeared not to be a recombinant in the targeted region. Intriguingly, most of the individuals recombining in the *Dp-fl* locus region were susceptible to RAA (6/7).

The resistance region was further restricted by developing two additional markers (SSR_377 and SSR_4). Analyzing the graphical genotyping in detail, five recombinants on the SSR_377 side were identified that precisely marked one border of the resistance locus. Indeed, the presence of one or more alleles coupled with resistance for markers SNP_104, SSR_C, SSR_F, SCAR_1 and SSR_377 was not associated with the resistant phenotype, so the *Dp-fl* locus should be located downstream of the marker SSR_377. However, the two remaining recombinant individuals (RF_X-9104-8 and PF_001) were also carrying alleles coupled with resistance for both markers SSR_4 and SNP_585, but RF_X-9104-8 was susceptible whereas PF_001 was resistant to RAA. The position of the *Dp-fl* locus upstream of SSR_4 was thus less strongly supported by phenotypic data. A final conclusion on this aspect may be reached by identifying new recombinants on the SSR_4 border.

New markers have been developed during the genome walking approach. Especially, the new SSR markers SSR_377 and SSR_4 can be used to predict the presence/absence of the favorable 'resistant' allele inherited from Florina and *M. floribunda* #821 in new genetic material and they can thus be extremely useful for more efficient breeding.

3.5.2 Candidate genes for the tolerance phenotype

The search of the genetic bases of RAA tolerance of Florina can be facilitated by the availability of the full sequences from both the 'resistance' and 'susceptibility' regions at the *Dp-fl* locus. Despite the structural differences between the *Dp-fl* sequence of Florina and GDDH13, no functional genes were predicted within these differential regions. From the annotation computed for the Florina sequence,

seven putatively functional coding genes were predicted. The lack of marked differences between the protein sequences from Florina and GDDH13 are not sufficient to exclude any of the seven candidate genes. Furthermore, differences in the expression of each gene due to differences in the promoter sequences cannot be excluded.

Unexpectedly, none of the seven predicted genes code for known resistance proteins including genes of the NBS-LRR family that were already reported as involved in the aphid resistance mechanisms of tomato and melon (Rossi et al. 1998; Dogimont et al. 2008). Therefore, a different resistance mechanism for RAA have to be hypothesized in Florina. Among the seven candidate genes, four have a biological function that could be related to the RAA attack.

Gene FLO-3 encodes for a protein belonging to the C2 calcium/lipid-binding plant phosphoribosyltransferase family. A member of this family in *A. thaliana*, the Quirky (QKY) protein, is involved in cell-to-cell communications that control cell patterning, organ shape and development. It has been shown that *qky*-mutants of *A. thaliana* have twisted petals and leaves, and an aberrant floral phyllotaxis (Fulton et al. 2009). A recent study co-locates QKY protein at the plasmodesmata with an LRR-receptor-like kinase protein called STRUBBELIG, but independently from one another. This positioning seems to play a central role in cell-to-cell communication and in the growth of plant cells (Vaddepalli et al. 2014). In vascular plants, such as apple, plasmodesmata permit the movement of sap through the sieve element. During probing and feeding, the aphids secrete saliva directly into the phloem (Miles 1999), causing leaf and shoot modifications that could be correlated with the QKY gene. In fact, a modification occurring at the plasmodesmata level could negatively or positively affect phloem transport and hence aphid feeding. The involvement of the QKY protein could also be correlated to the symptoms, such as leaf deformations, due to the salivary secretion of RAA.

FLO-5 encodes for a protein containing a CBS/octiosapeptide/Phox/Bemp1 (PB1) domain that is a structural component of the membrane. CBS domains are small intracellular modules that pair together to form a stable globular domain. Instead, the PB1 domain is present in many eukaryotic cytoplasmatic signaling proteins and is involved in specific protein-protein interactions. Kushwaha et al. (2009) suggested that these proteins might be involved in cellular signaling processes through interaction with other proteins and/or ligands (ATP, ADP or SAM). The CBS domain is known to be involved in a regulatory role for many enzymes and thus helps in maintaining the intracellular redox balance that is suggested as being directly related with plant stress response in *A. thaliana* and *Oryza sativa* (Kushwaha et al. 2009).

FLO-6 encodes for a Protease-associated (PA) RING/U-box zinc finger protein that belongs to a family of vacuolar sorting receptors, which in *A. thaliana* seems to be involved in the secretory pathway of the Trans-Golgi network. In *A. thaliana*, the expression of a gene with a similar structure

(RING-H2zinc-finger gene - ATL2) has been shown to be directly involved in defense against abiotic and biotic stresses (Serrano and Guzmán 2004).

Another very interesting gene is FLO-8 encoding for a SAC domain phosphoinositide (3,5) P2 phosphatase. This enzyme is involved in the release of free inositol, an important metabolite required for normal cell growth and other critical functions. Bohnert et al. (1995) reported that inositol and inositol-l-phosphate enzyme increase the production of compounds, such as gums, mucilagins, cell wall-located carbohydrates and carbohydrates in glycoproteins, which have been correlated with stress tolerance. In the study by Qubbaj et al. (2005), a phosphoinositide phosphatase was shown to be upregulated in Florina after aphid infestation.

3.6 CONCLUSION

In this research the RAA resistance locus has been mapped in a specific region of about 95 Kb of chromosome 8 of Florina and specific molecular markers linked to this trait have been developed. Four genes putatively involved in the RAA resistance have been identified. Some of these candidate genes have a putative biological function that might explain the leaf deformations occurring in susceptible cultivars under aphid attack in comparison with Florina. Among these, the FLO-3 gene, which is similar to the Quirky gene, seems to be interesting for its location at the plasmodesmata level, which could affect phloem sap movement and its availability to RAA. Moreover, its involvement in organ development and shaping could explain the typical leaf, stem and fruit deformations after RAA feeding. The phosphoinositide phosphatase gene FLO-8 could also be an interesting candidate gene given that a gene with a similar function was found to be upregulated in Florina after RAA infestation (Qubbaj et al. 2005). Further functional studies by genetic transformation are needed to fully understand the role of these genes in RAA resistance.

3.7 Additional material

Additional Material 3.1. Primer names, sequences, amplicon sizes and the different sources used for the experimental design. Golden Delicious v 1.0 is available at the GDR database (Velasco et al. 2010); GDDH13 corresponds to the Golden Delicious doubled-haploid 13 sequence (Daccord et al. 2017).

Primer Name		Primer sequence 5'-3'	Amplicon size (bp)	Source
SSR_C	For	TGGCGGTCTCCTTTTGTTCA	259	Golden Delicious v 1.0 genome
	Rev	ACCCATCAATCATCATCCTACCA		
SSR_F	For	GTGGTTTGGTAGTGGCTGCT	200	Golden Delicious v 1.0 genome
	Rev	ATTTCCCAGCCTCCAGTTGG		
SCAR_1	For	TGACTCAGACAATGATCCCAAT	741/246	Golden Delicious v 1.0 genome
	Rev	AGGAGGCATTTCGGGTATCT		
	TSP	tgcTAAGGAATGCATGGa		
SSR_56	For	ATTCCTCGTACCCAGGTCCT	380	BAC end 56D11_T7 sequence
	Rev	CCAGCACTAGATCCTTGCCC		
TSP_57	For	TCAAGACCCCAGATTTCTAGACA	435/241	BAC end 57P6_T7 sequence
	Rev	CCCCTATGCGTCAAACAATCA		
	TSP	tgcACCTGGATTAACAGT		
SSR_377	For	TGAAGCTCTTGGCTTCAATGAC	238/270	GDDH13 genome
	Rev	GGTTCTCAGCCCTCTTATCTTCC		
SSR_4	For	CTCCACAATGATGCTTAGTTGGT	228/266	GDDH13 genome
	Rev	GGAATGAAATCAGCCAGATGGG		

Additional Material 3.2 Positions of the BAC ends and markers (highlighted in grey) on the Golden Delicious doubled haploid (GDDH13) genome and position in the contig of Golden Delicious genome v1.0. TSP markers were developed from the 20 K SNP array (Bianco et al. 2014; Pagliarani et al. 2016).

Marker/BAC end name	Notes	Position on GDDH LG8		Contig on the Golden Delicious v1.0
59G11T7	BAC end sequence	27,874,009	27,874,583	MDC021726.395
86D23T7	BAC end sequence	27,940,789	27,941,269	MDC002350.211
60G23T7	BAC end sequence	27,967,617	27,967,926	MDC022778.432
TSP_104	20k SNP array	27,971,550	27,971,669	MDC022778.432
59G11Sp6	BAC end sequence	27,977,314	27,976,549	MDC012449.339
86D23Sp6	BAC end sequence	28,021,809	28,021,336	MDC018566.86
63M14Sp6	BAC end sequence	28,038,627	28,039,139	MDC022778.400
56D11Sp6	BAC end sequence	28,039,035	28,039,093	MDC002590.316
SSR_C	SSR from GD v 1.0 genome	28,054,481	28,054,225	MDC016445.143
60G23Sp6	BAC end sequence	28,060,323	28,059,578	MDC016445.143
SSR_F	SSR from GD v 1.0 genome	28,081,669	28,081,868	MDC003432.281
88H21T7	BAC end sequence	28,110,734	28,111,165	MDC003432.281
SCAR_1	SSR from GD v 1.0 genome	28,111,560	28,112,043	MDC003432.281
83P17T7	BAC end sequence	28,111,726	28,112,168	MDC003432.281
SSR_56	SSR from BAC ends	28,123,599	28,123,248	MDC002590.316
56D11T7	BAC end sequence	28,123,609	28,122,904	MDC022778.400
63M14T7	BAC end sequence	28,128,840	28,128,548	MDC008847.370
SSR_377	SSR from GDDH13 genome	28,144,113	28,144,937	MDC002325.377
83P17Sp6	BAC end sequence	28,202,018	28,201,588	MDC010553.228
47A15T7	BAC end sequence	28,205,668	28,206,036	MDC010553.228
75L7Sp6	BAC end sequence	28,206,626	28,206,884	MDC010553.228
57P6T7	BAC end sequence	28,207,492	28,207,962	MDC010553.228
TSP_57	SNP from BAC end	28,207,718	28,207,956	MDC010553.228
88H21Sp6	BAC end sequence	28,255,389	28,254,691	MDC012135.190
SSR_4	SSR from GDDH13 genome	28,260,689	28,260,934	MDC034963.3
TSP_585	20k SNP array	28,292,024	28,291,743	MDC015891.57
75L7T7	BAC end sequence	28,300,559	28,300,073	MDC043089.3
57P6Sp6	BAC end sequence	28,352,809	28,352,535	MDC002411.325
47A15Sp6	BAC end sequence	28,354,803	28,354,421	MDC002411.325

Additional Material 3.3 Percentage (%) of similarity and identity between each GDDH13 and Florina homologous gene performed by comparing the protein sequences on blast2seq software. In the first two columns are detailed the names and sizes of each protein sequence for both Florina and GDDH13.

Florina gene/protein (size)	GDDH13 gene/protein (size)	Identity	Similarity
FLO-3 (1035 aa)	MD08G1219200 (1036 aa)	99.03 %	99.61 %
FLO-4 (631 aa)	MD08G1219300 (631 aa)	98.89 %	99.20 %
FLO-5 (542 aa)	MD08G1219400 (542 aa)	99.63 %	99.81 %
FLO-6 (447 aa)	MD08G1219500 (447 aa)	99.32 %	99.55 %
FLO-8 (905 aa)	MD08G1219600 (905 aa)	98.89 %	99.56 %
FLO-10 (315 aa)	MD08G1219800 (328 aa)	92.99 %	93.63 %
FLO-11 (107 bp)	MD08G1219900 (107 bp)	94.39 %	-

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403-410.

Andreev R, Kutinkova H, Rasheva D (2012) Non-chemical control of *Aphis spiraecola* Patch. and *Dysaphis plantaginea* Pass. on apple. J Biopestic 5:239-242.

Angeli G, Simoni S (2006) Apple cultivars acceptance by *Dysaphis plantaginea* Passerini (Homoptera : *Aphididae*). J Pest Sci, 79:175-179.

Arnaoudov V, Kutinkova H (2006). Susceptibility of some apple cultivars to infestation by the rosy apple aphid (*Dysaphis plantaginea* Pass., Homoptera: *Aphididae*). J Fruit Ornam Plant Res 14, 137-142.

Aslan B, Karaca I (2005). Fruit tree aphids and their natural enemies in Isparta region, Turkey. J Pest Sci 78:227-229.

Baumgartner IO, Kellerhals M, Costa F, Dondini L, Pagliarani G, Gregori R, Tartarini S, Leumann L, Laurens F, Patocchi A. (2016). Development of SNP-based assays for disease resistance and fruit quality traits in apple (*Malus* × *domestica* Borkh.) and validation in breeding pilot studies. Tree Genet Genom 12:1-21.

Berardini T, Reiser L, Li D, Mezheritsky Y, Muller R, Strait E, Huala E (2015) The *Arabidopsis* Information Resource: Making and mining the "gold standard" annotated reference plant genome. Genes 53:474-485.

Bianco L, Cestaro A, Sargent DJ, Banchi E, Derdak S, Di Guardo M, Salvi S, Jansen J, Viola R, Gut I, Laurens F, Chagné D, Velasco R, van de Weg E, Troggio M (2014) Development and validation of a 20K single nucleotide polymorphism (SNP) whole genome genotyping array for apple (*Malus* x *domestica* Borkh). PLoS One 9(10) e110377. doi:10.1371/journal.pone.0110377

Bohnert HJ, Nelson DE, Jensen RG (1995). Adaptations to environmental stresses. Plant Cell, 7(7), 1099.

Brown MW, Mathews CR (2007) Conservation biological control of rosy apple aphid, *Dysaphis plantaginea* (Passerini), in Eastern North America. Environ Entomol, 36:1131-9.

Carver T, Berriman M, Tivey A, Patel C, Böhme U, Barrell BG, Parkhill J, Rajandream MA (2008) Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. Bioinform 24:2672-6.

Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler E E, Turner SW, Kornach J (2013) Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563-569.

Cova V (2008) Isolamento di marcatori microsatelliti strettamente associati al gene di resistenza a ticchiolatura *Vm* in melo. Dissertation, Università di Bologna

Cova V, Bandara NL, Liang W, Tartarini S, Patocchi A, Troggio M, Velasco R, Komjanc M (2015) Fine mapping of the *Rvi5* (*Vm*) apple scab resistance locus in the 'Murray' apple genotype. Mol Breed, 35:200.

Daccord N, Celton JM, Linsmith G, Becker C, Choisne N, Schijlen E, van de Geest H, Bianco L, Micheletti D, Velasco R, Di Pierro EA, Gouzy J, Rees DJG, Guérif P, Muranty H, Durel CE, Laurens F, Lespinasse Y, Gaillard S, Aubourg S, Quesneville H, Weigel D, van de Weg E, Troggio M, Bucher E (2017) High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. Nat Genet 49:1099-1106.

Dangl JL, Jones JD (2001). Plant pathogens and integrated defence responses to infection. Nat 411:826-833.

Dapena E, Miñarro M, Blasquez MD (2009) Evaluation of the resistance to the rosy apple aphid using a genetic marker. ISHS Acta Hortic 814:787-790.

De Berardinis E, Baronio P, Baumgartner J (1994) The effect of aphid *Dysaphis plantaginea* (Pass, Hom, Aphididae) feeding on apple fruit growth. Ecol Modell 72:115-127.

Dib H, Simon S, Sauphanor B, Capowiez Y (2010). The role of natural enemies on the population dynamics of the rosy apple aphid, *Dysaphis plantaginea* Passerini (Hemiptera: *Aphididae*) in organic apple orchards in south-eastern France. Biol Control 55:97-109.

Dogimont C, Chovelon V, Tual S, Boissot N, Rittener-Rüff V, Giovinazzo N, Bendahmane A, Pitrat M (2008). Molecular diversity at the *Vat*/*Pm*-*W* resistance locus in melon. Cucurbitaceae 2008 219-228.

Doyle JJ (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11-15.

Fidantsef AL, Stout MJ, Thaler JS, Duffey SS, Bostock RM (1999). Signal interactions in pathogen and insect attack: expression of lipoxygenase, proteinase inhibitor II, and pathogenesis-related protein P4 in the tomato, *Lycopersicon esculentum*. Physiol Mol Plant Pathol, 54:97-114.

Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang HY, Dosztányi Z, El-Gebali S, Fraser M, Gough J, Haft D, Holliday GL, Huang H, Huang X, Letunic I, Lopez R, Lu S, Marchler-Bauer A, Mi H, Mistry J, Natale DA, Necci M, Nuka G, Orengo CA, Park Y, Pesseat S, Piovesan D, Potter SC, Rawlings ND, Redaschi N, Richardson L, Rivoire C, Sangrador-Vegas A, Sigrist C, Sillitoe I, Smithers B, Squizzato S, Sutton G, Thanki N, Thomas PD, Tosatto SC, Wu CH, Xenarios I, Yeh LS, Young SY, Mitchell AL (2017) InterPro in 2017-beyond protein family and domain annotations. Nucleic Acids Res 45:D190-D199.

Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A (2016) The *Pfam* protein families database: towards a more sustainable future. Nucleic Acids Res 44:D279-85.

Foissac S, Bardou P, Moisan A, Cros MJ, Schiex T (2003) EUGENE'HOM: A generic similarity-based gene finder using multiple homologous sequences. Nucleic Acids Res. 31:3742-5.

Fulton L, Batoux M, Vaddepalli P, Yadav RK, Busch W, Andersen SU, Jeong S, Lohmann JU, Schneitz K. (2009). DETORQUEO, QUIRKY, and ZERZAUST represent novel components involved in organ development mediated by the receptor-like kinase STRUBBELIG in *Arabidopsis thaliana*. PLoS Genet 5(1), e1000355.

Galli P, Broggini GAL, Kellerhals M, Gessler C, Patocchi A (2010). High-resolution genetic map of the *Rvi15* (*Vr2*) apple scab resistance locus. Mol Breed, 26: 561-572.

Gianfranceschi L, Seglias N, Tarchini R, Komjanc M, Gessler C (1998). Simple sequence repeats for the genetic analysis of apple. Theor Appl Genet 96:1069-1076.

Gygax M, Gianfranceschi L, Liebhard R, Kellerhals M, Gessler C, Patocchi A (2004). Molecular markers linked to the apple scab resistance gene *Vbj* derived from *Malus baccata* jackii. Theor Appl Genet, 109:1702-1709.

Hayden M. J, Nguyen T. M, Waterman A, McMichael GL, Chalmers KJ (2008). Application of multiplex-ready PCR for fluorescence-based SSR genotyping in barley and wheat. Mol Breed, 21:271-281.

Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S (2014) InterProScan 5: genome-scale protein function classification. Bioinform. 30:1236-40.

Kushwaha HR, Singh, AK, Sopory SK, Singla-Pareek SL, Pareek A (2009). Genome wide expression analysis of CBS domain containing proteins in *Arabidopsis thaliana* (L.) Heynh and *Oryza sativa* L. reveals their developmental and stress regulation. BMC Genom 10:200.

Lespinasse Y, Olivier JM, Lespinasse JM, Le Lezec M (1985) Florina Querina®, la résistance du pommier à la tavelure.- Arboric Fruit 33:43-46.

Marchetti E, Civolani S, Leis M, Chicca M, Tjallingii WF, Pasqualini E, Baronio P (2009). Tissue location of resistance in apple to the rosy apple aphid established by electrical penetration graphs. Bull Insectology, 62:203-208.

Miles PW (1999) Aphid saliva. Biol Rev Camb Philos Soc 74:41-85.

Miñarro M, Dapena E (2007). Resistance of apple cultivars to *Dysaphis plantaginea* (Hemiptera: Aphididae): role of tree phenology in infestation avoidance. Environ Entomol 36:1206-1211.

Miñarro M, Dapena E. (2008) Tolerance of some scab-resistant apple cultivars to the rosy apple aphid, *Dysaphis plantaginea*. Crop Prot 27:391-395.

Miñarro M, HemptinneJ-L, Dapena E (2005). Colonization of apple orchards by predators of *Dysaphis plantaginea*: sequential arrival, response to prey abundance and consequences for biological control. BioControl 50:403-414.

Martín B, Collar JL, Tjallingii WF, Fereres A (1997). Intracellular ingestion and salivation by aphids may cause the acquisition and inoculation of non-persistently transmitted plant viruses. J Gen Virol 78:2701-2705.

Ng J C, Perry K L (2004). Transmission of plant viruses by aphid vectors. Mol Plant Pathol 5:505-511.

Pagliarani G, Dapena E, Miñarro M, Denancé C, Lespinasse Y, Rat-Morris E, Troggio M, Durel CE, Tartarini S (2016). Fine mapping of the rosy apple aphid resistance locus *Dp-fl* on linkage group 8 of the apple cultivar 'Florina'. Tree Genet Genom, 12:1-12.

Parisi L, Gros C, Combe F, Parveaud CE, Gomez C, Brun L. (2013). Impact of a cultivar mixture on scab, powdery mildew and rosy aphid in an organic apple orchard. Crop Prot 43:207-212.

Patocchi A, Vinatzer BA, Gianfranceschi L, Tartarini S, Zhang HB, Sansavini S, Gessler C (1999) Construction of a 550 kb BAC contig spanning the genomic region containing the apple scab resistance gene *Vf*. Mol Gen Genet 262:884-891.

Qubbaj T, Reineke A, Zebitz CPW (2005) Molecular interactions between rosy apple aphids, *Dysaphis plantaginea*, and resistant and susceptible cultivars of its primary host *Malus domestica*. Entomol Exp Appl, 115:145-152.

Rat-Morris E (1993) Development of rosy aphid *Dysaphis plantaginea* Pass. on a tolerant apple cultivar 'Florina'. IOBC/WPRS Bull 16 :91-100.

Rat-Morris E (1994) Analyse des relations entre *Dysaphis plantaginea* Pass. (Insecte Homoptère) et sa plante hôte *Malus* x *domestica* Borkh.: étude de la résistance du cultivar Florina. PhD thesis, Faculté des Sciences, Université de Tours, France (114 p).

Rat-Morris E, Lespinasse Y (1995). Pommier: la résistance au puceron cendré associée à la résistance à la tavelure. Une nouvelle source de résistance mise en évidence chez le cultivar Florina. Phytoma-La Défense des végétaux, 471:15-17.

Rossi M, Goggin FL, Milligan SB, Kaloshian I, Ullman DE, Williamson VM (1998). The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. Proc Natl Acad Sci, 95 :9750-9754.

Serrano M, Guzmán P. (2004). Isolation and gene expression analysis of Arabidopsis thaliana mutants with constitutive expression of ATL2, an early elicitor-response RING-H2 zinc-finger gene. Genet, 167:919-929.

Solovyev V, Kosarev P, Seledsov I, Vorobyev D (2006) Automatic annotation of eukaryotic genes, pseudogenes and promoters. Genome Biol 7, Suppl 1: P. 10.1-10.12.

Tabone T, Mather DE, Hayden MJ (2009). Temperature switch PCR (TSP): Robust assay design for reliable amplification and genotyping of SNPs. BMC Genomics, 10:580.

Untergasser A (2006) Maxiprep - Alkaline Lysis Untergasser's Lab. Summer 2006. http://www.untergasser.de/lab/protocols/maxiprep_alkaline_lysis_v1_0.htm

Vaddepalli P, Herrmann A, Fulton L, Oelschner M, Hillmer S, Stratil TF, Fastner A, Hammes UZ, Robinson DG, Schneitz K (2014). The C2-domain protein QUIRKY and the receptor-like kinase STRUBBELIG localize to plasmodesmata and mediate tissue morphogenesis in *Arabidopsis thaliana*. Dev 141:4139-4148.

Van der Westhuizen AJ, Qian XM, Botha AM (1998) Differential induction of apoplastic peroxidase and chitinase activities in susceptible and resistant wheat cultivars by Russian wheat aphid infestation. Plant Cell Rep 18:132-137.

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A et al. (2010). The genome of the domesticated apple (*Malus* [times] *domestica* Borkh.). Nat Genet 42:833-839.

Vinatzer BA, Zhang HB, Sansavini S (1998) Construction and characterization of a bacterial artificial chromosome library of apple. Theor Appl Genet, 97:1183-1190.

Young ND, Tanksley SD (1989). Restriction fragment length polymorphism maps and the concept of graphical genotypes. TAG Theor Appl Genet, 77:95-101.

Zhong Y, Yin H, Sargent DJ, Malnoy M, Cheng ZM (2015) Species specific duplications driving the recent expansion of NBS-LRR genes in five Rosaceae species. BMC Genomics 16:77. doi:10.1186/s12864-015-1291-0

Chapter 4 –Identification of the Quirky gene in Florina and its role in RAA defense

4.1 ABSTRACT

To date the control of aphid populations is based on the use of pesticides, but the development of new resistant cultivars by breeding is considered a good approach to reduce chemical applications. Florina is the best well studied cultivar for its resistance to rosy apple aphid (RAA) *Dysaphis plantaginea* (Passerini), characterized by tolerance, antibiosis and antixenosis. Recently a genomic region of about 279 Kb from the Florina BAC library, encompassing the *Dp-fl* locus conferring resistance against rosy apple aphid, was successfully sequenced. Through the development of new polymorphic markers the *Dp-fl* locus was narrowed to a region of 97 Kb. The complete annotation of this sequence resulted in the identification of four candidate genes putatively involved in the rosy apple aphid resistance. The principal aim of this work is to identify a candidate gene conferring resistance against the *D. plantaginea* in the cultivar Florina. New markers were developed and a single candidate gene was identified in the *Dp-fl* region. The identified candidate gene is predicted to code for a protein belonging to the C2 calcium/lipid-binding phosphoribosyltransferase family, annotated in *Malus domestica* genomes (Velasco et al. 2010; Daccord et al. 2017), but also similar to the Quirky gene of *Arabidopsis*. In order to confirm the role of the Quirky gene in *D. plantaginea* resistance, various approaches were used. The Quirky gene was fully sequenced and a gene expression study was performed on Florina and Golden Delicious leaves after aphid infestation. Finally, to validate the possible involvement of this gene in aphid resistance, a construct for *Agrobacterium*-mediated transformation was prepared and a first test of genetic transformation is reported.

4.2 INTRODUCTION

Dysaphis plantaginea (Passerini) or rosy apple aphid (RAA) is a dioecious species whose primary host until mid-summer is apple and its secondary host are herbaceous plants of the *Plantago* genus (Blommers et al. 2004). Geographically, the rosy apple aphid is spread all over Asia, North Africa, North America and Europe including the whole Italian territory (Barbagallo et al. 1996). During their cycle, the apterous virginoparae settle positionate on the adaxial side of leaves, causing severe damages, such as petal fall, abscission and deformation of growing shoots, but the most relevant damage is the deformation of the developing fruits that loss their economic value (Faccioli et al. 1985; Pasqualini et al. 1996). Because of the significant economic losses that it infers, RAA is considered one of the major insect pest of European apple orchards. The damages are due to the salivary secretion released while probing intercellularly during the food-plant selection process and while feeding in the

phloem. The saliva contains peroxidases, β-glucosidase and other potential signal-generation enzymes (Miles 1999). Signals arising from the phloem feeding are able to alter the expression of inducible plant physiological factors similar to those involved in defense against pathogens (Van Der Westhuizen et al. 1998; Fidantsef et al. 1999). Two RAA resistances have been described in *Malus* spp.; *smh* in *Malus robusta* and *Dp-fl* in Florina (Alston and Briggs, 1970; Rat-Morris and Lespinasse, 1995).

Florina is the best well studied cultivar for its resistance to RAA, characterized by tolerance, antibiosis and antixenosis. When attacked by *D. plantaginea* this cultivar do not show the typical leaf and shoot deformations. After feeding on Florina plants RAA has also been shown to be less fecund and with a high mortality (Rat-Morris 1994). On Florina, RAA has been observed moving from the leaves to the stems, suggesting repellent compounds released by the leaves or a more difficult stylet penetration (Angeli and Simoni 2006). Consistently, electrical penetration graphs demonstrated that RAA stays on Florina leaves without stylet penetration for a longer period before the first probe and with a reduced duration of sap ingestion, thus indicating mechanical, biochemical or vascular resistance (Marchetti et al. 2009).

Apple resistance to *D. plantaginea* is monogenic and inherited as a dominant trait. Plant populations were obtained from crosses of Florina and aphid-susceptible parents to map and link the RAA resistance gene (*Dp-fl* locus). This result was successfully obtained by using different types of molecular markers and developing chromosome maps of the resistance locus (Pagliarani et al. 2016). Results obtained from Pagliarani work were a starting point for the identification of the gene conferring resistance against RAA in Florina. Firstly, the locus has been mapped in a specific region of about 95 Kb on the chromosome 8 of Florina and specific molecular markers linked to this trait have been developed. Four genes putatively involved in the RAA resistance have been identified in this region. One of these was the Quirky gene, interesting for its location at the plasmodesmata level that could affect phloem sap movement and its availability for the RAA. Moreover, its involvement in organ development and shaping could explain the typical leaf, stem and fruit deformations after RAA feeding (See Chapter 3).

Up to now only two aphid resistance genes have been identified, isolated and cloned in two different species: the tomato *Mi-1.2* gene and the melon *Vat* gene (Rossi et al. 1998; Dogimont et al. 2008). Both genes are constitutively expressed at low levels in the plant and encode for a protein located in the cytoplasm that belong to the NBS-LRR family, to which belongs the majority of plant resistance genes isolated so far (Dangl and Jones 2001). Aphid resistance encoded by such genes seems to be common in plants; genetic analysis of other plant-aphid interactions has shown tight linkage between resistance loci and NBS-LRR gene sequences (Brotman et al. 2002; Cevik and King 2002; Klingler et al. 2005; Seah et al. 1998). However, the mechanisms by which resistance genes recognize and defend the plant against aphids remain unknown.

There has been a growing interest in studying the defense-signaling pathways that operate during plant-aphid interactions and much of the progress to date has come from susceptible interactions. Analysis of gene expression profiling upon aphid infestation has shown upregulation by both salicylic acid (SA) and jasmonic acid (JA) responsive genes (Gao et al. 2007). These compounds have been shown to mediate resistance to a number of chewing insects (Halitscheke and Baldwin 2003; Kessler et al. 2004; Li et al. 2004; McConn et al. 1997; Royo et al. 1999) and also to cell content-feeding mites and thrips in tomato (Li et al. 2002). It remains unknown whether naturally derived aphid resistance genes operate by a similar pathway. Other pathogens related proteins were also differently expressed between resistant (R) and susceptible (S) cultivars in response to aphids. For example, in wheat the activities of β-1,3-glucanase, peroxidase and chitinase were induced to higher levels in resistant than in susceptible cultivars (Van Der Westhuizen et al. 1998). Russian wheat aphid infestation also induced SA accumulation and peroxidase activity in R compared with S wheat (Mohase and Van Der Westhuizen, 2002).

In this work is reported the comparison of the sequences of Quirky gene in different alleles and the changes in its expression after the infestation with *Dysaphis plantaginea* in greenhouse conditions. The Quirky gene was cloned in a vector for *Agrobacterium*-mediated transformation in order to validate its involvement in the RAA resistance. First steps of genetic transformation of apple are here reported.

4.3 MATERIALS AND METHODS

4.3.1 *Dp-fl* fine mapping of the Quirky gene in Florina

In Chapter 3, through the development of new polymorphic markers, the *Dp-fl* region was narrowed down to 95 Kb, flanked by two SSRs (SSR_377 and SSR_4). From the available sequences of both Florina and GDDH13 (Daccord et al. 2017), two new SSR markers within the *Dp-fl* region were developed (SSR_T and SSR_228) for the shrinkage of the *Dp-fl* window (Additional Material 4.1). The two SSR markers (SSR_T and SSR_228) were identified using the 'Tandem Repeats Finder' tool (https://tandem.bu.edu/trf/trf.html) setting the maximum period size as 3 and the minimum copy number at 12. All the primers developed were designed with Primer3 (http://primer3.ut.ee, version 4.0.0,) using default parameters. Primers are shown in the Additional Material 4.1. Finally, the fine-mapping process was completed by direct sequencing the two candidate genes identified inside the *Dp-fl* region (Quirky and tRNA genes) in each recombinant plant.

4.3.2 QKY sequencing

The two Quirky allele sequences from 'resistant' and 'susceptible' chromosomes of Florina were necessary. The complete sequence of the resistant locus of Florina was already available from the contig assembled in Chapter 3 but no sequences were available from the susceptible allele of Florina. For this purpose a set of well-distributed primer pairs have been designed on the available sequence to obtain a good overlap between flanking sequences (Additional Material 4.1). In Figure 4.1 are showed the different fragments amplified with the designed couples of primes on the Quirky sequence. In order to avoid possible unspecific amplifications, the BAC clone 83P17 from the susceptible allele of Florina was used as template for amplification and sequencing. Amplifications were performed in a 17.5 µl of volume containing 1.5 µl of diluted plasmid, 100 nM of primers [25 µM each], 1.5 µM of $MgCl_2$, 100 µM dNTPs, 0.5 Unit AmpliTaq Gold® DNA Polymerase (Applied Biosystems) and 1X reaction buffer. The reaction included an initial 10 min denaturation at 95°C, followed by 33 PCR cycles (45s at 58°C, 2 min at 72°C and 30s at 95°C), with a final extension of 7 min at 72°C. The amplicons were sequenced by using both forward and reverse primers by an external service (BIOFAB Research, Rome, Italy). The analysis of the sequences was performed with Codon Code Aligner software (version 7.0.1). For sequence comparison the Quirky gene sequence of the GDDH13 was also considered.

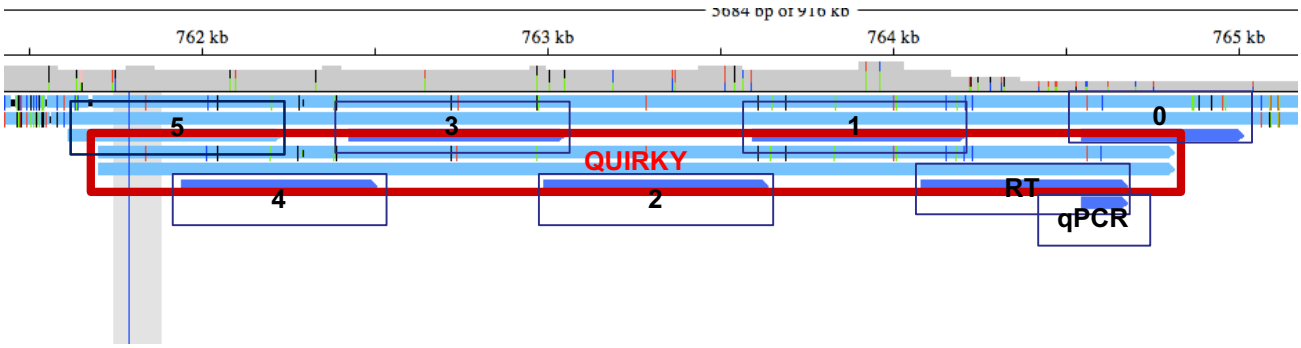

Figure 4.1 Codon Code Aligner software image show the alignment of the Quirky sequences from the resistant allele of Florina, highlighted in the red box, and the different fragments from the susceptible allele, highlighted in blue boxes. One of these sequences is named RT which is the sequence that has been utilized for the Real Time PCR. The box containing the qPCR fragment comprehend only a part of the sequence used for the Real Time PCR.

4.3.3 Resistance test assessment and qPCR analysis of Quirky gene

For the evaluation of the expression of Quirky gene, two infestations were performed on Florina and Golden Delicious plants, respectively resistant and susceptible to RAA. The two tests were performed under controlled conditions in a greenhouse at INRA, Angers, France, in two different years (2016 and 2017), as follows. Plants of each genotype were grown in pots at a temperature of about 22°C. Plants were artificially infested by RAA in mid-June when the young shoots were about 20-30 cm. Infestations were performed placing young adult apterous virginiparous females on the youngest and well-expanded leaf of each plant with a paint brush and closing the single infested leaf inside a cage (Figure 4.2). All aphids were derived from a clonal aphid line of RAA obtained from one founder collected in the field and reared on seedlings of Golden Delicious. Mock controls were also made by positioning only the cages on a leaf without aphids. The tests were carried on for 72 hours from the infestation.

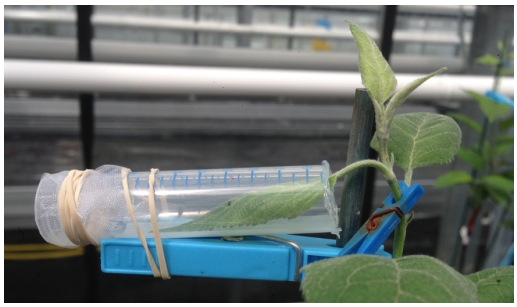

Figure 4.2 Picture of the cage used for the aphid infestation. Each cage was fixed to a support and used for a single leaf infestation. Each infested leaf was signed with a small clip. The leaf was closed into the cage with parafilm on the petiole side and with a fine net on the opposite side.

In spring 2016, the test was performed on 16 replicates of each genotype infested with five aphids. The leaf sampling was done before the infestation (T0) and after 72 hours (T72) for both infested plants (TR) and mock plants (NT). At the T0 sampling only two biological replicates (BR) were available; while, after T72, three BR were collected. Each BR is a pool of two leaves collected from two different plants.

In spring 2017, the test was performed on 40 replicates for each genotype infested with ten aphids. The leaf sampling was done at T0, T48 and T72 for both TR plants and NT plants. At each time point, four BR were collected, each one as a pool of two leaves collected from two different plants.

The sampling was performed with liquid nitrogen and RNA was extracted with NucleoSpin® RNA plant extraction kit (Macherey-Nagel GmbH & Co KG, Germany). Quantification of RNA was performed with Nanodrop and diluted for obtaining about 1,5 ng of RNA for each sample. Then, starting with the

same quantity of material was perfomed retrotrascription with GoScript™ Reverse Transcriptase kit (Promega Corporation).

The gene expression was analyzed by qPCR after the designing of a pair of primers in a conserved region of the Quirky gene (Additional Material 4.1). Actin gene, forward primer 5'-ctatgttccctggtattgcagacc-3', reverse primer 5'- acaatctgcctccaccaaacta-3' (Iorio et al. 2012) was utilized as external reference for the analysis of expression in 2016. In the analysis of 2017 also the UBiquitin-Cojugating enzyme was added as reference (forward primer 5'- cgaatttgtccgaaggcgt-3', reverse primer 5'- caatgattgtcacagcagcca-3'; Pagliarani et al. 2013). Firstly, qPCR primers were tested on 1:10 diluted cDNA samples by end-point PCR. End-point PCR amplifications were performed in a 17.5 µl of volume containing 1.5 µl of diluted cDNA, 100 nM of primers, 1.5 µM of $MgCl_2$, 100 µM dNTPs, 0.5 Unit AmpliTaq Gold® DNA Polymerase (Applied Biosystems) and 1X reaction buffer. The reaction included an initial 10 min denaturation at 95°C, followed by 33 PCR cycles (45s at 58°C, 2 min at 72°C and 30s at 95°C), with a final extension of 7 min at 72°C. The amplicons were visualized on a Kodak Image station 440 CF after electrophoresis on a 1.5% (w/v) agarose gel.

qPCR reactions were prepared in a final volume of 10 µl containing 5 µl of SYBR® Green Master Mix (Applied Biosystem), 80 nM of primers, 1.7 of RNase-free water and 2.5 µl of diluted 1:10 cDNA. Reactions were performed with a Mx3000P real-time PCR system (Stratagene) with the following program: 50°C for 2 min and 95°C for 10 min followed by 40 cycles at 95°C for 15s and 58°C for 1 min. To ensure the absence of aspecific PCR products and primer dimers, a heat dissociation protocol (from 60°C to 95°C) was also performed and a dissociation curve for each sample was generated. Each expression value was determined from the mean of three technical replicates. Amplification efficiency was calculated from raw data using LinRegPCR software (Ramakers et al. 2003). The mean normalized expression (MNE)-value was calculated foe each sample referred to the housekeeping expression, in case of two housekeeping genes (2017) it was calculated the geometrical average of the MNE. Standard error (SE)-values were calculated among the biological replicates.

In the first analysis of qPCR (2016) to deeply investigate the role of the gene related to the resistant/susceptible allele it was performed also a qPCR using primers specific for resistant and susceptible allele of Florina. The primer specificity was performed firstly by using 1 ng of extracted plasmid BAC clones in the reaction mix. BAC clones coming from the resistant (88H21) and the susceptible (83P17) chromosomes of Florina were used.

4.3.4 Construct preparation and plant transformation

To validate the role of Quirky gene and confirm its possible involvement in the response to RAA, a construct to over-express the Quirky gene was prepared as follows. First step was to isolate the Quirky gene by PCR, using the BAC clone covering the resistant allele of Florina (88H21) as template. A Herculase II Fusion DNA polymerase (Agilent Technologies) was utilized for a high-fidelity amplification of the gene. Forward and reverse primers were designed specifically for cloning the Quirky gene by adding a 5'-CACC-3' end on the forward primer before the starting codon ATG (Additional Material 4.1). A total volume of 75 μl was prepared for the reaction: 45.75 μl of PCR water, 15.0 μl of 5X Herculase II reaction buffer, 0.75 μl of oligonucleotides [25 μM each], 4.5 μl of BAC 88H21 plasmid [80 ng/μl], 4.0 μl of forward and reverse primers [10 μM], 2.0 μl of Herculase II Fusion DNA polymerase, 2.25 μl of DMSO kit solution and 0.75 μl of bovine serum albumin (BSA) solution [10%]. Amplification was performed setting the following program: 95°C for 1 min and 30 s of denaturation followed by 30 cycles (95 °C for 10s, 55 °C for 30s and 72 °C for 1 min and 30s) and 72 °C for 5 min. The amplicons were visualized on a Kodak Image station 440 CF after electrophoresis on a 1.5% (*w/v*) agarose gel. Quantification with Nanodrop® was performed to evaluate quality of the amplification. Second step for the cloning of the gene was performed following the steps described in the pENTR directional TOPO Cloning kit (ThermoFisher Scientific) using the One Shot® TOP10 Chemically competent *E. coli* cells (ThermoFisher Scientific). After the selction of the bacterial colonies, a plasmid mini-prep alkaline lysis extraction protocol (Untergasser et al. 2006) was performed. To verify the insertion of the gene in the cloning vector, a colony PCR reaction was performed with both M13 forward and reverse primers, provided in the kit, and using Quirky forward and reverse cloning primers. In order to confirm the complete gene insertion in the donor vector, one of the colonies resulted positive was then sequenced by an external service (BIOFAB Research, Rome, Italy) using the same procedure used for the complete sequencing of the gene described before. Final step was the LR reaction that was performed accordingly with the reaction protocol Gateway® LR Clonase II Enzyme mix (ThermoFisher Scientific) and using as final vector the pK7WG2D vector from the Gent University (https://gateway.psb.ugent.be/; Figure 4.3). Positive colonies were than extracted with mini-prep extraction protocol (Untergasser et al. 2006) and the insertion of the Quirky gene was again performed by PCR amplification using a p35S forward primer 5'-cttcgtcaacatggtggagcacgaca-3' and QKY5 reverse primer (Additional Material 4.1). The selected positive colony containing the final construct was then sequenced again with the same described protocol to confirm the complete insertion of the Quirky gene in the final vector pK7WG2D.

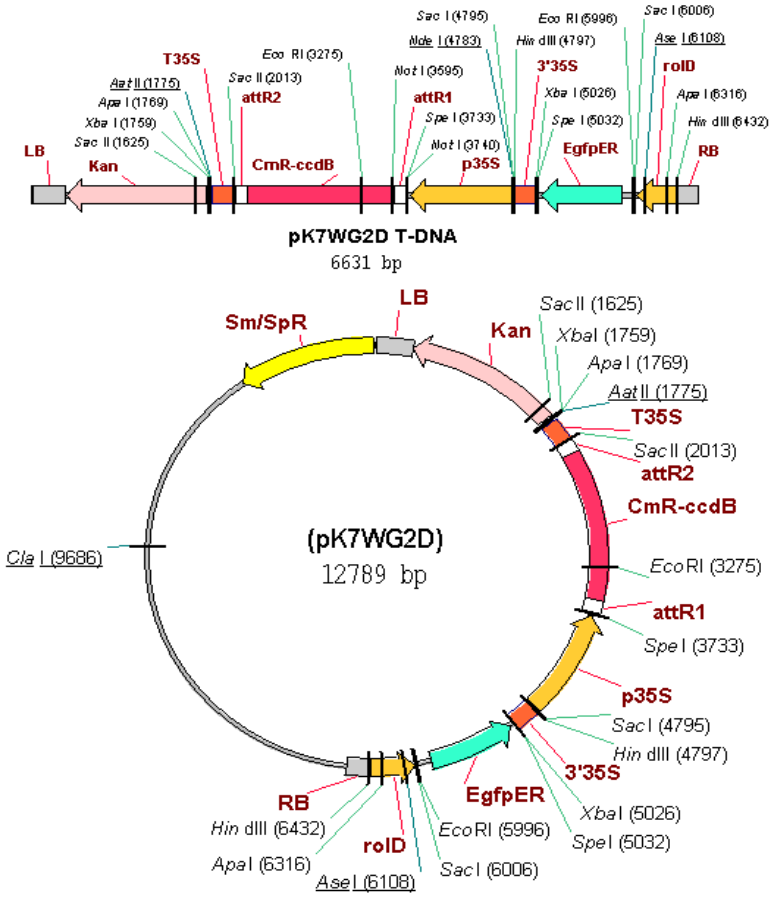

Figure 4.3 Details of the clone pK7WG2D (from Gent University) utilized for the transformation protocol. As showed in the picture the vector contains a 35S promoter for the gene of interest and contains also two reference genes: the GUS and the GFP protein.

Plasmidic DNA from the final vector was then inserted by electroporation in *Agrobacterium tumefaciens* strain EHA-105 with pBBR1MCS vector (Kovach et al. 1995); a volume of 200 µl of fresh *A. tumefaciens* EHA-105 colture and about 600 ng of extracted plasmid were then diluted in 2 ml of LB medium and left in slow agitation at 28°C for 1 hour. Then, aliquots of 50 µl and 500 µl were plated into Petri dishes containing LB with antibiotics: Gentamycin for the *Agrobacterium* selection and Spectinomycin for the plasmid selection. The Petri dishes were left in incubation at 28 °C for 2 days. Only a colony that grew on selective media was used for the following genetic transformation.

In vitro plantlets of Gala and Florina were used for the transformation because of their different interaction with RAA, respectively susceptible and resistant. Two independent experiments were performed. Gala plants were firstly maintained for 15 days in cold (4°C), then about 1 month before the transformation were transferred in standard growing chamber conditions. Florina was multiplied less than one month before the transformation and shoots grew only in the growing chamber. Media preparation for plant regeneration and micropropagation are explained in the Additional Material 4.5.

For transformation process, the day before the infiltration the leaves were picked up and prepared on petri dishes for a total of 200 leaves from Gala and 175 leaves from Florina. The leaves remained one day in dark for preparation. The inoculum was prepared using LB medium added with Spectinomycin and Gentamycin, the inoculum was firstly prepared on solid medium and then transferred into liquid media. The suspension was prepared for a final concentration of $1x10^8$ cells in liquid LB media added with Acetosyringon 150 μM. The inoculum was left for 4 hours at RT to activate the virulence of the bacteria. Leaves were prepared by inferring wounds on the main veins of the adaxial side. A part of the Gala leaves were not wounded to estimate the regeneration efficiency. About 30 leaves of Gala were prepared for a control transformation with the same *Agrobacterium* tumefaciens strain, with a pKGWFS7,0 vector from Gent University containing a 35S promoter for GFP and a GUS reporter. The leaves were agroinfiltrated for 1 min and then left in co-colture in liquid medium added with 100 μM of Acetosyringone for 2 days in the darkness. At the end of the co-colture the leaves were transferred to a new media with Cefotaxime 450 mg/l and Kanamycin 150 mg/l. In table 4.1 there are the effective leaves that were utilized in the experiment and the treatment of each one.

Table 4.1 Number of leaves used in the transformation experiment divided for genotype (Florina or Gala). For each transformation is detailed the vector used in the infiltration or is written no for the control. In the last column is detailed that for all the transformed plants there was a co-colture period after the inoculum.

Genotype	Infiltration	N. of leaves	Co-culture
Gala	QUIRKY	140	yes
Florina	QUIRKY	175	yes
Gala	pKGWFS7,0	30	yes
Gala	no	30	no

4.4 RESULTS

4.4.1 Quirky gene as candidate gene for RAA resistance in Florina

The *Dp-fl* locus was firstly mapped within the two SSR markers SSR_377 and SSR_4, bracketing a region of only about 95 Kb. Thanks to the mapping of two new SSRs developed within the region (SSR_T and SSR_228), it was possible to further restrict the region to about 56 Kb. In graphical genotype are showed genotypic and phenotypic data of the recombinants plants (Figure 4.4). By the addition of the two last markers, a further individual with a genotype-phenotype incongruence (GPI) emerged (PF_P001). This individual, with resistant phenotype, showed an incongruent 'susceptible' genotype for marker SSR_228. Anyway, thanks to the individual RF_X-9104-8 with 'ab' genotype and the susceptible phenotype at the SSR_228 position made possible to delimit the *Dp-fl* locus before the SSR_228. On the opposite site, the individual FR_154 allowed us to define the SSR_T as the left border of the *Dp-fl* region with 'ab' genotype at the SSR_T locus and resulting susceptible for the phenotype (Figure 4.4).

RECOMBINANT PLANTS	SNP_104	SSR Df_C	SSR DF-F	SCAR_1	SSR 377	SSR_T	QUIRKY	*Dp-fl*	tRNA	SSR_228	SSR_4	SNP_585
FR_154	ab	ab	ab	ab	ab	ab	aa	**S**	aa	aa	aa	aa
FP_21	ab	ab	ab	ab	aa	aa	aa	**S**	aa	aa	aa	aa
GD_4	ab	ab	ab	aa	aa	-	aa	**S**	aa	aa	aa	aa
MF_7321	ab	ab	aa	aa	aa	-	aa	**S**	aa	aa	aa	aa
MF_F145	ab	ab	aa	aa	aa	aa	aa	**S**	aa	aa	aa	aa
RF_X-9104-8	aa	aa	aa	aa	aa	aa	aa	**S**	ab	ab	ab	ab
PF_P001	aa	aa	aa	aa	aa	aa	aa	**R**	aa	aa	ab	ab
PF_X-9504-33	aa	aa	aa	aa	aa	aa	aa	**R**	aa	aa	aa	aa

Figure 4.4 Graphical genotyping of the available recombinant individuals for the all the markers developed in the *Dp-fl* region. The allele linked to susceptibility is 'aa' (red box) or 'ab' (green box) when linked to resistance. The *Dp-fl* locus column show the results of the phenotypic evaluations: 'S' in a red box indicate RAA susceptibility while 'R' in a yellow box highlight the two GPI individuals resistant to RAA. SSR_T is an 'abxab' marker in some progenies, therefore the genotype of some recombinants cannot be deduced (white box).

Using the predictions performed with Fgenesh software, reported in Chapter 3, it was possible to focus the attention in the elements included in the 56 Kb region. In this region only a transposable element (TE) and three genes were predicted (Figure 4.5).

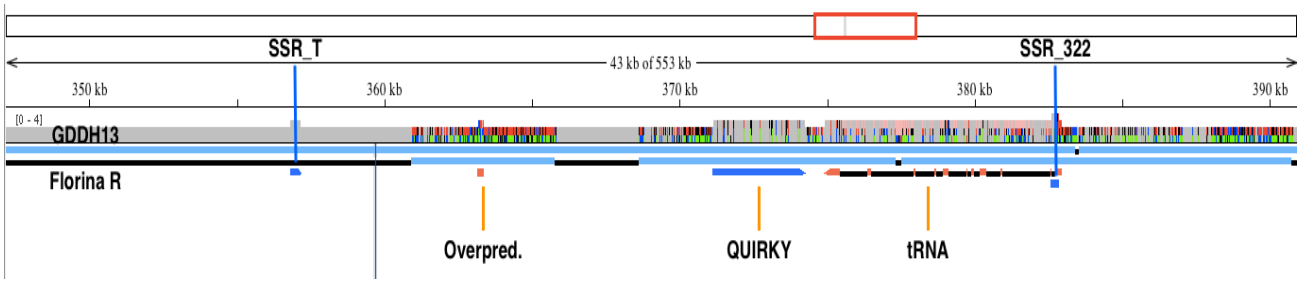

Figure 4.5 Schematic representation of the alignment of the GDDH13 sequence (upper part) and the sequence from the resistant allele of Florina (lower part) by Codoncode Aligner. Markers are shown with blue arrows and genes with orange arrows.

With Codoncode aligner software, it was possible to align the assembled *Dp-fl* contig sequences of Florina (lower part) and the homologous sequence of GDDH13 (upper part). This comparison made it possible to identify many regions not shared by the two sequences showed by the black lines in Figure 4.5. In detail, Florina lack two big parts of the sequence of GDDH13, and one of these includes the SSR_T. The SSR_322 is included in the third candidate gene. Among the three predicted genes one had a very short sequence (162 bp) and putatively was a overpredicted gene by the software Fgenesh. In addition, it was not possible to predict a function for this gene and for this reason it was not further considered. Another candidate gene was predicted to encode for a protein similar to the Quirky gene (see Chapter 3). Quirky is about 3108 bp, do not include introns and is coding for a protein of 1035 aa. The last candidate gene is predicted to encode for a Threonylcarbamoyladenosine tRNA methylthiotrasferase, simply called in this work tRNA. The length of this gene is 1896 bp, containing many introns and encoding for a protein of 631 aa. tRNAs are central adaptors in the translation process responsible for decoding mRNAs; tRNAs harbor numerous post-trascriptional modifications that are reported to fine-tune their function.

Through the sequencing of the Quirky and tRNA candidate genes it was possible to identify the differences between the susceptible and resistant alleles in order to fine map the *Dp-fl* locus. The sequences from the two BAC clones covering the resistant and the susceptible chromosome of Florina (88H21 and 83P17) were compared with the corresponding sequences of the recombinant plants. SNPs differences in the recombinants are showed in the graphical genotyping represented in Figure 4.4. From the comparison of the sequences of the Quirky gene, three different SNPs were identified but analyzing the sequences of the recombinant plants all of them were presenting the 'aa' genotype at the Quirky position (Figure 4.4). Then, considering the tRNA gene sequences two SNPs differences were highlighted. From the analysis of the SNPs found on the tRNA gene sequences from the recombinants,

the individual RF_X-9104-8 was found to be 'ab' for the tRNA gene (Figure 4.4). This data made possible to exclude this gene from the list of the candidate genes. Finally, the phenotypic data were completely in accordance with the genotypes at the Quirky locus, suggesting this gene as the only candidate gene in the *Dp-fl* region. Noteworthy, genotypes of the recombinant plants used for the fine mapping are all with a susceptible phenotype since the resistant phenotype resulted as GPI.

4.4.2 Quirky sequencing and sequences comparison

To prove the possible involvement of the Quirky candidate gene in RAA resistance of Florina, we compared the sequences of the resistant and susceptible alleles. The sequence on the resistant allele was already available from the assembled BAC clones contig, the sequence of the double haploid Golden Delicious (GDDH13) was available from the web, and finally, the complete sequence of the susceptible genotype was obtained by direct sequencing of the whole Quirky gene from the 83P17 clone covering the susceptible chromosome of Florina. The complete alignment was performed with Clustal Omega software (v. 1.2.4, https://www.ebi.ac.uk/Tools/msa/clustalo) and is available in the Additional Material 4.2. Between the Golden Delicious and the Florina susceptible allele only one SNP was detected. In particular, this SNP was not significant because it doesn't change the translation, indeed the resulting protein is identical . Instead, looking at the differences between the 'susceptible' and the 'resistant' allele of Florina, a total of 23 SNPs and a 3 bp indel were identified (Additional Material 4.2). Looking then at the differences at the protein level, a total of 9 amino acids changes were identified: mostly are conservative replacements (no. 8) and only one is determining a change between two amino acids with different properties. For instance, the Glutamic Acid (acidic) at position 347 of the 'resistant' allele is substituted by a Valine (aliphatic) in the 'susceptible' allele. In addition, the Florina 'resistant' allele lack of a single amino acid (Lys-197) because of the presence of a three bp indel (Additional Material 4.3).

The three promoter sequences of the Quirky gene (GDDH13, 'resistant' and 'susceptible' alleles of Florina) were also compared by alignment with Clustal O software (v. 1.2.4) (Additional Material 4.4). The main difference in promoter regions is the lack of the first 115 bp before the starting codon in both the alleles of Florina compared to the GDDH13 sequence. This big difference in the promoter sequence, that occurs next to the start codon, can affect the expression of the Quirky gene of Florina and Golden Delicious. Nevertheless, the two alleles of Florina are both lacking this sequence before the ATG. Possible differential expression pattern between the two Quirky alleles in Florina has to be ascribed to the presence of other differences in the promoter sequence (highlighted in green in Additional Material 4.4). Further analysis will be necessary to link the differences in the promoter region and the expression pattern of the gene.

4.4.3 qPCR analysis of Quirky gene

Results obtained in 2016 are shown in Figure 4.6, the expression of the Quirky (QKY) gene is low and quite similar in the two genotypes (Golden Delicious and Florina) and at the two times after the infestation (T0 and T72).

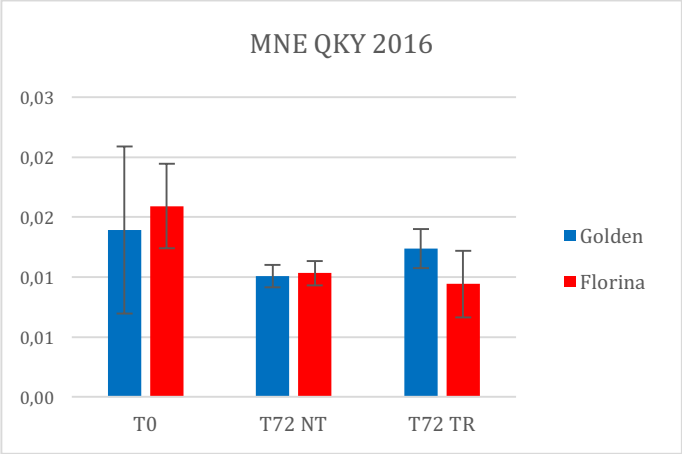

Figure 4.6 Mean normalized expression (MNE) of Quirky (QKY) gene of tests performed in 2016 for Florina and Golden genotypes. T0- Before infestation; T72 NT- Non-infested plants after 72 hours from infestation; T72 TR – Infested plants after 72 hours from the infestation. Bars represent the standard deviation between the BR.

By using the primer specific for the two Quirky gene sequences from the two alleles of Florina, the 'resistant' allele was not expressed at all in the Golden Delicious genotype, as expected. Regarding the Florina samples, the expression of the 'resistant' allele (QKY-R) is more expressed at T0 and in the mock control at T72, while the plants after 72 hours from infestation with aphids were not showing any expression of the Quirky 'resistant' allele (Figure 4.7).

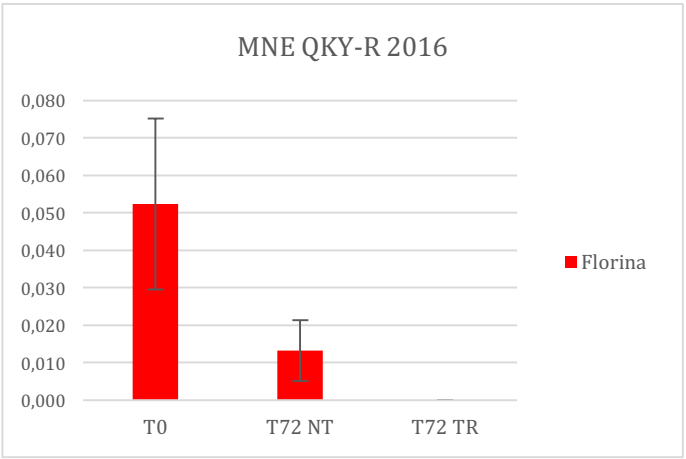

Figure 4.7 Mean normalized expression (MNE) of Quirky (QKY) gene with primers specific for the resistant allele (R); tests performed in 2016. Are showed only results obtained with Florina genotypes because Golden resulted with no expression, as expected. T0- Before infestation; T72 NT- Non-infested plants after 72 hours from the infestation; T72 TR – Infested plants after 72 hours from the infestation. Bars represent the standard deviation between the BR.

The expression of QKY gene of the 'susceptible' allele (QKY-S) of Florina and Golden Delicious is showed in Figure 4.8. The QKY-S allele is almost not expressed at T0 and T72 in the not-treated plants, but it's expression increase at 72 hours after the RAA infestation but only in Golden Delicious while in Florina the expression is not different from the mock control at T72.

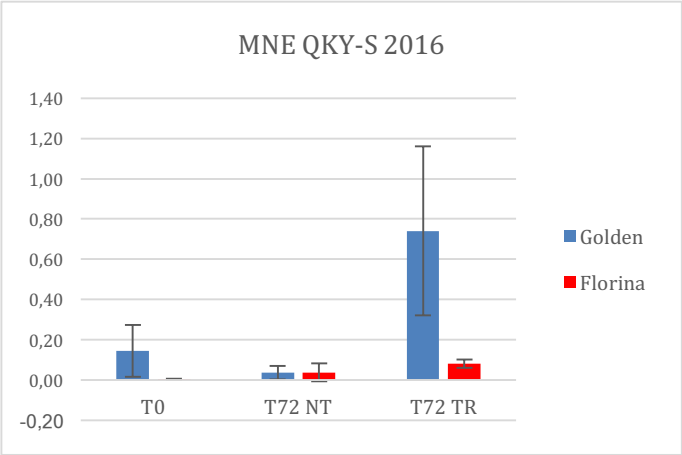

Figure 4.8 Mean normalized expression (MNE) of Quirky gene with primers specific for the susceptible allele (S); test performed in 2016 on Florina and Golden genotypes. T0- Before infestation; T72 NT- Non-infested plants after 72 hours from infestation; T72 TR – Infested plants after 72 hours from infestation. Bars represent the standard deviation between the biological replicates.

The high variability of expression between biological replicates observed in 2016 suggested to repeat the RAA infestation on a larger set of plants. Therefore, in 2017 the number of biological replicates was increased to 4 and an additional sampling timing was added (T48). The results in Figure 4.9 show that the expression levels at T0, in the two mock controls (T48NT and T72NT) and in the infested samples at T48 (T48TR) are generally very low but with a slightly higher expression in Florina than in Golden. At T72 after the RAA infestation, a clear increase of the expression of QKY gene is shown in Golden Delicious compared to Florina. In particular, QKY expression is about 3 times higher in Golden Delicious than in Florina (Figure 4.9).

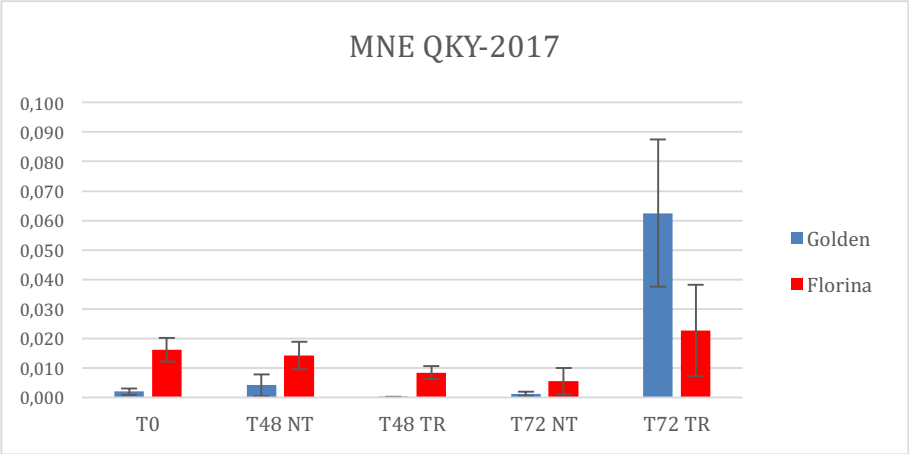

Figure 4.9 Mean normalized expression (MNE) of Quirky gene, tests performed in 2017 in Florina and Golden genotypes. Bars represent the standard deviation between the biological replicates.

In general, all the value of MNE are low, so very little differences between the biological replicates can influence the significance of the expression. In conclusion, after 72 hours from the infestation, the QKY gene seems to be more expressed in Golden Delicious than Florina.

4.4.4 In vitro transformed plants

To better understand the role of the Quirky gene, a QKY construct was prepared and used for the genetic transformation of two apple cultivars: Gala and Florina. Gala is a RAA susceptible cultivar with a good attitude to genetic transformation, while Florina is a RAA resistant cultivar is a cultivar that is not usually used for the transformation tests (Radchuk and Korkhovoy, 2005). Herewith we report only the preliminary results of the transformation experiment made in during last months of my PhD.

After one month from the transformation, in Gala, some precocious independent regenerations were observed. In particular, at least four well-growing shoots and other small shoots were obtained on Gala explants (Figure 4.10), while no regenerations were visible in Florina (Figure 4.11). In Gala explants transformed with the pKGWFS7,0 vector control construct there were also two regenerations.

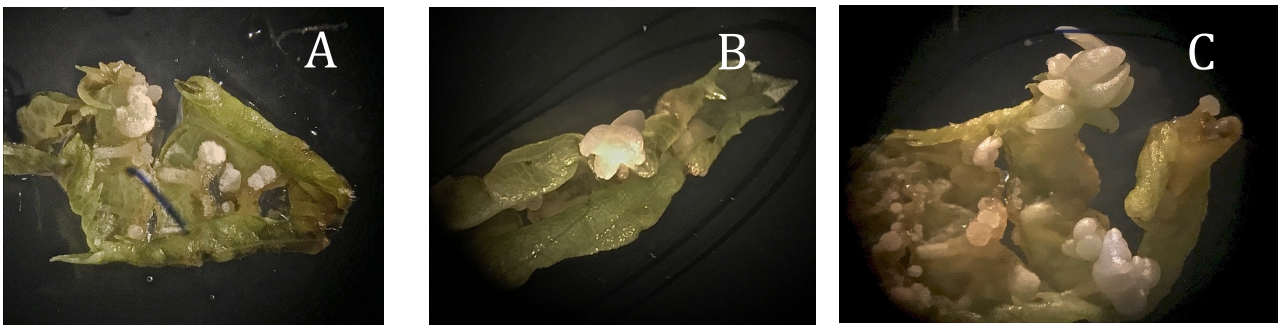

Figure 4.10 A B and C Gala leaves after 1 month from transformation with Quirky gene show some growing shoots.

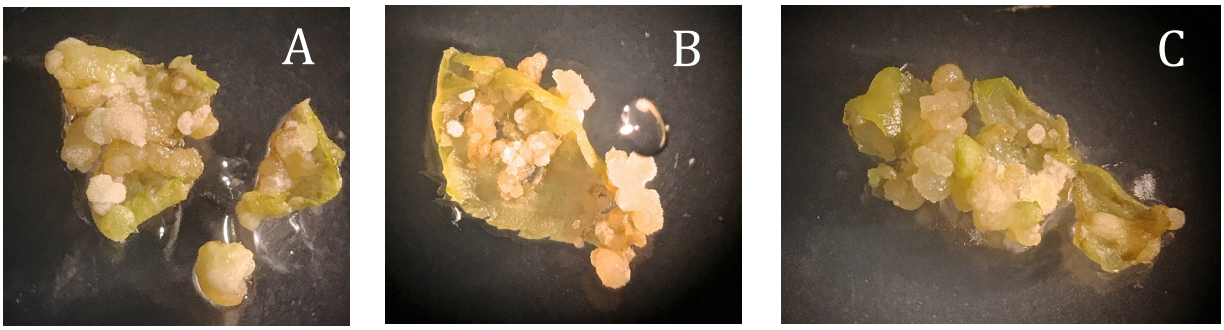

Figure 4.11 A B and C Florina leaves after 1 month from transformation with Quirky gene show an abundant formation of callus.

After five month from the transformation, Florina leaves produced only some callus, no regenerations were observed. In contrast, Gala explants continue to produce new regenerations. The best regenerated shoots were transferred to a propagation medium in order to propagate each independent regeneration (Table 4.2). In Figure 4.12 are showed the Gala plants transferred to propagation media four months after transformation with Quirky, the plants do not present particular phenotype.

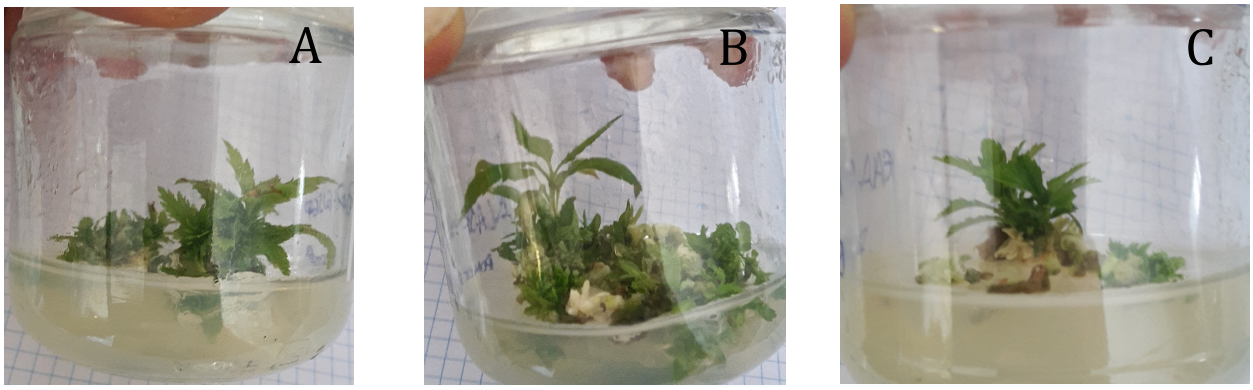

Figure 4.12 A B and C Gala plants in propagation media after 5 month from transformation with Quirky gene.

A very high percentage of putatively transformed plants was obtained in Gala, about 9% of the explants transformed with the Quirky gene showed regenerations, as well as the 7.5 % of the explants transformed with the pKGWFS7,0 control vector (Table 4.2).

Table 4.2 Percentage (%) of transformed plants after five months. In the second column is indicated with which vector has been performed the infiltraton: QUIRKY - plants infiltrated with *A. tumefaciens* strain EHA-105 with pBBR1MCS vector and pK7WG2D vector carrying the Quirky gene; pKGWFS7,0 – plants transformed with *A. tumefaciens* strain EHA-105 with pBBR1MCS vector pKGWFS7,0 control vector.

Genotype	Infiltration	N. of transformants	% transformants
Gala	QUIRKY	13	13/140= 9%
Florina	QUIRKY	0	0/175= 0%
Gala	pKGWFS7,0	4	4/30= 7.5%

To date, it was not possible to verify the effective transformation of each regenerant line because of the small size of the plantlets. The molecular characterization of the first regenerants will be performed after the completation of the propagation step of each transformed line and the subsequent rooting and acclimation.

4.5 DISCUSSION

In this work the fine mapping of the *Dp-fl* locus of Florina was accomplished. Thanks to the development of new polymorphic markers it was possible to reduce the region to a window of only 56 Kb, but a GPI individual was identified (PF_P001, Figure 4.4). GPI individuals have been already reported in the cloning of other resistance gene in apple but their elimination do not affect the marker order (Gigax et al. 2004; Erdin et al. 2006; Soriano et al. 2009). A GPI individual (PF_X-9504-33) was already identified in Chapter 3 in Perico x Florina population. This individual was discarded because of its resistant phenotype coupled with a 'susceptible' genotype on both sides of the *Dp-fl* locus. Unfortunately, both GPI plants were no more available at SERIDA (Villaviciosa, Spain) and therefore the phenotypic evaluation cannot be repeated.

Interestingly, all the recombinant plants that locate the *Dp-fl* locus present all a susceptible phenotype and in the phenotyping tests the susceptible phenotype is always more clear to identify than the resistant one. The available graphical genotypes made it possible to locate the resistance gene inside the *Dp-fl* window, but an asymmetric distribution of recombination events in the identified individuals was observed. Five recombinants on one side of the resistance gene strongly support the left limit of the *Dp-fl* window; while, on the opposite side, only one individual that recombine in this position is available. Therefore, the search of new recombinants is needed to further confirm the location of the resistance gene.

Unexpectedly, none of the genes identified in the *Dp-fl* windows is similar to one of the known classes of resistance genes identified so far (Gururani et al. 2012), including the NBS-LRR proteins that are able to confer resistance to aphids in other species (Dogimont et al. 2010). If the fine-mapping results will be confirmed by the analysis of new recombinants, a different mechanism of resistance can be hypothesized. To date, inside the 56 Kb region only one candidate gene, the Quirky gene was identified. The Quirky gene was already studied in *A. thaliana*, where it was demonstrated that is located at plasmodesmata level (Fulton et al. 2009) and this position can affect the movement of the sap through the sieve elements and its availability for RAA. Moreover, the involvement of the Quirky gene in tissue development in *A. thaliana* could explain the typical leaf, stem and fruit deformations after RAA feeding. A recent study co-locates the QKY protein at the plasmodesmata level together with an LRR-receptor-like kinase protein called STRUBBELIG (Vaddepalli et al. 2014). Therefore, a gene with an LRR-receptor like kinase structure interacting with the Quirky gene can be hypothesized, supporting a second possibility where the presence of a resistance gene collaborate with the Quirky explaining also the observed genotype-phenotype incongruences.

Sequences of the Quirky gene from both 'resistant' and 'susceptible' alleles of Florina and from GDDH13 were obtained in this work. These sequences were compared and resulted very similar in the coding regions, but quite different in promoter region, where a region of 154 bp before the start codon was present only in GDDH13. The differences among promoters could influence the expression pattern of this gene. In particular, Quirky gene is more expressed in Golden Delicious than in Florina; its expression increase at 72 hours after RAA infestation. This data can explain the difference among the promoter sequence, that can influence in the efficiency of expression. Finally, the higher expression in Golden suggest that this gene could be more linked to susceptibility than resistance. Indeed, the recombinant plants that define the *Dp-fl* windows present all a susceptible phenotype.

The gene validation by gene transformation in both Gala and Florina has been tested, but only Gala for the moment promise good results that have still to be confirmed.

The undertaken validation of gene function by genetic transformation in both Gala and Florina should shed light on the real involvement of this gene in RAA resistance. To date preliminary results indicate that Florina is more recalcitrant to regenerate than Gala, that confirmed its high regeneration attitude. Further analysis will be necessary to confirm the integration and the expression of the transgene in the two different genetic backgrounds. Because of the observed Quirky expression patterns, the overexpression of the Quirky gene in both GM-genotypes could increase their susceptibility. On the other hand, a co-suppression mechanism of the endogenous Quirky gene cannot be excluded in the Gala regenerants. The resulting Quirky gene silencing in Gala could possibly increase its resistance to RAA. To date no alterations of the shoot morphology was observed in the first regenerants.

4.6 CONCLUSION

In this work, a candidate gene responsible for the RAA response in Florina has been identified. The promoter sequence analysis and the gene expression suggest that its role could be more related to susceptibility than to resistance. The identification of GPI individuals suggests to not exclude the presence of a resistance gene that may influence the function of the Quirky gene, but further studies are needed. The identification of new recombinant plants in the region, hopefully with both the resistant and susceptible phenotype, and the results obtained from the genetic transformation with the Quirky gene should increase our knowledge on RAA resistance.

4.7 Additional material

Additional Material 4.1 List of the primers. In the first group are listed the two SSR primers used for the fine mapping of the gene. In the second group are listed the primers used for the sequencing of the whole Quirky gene including part of the promoter region. QKYRT_F and the QKYqPCR_R have been used for the end point PCR analysis For the promoter analysis two couples of primers were developed and specific primers were developed for the cloning of the gene. Are here included also the two primers used for the sequencing of the tRNA gene. In the final box are showed the primers used in the qPCR analysis, the not allele specific primer (QKYqPCR_F), the QKYqPCR(S)_F specific for the susceptible and QKYqPCR(R)_F for the resistant allele of Florina.

Analysis	Primer name	Sequence (5'-3')	Amplicon length (bp)
***Dp-fl* fine mapping**	SSR 228 F	ACCTGGTTGTGTGAGCATCC	191/231
	SSR 228 R	ACCAACGCCAAGCTATCTCA	
	SSR T F	GGAGGGGAAATAGGGAATTGTGT	355
	SSR T R	CCACGTTGGCTTGAACCGA	
Quirky sequencing	QKY1_F	CTTGGGATCCTGAGTGCCA	616
	QKY1_R	TGCGGTGGAAGTTGGCTT	
	QKY2_F	AGTCCTTTCCCGATGCTTGG	649
	QKY2_R	AGTCCTTTCCCGATGCTTGG	
	QKY3_F	CGCCCACCGTTGAAACAAG	626
	QKY3_R	GGCGAGATTGACGTGGCTAA	
	QKY4_F	TGGACACCATACGAACTTCCTC	562
	QKY4_R	AAATTCCGGGTTCTGCTGCT	
	QKY5_F	CTCCCACTCCTTCTAGCGAT	617
	QKY5_R	TTGATTTCCTGCAGCGGAGG	
	QKYqPCR_F	CAGGGTGCAGATGGTGGT	467
	QKY0_R	TGTTGTGCGCTTCTTTCGTG	
	QKYRT_F	AATGCAGATTGTAGCAACGAGG	451
	QKYqPCR_R	AGAAGAGCCCTATCGACGAAAC	
Quirky promoter sequencing	QKYProm1_F	CATGTCGACAACTTTTGCCAAAC	1402
	QKYProm1_R	CTTGCGTCGTGAACTTCCAC	
	QKYProm2_F	ACCCAAGACAATCGACTTGTC	713
	QKYProm2_R	CTTGCGTCGTGAACTTCCAC	
Quirky gene cloning	QKY_cloningF	caccATGGCCAATACTAAACTT	3112
	QKY_cloningR	TCAAGATAGTAGCATATCTGAC	
tRNA gene sequencing	tRNA_F	TTCTTGTAGCCGCTCCAGAAAG	451
	tRNA_R	sAGAAGAGCCCTATCGACGAAAC	
Expression analysis	QKYqPCR_F	CAGGGTGCAGATGGTGGT	132
	QKYqPCR(S)_F	CAGGGTGCAGATGGTGGTA	
	QKYqPCR(R)_F	CAGGGTGCAGATGGTGGTT	
	QKYqPCR_R	AGAAGAGCCCTATCGACGAAAC	

Additional Material 4.2 Clustal O alignment of the sequences of Quirky gene from the resistant chromosome (QKY_FlorinaR) from the GDDH13 genome (Golden) and from the susceptible allele of Florina (QKY_S). Highlighted in red the differences between the quirky from resistant allele of Florina and the other two susceptible alleles. Highlighted in light blue the only difference between the quirky gene from the susceptible allele of Florina and the sequence coming from the Golden Delicious double haploid genome.

```
QKY_FlorinaR      ----------------------ATGGCCAATACTAAACTTGTGGTGGAAGTTCACGACGCA
Golden            ----------------------ATGGCCAATACTAAACTTGTGGTGGAAGTTCACGACGCA
QKY_S             CTCAACTcCGCCGAGCTAGCCATGGCCAATACTAAACTTGTGGTGGAAGTTCACGACGCA
                                        ************************************

QKY_FlorinaR      AGCGACCTGATGCCGAAAGACGGCGACGGTTTTGCGAGTCCCTTCGTGGAGGTAAACTTT
Golden            AGCGACCTGATGCCGAAAGACGGCGACGGTTTTGCGAGTCCCTTCGTGGAGGTAAACTTT
QKY_S             AGCGACCTGATGCCGAAAGACGGCGACGGTTTTGCGAGTCCCTTCGTGGAGGTAAACTTT
                  ************************************************************

QKY_FlorinaR      GAAGGGGAGCGGCAGCGGACTCAGACCAAGCCAAAAGACCTCAATCCTAACTGGAACGAG
Golden            GAAGGGGAGCGGCAGCGGACTCAGACCAAGCCAAAAGATCTCAATCCTAACTGGAACGAG
QKY_S             GAAGGGGAGCGGCAGCGGACTCAGACCAAGCCAAAAGATCTCAATCCTAACTGGAACGAG
                  ************************************ ********************

QKY_FlorinaR      AAGCTCGTCTTCAACATCAACGACCGTTCTCACCTCCCCCACAAGACCGTCGACATTGTC
Golden            AAGCTCGTCTTCAACATCAACGACCGTTCTCACCTCCCCCACAAGACCGTCGACATTGTC
QKY_S             AAGCTCGTCTTCAACATCAACGACCGTTCTCACCTCCCCCACAAGACCGTCGACATTGTC
                  ************************************************************

QKY_FlorinaR      GTTTACAATGACAGACAAACTGGACACCATACGAACTTCCTCGGCCGAGTCAGAATCTCC
Golden            GTTTACAATGACAGACAAACTGGACACCATACGAACTTCCTCGGCCGAGTCAGAATCTCC
QKY_S             GTTTACAATGACAGACAAACTGGACACCATACGAACTTCCTCGGCCGAGTCAGAATCTCC
                  ************************************************************

QKY_FlorinaR      GGCGTCTCCGTCCCTTTCTCCGAGTCTCAGGCCACTATCCAACGGTACCCGCTCGATAAG
Golden            GGCGTCTCCGTCCCTTTCTCCGAGTCTCAGGCCACCATCCAACGGTACCCGCTCGATAAG
QKY_S             GGCGTCTCCGTCCCTTTCTCCGAGTCTCAGGCCACCATCCAACGGTACCCGCTCGATAAG
                  ********************************** ************************

QKY_FlorinaR      CGCGGTCTCTTCTCTCATGTCAAAGGCGATATTGCCCTCAGAATCTACGCTATTCAAGAT
Golden            CGCGGTGTCTTCTCTCATGTCAAAGGCGATATTGCCCTCAGAATCTACGCTATTCAAGAT
QKY_S             CGCGGTGTCTTCTCTCATGTCAAAGGCGATATTGCCCTCAGAATCTACGCTATTCAAGAT
                  ****** ****************************************************

QKY_FlorinaR      TACATCAACAATGGCGACTTTGCTCCAACACCAGCACCACCCCCACCTACACTAAATGAT
Golden            TACATCAACAATGGCGACTTTGCTCCAACACCAGCACCACCCCCACCTACACTAAATGAT
QKY_S             TACATCAACAATGGCGACTTTGCTCCAACACCAGCACCACCCCCACCTACACTAAATGAT
                  ************************************************************

QKY_FlorinaR      GAATTTGTTACTAATAGTACTGGTGGTGCTGCTGGGACTCCTCGTCCTCCTCCGCTGCAG
Golden            GAATTTGTTACTAATAGTACTGGTGGTGCTGCTGGGACTACTCGTCCTCCTCCGCTGCAG
QKY_S             GAATTTGTTACTAATAGTACTGGTGGTGCTGCTGGGACTACTCGTCCTCCTCCGCTGCAG
                  ************************************ ********************
```

```
QKY_FlorinaR      GAAATCAATACTAATAGGATCGTTGAGGAGATTCATCACCACCATTTTGGGGGAGAGAAA
Golden            GAAATCAATACTAATAGGATCGTTGAGGAGATTCATCACCACCATTTTGGGGGAGAGAAG
QKY_S             GAAATCAATACTAATAGGATCGTTGAGGAGATTCATCACCACCATTTTGGGGGAGAGAAG
                  ************************************************************

QKY_FlorinaR      ATCAAGA---AGAAGGAAAAAGAAGTGAGAACTTTCCACTCCATCGGCACTGGGATGGGT
Golden            ATCAAGAAGAAGAAGGAAAAAGAAGTGAGAACTTTCCACTCCATCGGCACTGGGATGGGT
QKY_S             ATCAAGAAGAAGAAGGAAAAAGAAGTGAGAACTTTCCACTCCATCGGCACTGGGATGGGT
                  *******   ************************************************

QKY_FlorinaR      GGTGGTGGTGGCGGCGGCGGTGGTTCTCATCCTCCTCCTCCTATGTCTTCCGGATTCGGA
Golden            GGTGGTGGTGGCGGCGGCGGTGGTTCTCATCCTCCTCCTCCAATGTCTTCGGGATTCGGA
QKY_S             GGTGGTGGTGGCGGCGGCGGTGGTTCTCATCCTCCTCCTCCAATGTCTTCGGGATTCGGA
                  ************************************** ******** *********

QKY_FlorinaR      TTTGAGACAATGAAGGAGAAGGCGCCCACCGTTGAAACAAGGACGGATTTCGCTCGGGCG
Golden            TTTGAGACAATGAAGGAGAAGGCGCCCACCGTTGAAACAAGGACGGATTTCGCTCGGGCG
QKY_S             TTTGAGACAATGAAGGAGAAGGCGCCCACCGTTGAAACAAGGACGGATTTCGCTCGGGCG
                  ************************************************************

QKY_FlorinaR      GGTCCTGCCACGGTTATGCACATGCAGCAGCAGAACCCGGAATTTTCCCTGGTGGAGACA
Golden            GGTCCTGCCACGGTTATGCACATGCAGCAGCAGAACCCGGAATTTTCCCTGGTGGAGACA
QKY_S             GGTCCTGCCACGGTTATGCACATGCAGCAGCAGAACCCGGAATTTTCCCTGGTGGAGACA
                  ************************************************************

QKY_FlorinaR      GATCCACCATTGGCGGCGCGGCGGTACAGAGGTTTTGGAGGGGACAAGACCTCGAGCACA
Golden            GATCCACCATTGGCGGCGCGGCGGTACAGAGGTTTTGGAGGGGACAAGACCTCGAGCACA
QKY_S             GATCCACCATTGGCGGCGCGGCGGTACAGAGGTTTTGGAGGGGACAAGACCTCGAGCACA
                  ************************************************************

QKY_FlorinaR      TACGATCTGGTTGAGCAGATGCATTACTTGTACGTGAGTGTGGTGAAGGCAAGAGATCTT
Golden            TACGATCTGGTTGAGCAGATGCATTACTTGTACGTGAGTGTGGTGAAGGCAAGAGATCTT
QKY_S             TACGATCTGGTTGAGCAGATGCATTACTTGTACGTGAGTGTGGTGAAGGCAAGAGATCTT
                  ************************************************************

QKY_FlorinaR      CCAACCATGGATGTTACAGGAAGCCTTGATCCTTATGTGGAGGTGAAGCTTGGCAACTAC
Golden            CCAACCATGGATGTTACAGGAAGCCTTGATCCTTATGTGGAGGTGAAGCTTGGCAACTAC
QKY_S             CCAACCATGGATGTTACAGGAAGCCTTGATCCTTATGTGGAGGTGAAGCTTGGCAACTAC
                  ************************************************************

QKY_FlorinaR      AAAGGGGTGACCAAGCATGTGAACAAGGACCAGAACCCTGAGTGGCACCAGATTTTCGCC
Golden            AAAGGGGTGACCAAGCATGTGGACAAGGACCAGAACCCTGTGTGGCACCAGATTTTCGCC
QKY_S             AAAGGGGTGACCAAGCATGTGGACAAGGACCAGAACCCTGTGTGGCACCAGATTTTCGCC
                  ******************** ***************** ********************

QKY_FlorinaR      TTCTCGAAAGAGCGCGTGCAATCCAATTTGCTTGAAGTCACTGTCAAGGACAAGGATTTC
Golden            TTCTCGAAAGAGCGCGTGCAATCCAATTTGCTTGAAGTCACTGTCAAGGACAAGGATTTC
QKY_S             TTCTCGAAAGAGCGCGTGCAATCCAATTTGCTTGAAGTCACTGTCAAGGACAAGGATTTC
                  ************************************************************

QKY_FlorinaR      ACCAAGGATGATATCGTGGGGAGGCTACATTTCGATCTCAGCGAAGTCCCCCTTTGCATG
Golden            ACCAAGGATGATATCGTGGGGAGGCTACATTTCGATCTCAGCGAAGTCCCCCTTTGCATG
QKY_S             ACCAAGGATGATATCGTGGGGAGGCTACATTTCGATCTCAGCGAAGTCCCCCTTTGCATG
                  ************************************************************
```

```
QKY_FlorinaR      CCGCCTGACAGCCCTTTGGCTCCTCAGTGGTACGGGTTGCTGGACATGCACGGGAACAAG
Golden            CCGCCTGACAGCCCTTTGGCTCCTCAGTGGTACGGGTTGCTGGACATGCACGGGAACAAG
QKY_S             CCGCCTGACAGCCCTTTGGCTCCTCAGTGGTACGGGTTGCTGGACATGCACGGGAACAAG
                  ************************************************************

QKY_FlorinaR      GTCAGAGGGGAGCTTATGCTTGCTGTTTGGATGGGGACTCAGGCCGATGAGTCCTTTCCC
Golden            GTCAGAGGGGAGCTTATGCTTGCTGTTTGGGTAGGGACTCAGGCCGATGAGTCCTTTCCC
QKY_S             GTCAGAGGGGAGCTTATGCTTGCTGTTTGGGTAGGGACTCAGGCCGATGAGTCCTTTCCC
                  ***************************** * ************************

QKY_FlorinaR      GATGCTTGGCATTCCGATGCACATGACATTAGCCACGTCAATCTCGCCACCACTCGATCA
Golden            GATGCTTGGCATTCCGATGCACATGACATTAGCCACGTCAATCTCGCCACCACTCGATCA
QKY_S             GATGCTTGGCATTCCGATGCACATGACATTAGCCACGTCAATCTCGCCACCACTCGATCA
                  ************************************************************

QKY_FlorinaR      AAGGTTTACTTCTCCCCCAAGTTATATTACCTTCGAGTTCAAATTCTGCAAGCTCAGGAT
Golden            AAGGTTTACTTCTCCCCCAAGTTATATTACCTTCGAGTTCAAATTCTGCAAGCTCAGGAT
QKY_S             AAGGTTTACTTCTCCCCCAAGTTATATTACCTTCGAGTTCAAATTCTGCAAGCTCAGGAT
                  ************************************************************

QKY_FlorinaR      CTTGTTCCTTGGGATAGAAACCGCCCTTTGGATACATATGTCAAGGTACAGCTTGGGAAC
Golden            CTTGTTCCTTGGGATAGAAACCGCCCTTTGGATACATATGTCAAGGTACAGCTTGGGAAC
QKY_S             CTTGTTCCTTGGGATAGAAACCGCCCTTTGGATACATATGTCAAGGTACAGCTTGGGAAC
                  ************************************************************

QKY_FlorinaR      CAGCTGAGGGTCTCAAGGCCTTCCCAAGTGCATACTATTAACCCCGTTTGGAATGATGAC
Golden            CAGCTGAGGGTCTCAAGGCCTTCCCAAGTGCATACTATTAACCCCGTTTGGAATGATGAC
QKY_S             CAGCTGAGGGTCTCAAGGCCTTCCCAAGTGCATACTATTAACCCCGTTTGGAATGATGAC
                  ************************************************************

QKY_FlorinaR      CTCATGCTCGTGGCCTCCGAGCCTTTCGAAGATATCTTAGTTATAACAGTTGAGGACAGG
Golden            CTCATGCTCGTGGCCTCCGAGCCTTTCGAAGATATCTTAGTTATATCAGTTGAGGACAGG
QKY_S             CTCATGCTCGTGGCCTCCGAGCCTTTCGAAGATATCTTAGTTATATCAGTTGAGGACAGG
                  ******************************************** **************

QKY_FlorinaR      GTTGGTCCTGGAAAGGATGAGATATTAGGGAGGGTGATTCTTTCGGTTAAAGACCTTCCG
Golden            GTTGGTCCTGGAAAGGATGAGATATTAGGGAGGGTGATTCTTTCGGTTAAAGACCTTCCG
QKY_S             GTTGGTCCTGGAAAGGATGAGATATTAGGGAGGGTGATTCTTTCGGTTAAAGACCTTCCG
                  ************************************************************

QKY_FlorinaR      CAGAGAATTGACACTCATAAGCTCCCGGAGCCGATATGGTTCAATCTCCACAAGCCTTCA
Golden            CAGAGAATTGACACTCATAAGCTCCCGGAGCCGATATGGTTCAATCTCCACAAGCCTTCA
QKY_S             CAGAGAATTGACACTCATAAGCTCCCGGAGCCGATATGGTTCAATCTCCACAAGCCTTCA
                  ************************************************************

QKY_FlorinaR      GCAGCTGCTGAAGAGGAAACTAAAAGACAGAAGGAGAAGTTCTCAAGTAAGATTCATCTG
Golden            GCAGCTGCTGAAGAGGAAACTAAAAGACAGAAGGAGAAGTTCTCAAGTAAGATTCATCTG
QKY_S             GCAGCTGCTGAAGAGGAAACTAAAAGACAGAAGGAGAAGTTCTCAAGTAAGATTCATCTG
                  ************************************************************
```

```
QKY_FlorinaR      CGCCTCTGTTTAGACGTGGGTTATCATGTTCTTGATGAGTCCACACACTTTAGCAGCGAT
Golden            CGCCTCTGTTTAGACGTGGGTTATCATGTTCTTGATGAGTCCACACACTTTAGCAGCGAT
QKY_S             CGCCTCTGTTTAGACGTGGGTTATCATGTTCTTGATGAGTCCACACACTTTAGCAGCGAT
                  ************************************************************

QKY_FlorinaR      TTTCAGCCGTCGTCCAGGCACCTGAGGAAATCAGGCATTGGAATTCTTGAGCTTGGGATC
Golden            TTTCAGCCGTCGTCCAGGCACCTGAGGAAATCAGGCATTGGAATTCTTGAGCTTGGGATC
QKY_S             TTTCAGCCGTCGTCCAGGCACCTGAGGAAATCAGGCATTGGAATTCTTGAGCTTGGGATC
                  ************************************************************

QKY_FlorinaR      CTGAGTGCCAAAAAATTTCCAGCTTTGAAGGGAAATGAGGGTAGGACTGCTGATGCATAC
Golden            CTGAGTGCCAGAAAATTTCCAGCTTTGAAGGGAAATGAGGGTAGGACTACTGATGCATAC
QKY_S             CTGAGTGCCAGAAAATTTCCAGCTTTGAAGGGAAATGAGGGTAGGACTACTGATGCATAC
                  ********** ********************************** ***********

QKY_FlorinaR      TGCGTGGCCAAGTATGGAAACAAGTGGGTACGAACCAGAACGCTTCTCGACACTCTGTCT
Golden            TGCGTGGCCAAGTATGGAAACAAGTGGGTGCGAACCAGAACGCTTCTCGACACTCTGTCT
QKY_S             TGCGTGGCCAAGTATGGAAACAAGTGGGTGCGAACCAGAACGCTTCTCGACACTCTGTCT
                  ***************************** ******************************

QKY_FlorinaR      CCTCGCTGGAATGAGCAGTATACTTGGGAAGTTTATGATCCATGTACTGTAATCACCATC
Golden            CCTCGCTGGAATGAGCAGTATACTTGGGAAGTTTATGATCCATGTACTGTAATCACCATC
QKY_S             CCTCGCTGGAATGAGCAGTATACTTGGGAAGTTTATGATCCATGTACTGTAATCACCATC
                  ************************************************************

QKY_FlorinaR      GGTGTTTTCGACAATTGCCATACCAACGGAAGCAGGGAAGACTCGAGAGATCAAAGGATT
Golden            GGTGTTTTCGACAATTGCCATACCAACGGAAGCAGGGAAGACTCGAGAGATAAAAGGATT
QKY_S             GGTGTTTTCGACAATTGCCATACCAACGGAAGTAGGGAAGACTCGAGAGATAAAAGGATT
                  ******************************** ****************** ********

QKY_FlorinaR      GGGAAGGTGAGAATTCGATTATCGACTTTAGAAATTCATCGAGTTTATACGCATTTCTAT
Golden            GGGAAGGTGAGAATTCGATTATCGACTTTAGAAATTCATCGAGTTTATACGCATTTCTAT
QKY_S             GGGAAGGTGAGAATTCGATTATCGACTTTAGAAATTCATCGAGTTTATACGCATTTCTAT
                  ************************************************************

QKY_FlorinaR      CCCCTGCTGATCCTCACACCCTCGGGTTTAAAAAAGCAAGGGGAACTTCAGTTAGCATTG
Golden            CCCCTGCTGATCCTCACACCCTCGGGTTTAAAAAAGCAAGGGGAACTTCAGTTAGCATTG
QKY_S             CCCCTGCTGATCCTCACACCCTCGGGTTTAAAAAAGCAAGGGGAACTTCAGTTAGCATTG
                  ************************************************************

QKY_FlorinaR      AGGTTCACTTGCTTTGCTTGGGTTAACATGTTAGCTCAATACGGAAGACCTTTGCTTCCA
Golden            AGGTTCACTTGCTTTGCTTGGGTTAACATGTTAGCTCAATATGGAAGACCATTGCTTCCA
QKY_S             AGGTTCACTTGCTTTGCTTGGGTTAACATGTTAGCTCAATATGGAAGACCATTGCTTCCA
                  ************************************** ******** *********

QKY_FlorinaR      AAGATGCATTATGTCCAGCCTATACCTATTAGGCACCTTGATTGGCTCCGCCACCAAGCA
Golden            AAGATGCATTATGTCCAGCCTATACCTATTAGGCACCTTGATTGGCTCCGCCACCAAGCA
QKY_S             AAGATGCATTATGTCCAGCCTATACCTATTAGGCACCTTGATTGGCTCCGCCACCAAGCA
                  ************************************************************

QKY_FlorinaR      ATGCAGATTGTAGCAACGAGGCTCGCCCGTTCAGAGCCACCGCTCAGGCGGGAGATTGTC
Golden            ATGCAGATTGTAGCAACGAGGCTCGCCCGTTCAGAGCCACCGCTCAGGCGGGAGATTGTC
QKY_S             ATGCAGATTGTAGCAACGAGGCTCGCCCGTTCAGAGCCACCGCTCAGGCGGGAGATTGTC
                  ************************************************************
```

```
QKY_FlorinaR      GAGTACATGTTAGATATAGACTACCATATGTTTAGTATGAGGAGGAGCAAAGCCAACTTC
Golden            GAGTACATGTTAGACATAGACTACCATATGTTTAGTATGAGGAGAAGCAAAGCCAACTTC
QKY_S             GAGTACATGTTAGACATAGACTACCATATGTTTAGTATGAGGAGAAGCAAAGCCAACTTC
                  ************** *************************** ***************

QKY_FlorinaR      CACCGCATAATGTCGGTTCTCAGCGGGGTTATGACTGTCTGCAGATGGTTTAATGACATT
Golden            CACCGCATCATGTCGGTTCTCAGCGGGGTCATGACTGTCTGCAGATGGTTTAATGACATT
QKY_S             CACCGCATCATGTCGGTTCTCAGCGGGGTCATGACTGTCTGCAGATGGTTTAATGACATT
                  ******** ******************** ******************************

QKY_FlorinaR      TGCAACTGGAGAAACCCGATCACAACGTGCCTCGTCCATATCTTGTTTGTGATATTAGTT
Golden            TGCAACTGGAGAAACCCGATCACAACGTGCCTCGTCCATATCTTGTTTGTGATATTAGTT
QKY_S             TGCAACTGGAGAAACCCGATCACAACGTGCCTCGTCCATATCTTGTTTGTGATATTAGTT
                  ************************************************************

QKY_FlorinaR      TGCTACCCAGAACTAATATTGCCCACAATTTTCCTCTACCTCTTTGTGATTGGTATATGG
Golden            TGCTACCCAGAACTAATATTGCCCACAATTTTCCTCTACCTCTTTGTGATTGGTATATGG
QKY_S             TGCTACCCAGAACTAATATTGCCCACAATTTTCCTCTACCTCTTTGTGATTGGTATATGG
                  ************************************************************

QKY_FlorinaR      AACTACCGGCTCAGGCCAAGGCACCCACTTCACATGGATGCTCGGCTTTCGCAGGCAGAG
Golden            AACTACCGGCTCAGGCCAAGGCACCCACTTCACATGGATGCTCGGCTTTCGCAGGCAGAG
QKY_S             AACTACCGGCTCAGGCCAAGGCACCCACTTCACATGGATGCTCGGCTTTCGCAGGCAGAG
                  ************************************************************

QKY_FlorinaR      GTTGCCCACGCGGATGAGTTGGACGAGGAATTTGACAGCTTCCCCACGGGTCGGCCCGCG
Golden            GTTGCCCACGCGGATGAGTTGGACGAGGAATTTGACAGCTTCCCCACGGGTCGGCCCGCG
QKY_S             GTTGCCCACGCGGATGAGTTGGACGAGGAATTTGACAGCTTCCCCACGGGTCGGCCCGCG
                  ************************************************************

QKY_FlorinaR      GACATTGTGAGGATGAGGTACGACAGGTTGCGTAGCGTGGCGGGCAGGGTGCAGATGGTG
Golden            GACATTGTGAGGATGAGGTACGACAGGTTGCGTAGCGTGGCGGGCAGGGTGCAGATGGTG
QKY_S             GACATTGTGAGGATGAGGTACGACAGGTTGCGTAGCGTGGCGGGCAGGGTGCAGATGGTG
                  ************************************************************

QKY_FlorinaR      GTAGGAGATTTGGCAACCCAAGGGGAAAGAGCACAAGCATTACTAAGCTGGAGGGATCCG
Golden            GTTGGAGATTTGGCAACCCAAGGGGAAAGAGCACAAGCATTACTCAGCTGGAGGGATCCG
QKY_S             GTTGGAGATTTGGCAACCCAAGGGGAAAGAGCACAAGCATTACTCAGCTGGAGGGATCCG
                  ** ***************************************** ***************

QKY_FlorinaR      AGAGCAACGGCAATCTTCATCATCTTCGCGTTGATCTGGGCCATGTTGATATACGTTACT
Golden            AGAGCAACGGCAATCTTCATCATCTTCGCGTTGATCTGGGCCATGTTGATATACGTTACT
QKY_S             AGAGCAACGGCAATCTTCATCATCTTCGCGTTGATCTGGGCCATGTTGATATACGTTACT
                  ************************************************************

QKY_FlorinaR      CCGTTCCGGCTTATAGCAGTGCTGTTCGGTATCTACCTTCTTCGGCATCCACGGTTCAGG
Golden            CCGTTCCGGCTTATAGCAGTGCTGTTCGGTATCTACCTTCTTCGGCATCCACGGTTCAGG
QKY_S             CCGTTCCGGCTTATAGCAGTGCTGTTCGGTATCTACCTTCTTCGGCATCCACGGTTCAGG
                  ************************************************************
```

```
QKY_FlorinaR      AGCAAGATGCATTCTGCACCAGTTAATTTCTTCAAGAGATTGCCATCCAAGTCAGATATG
Golden            AGCAAGATGCATTCTGCACCAGTTAATTTCTTCAAGAGATTGCCATCCAAGTCAGATATG
QKY_S             AGCAAGATGCATTCTGCACCAGTTAATTTCTTCAAGAGATTGCCATCCAAGTCAGATATG
                  ************************************************************

QKY_FlorinaR      CTACTATCTTGA------------------------------------------------
Golden            CTACTATCTTGA------------------------------------------------
QKY_S             CTACTATCTTGATTATTGGTATATTCGCCTAATTCTCGTGAAGAAAAGGTTGTTTAGTAA
                  ************
```

Additional material 4.3 Clustal O alignment of the protein sequences of Quirky from the resistant chromosome of Florina (R) and from the susceptible one (S). Highlighted in green the differences between the two sequences. When there is a ":" sign below the two amino acids it means that they are different but with omology, instead when nothing is showed under the two different amino acids it means that there is a big dissimilarity.

```
R       MANTKLVVEVHDASDLMPKDGDGFASPFVEVNFEGERQRTQTKPKDLNPNWNEKLVFNIN
S       MANTKLVVEVHDASDLMPKDGDGFASPFVEVNFEGERQRTQTKPKDLNPNWNEKLVFNIN
        ************************************************************

R       DRSHLPHKTVDIVVYNDRQTGHHTNFLGRVRISGVSVPFSESQATIQRYPLDKRGLFSHV
S       DRSHLPHKTVDIVVYNDRQTGHHTNFLGRVRISGVSVPFSESQATIQRYPLDKRGVFSHV
        ************************************************************:****

R       KGDIALRIYAIQDYINNGDFAPTPAPPPPTLNDEFVTNSTGGAAGTARPPPLQEINTNRI
S       KGDIALRIYAIQDYINNGDFAPTPAPPPPTLNDEFVTNSTGGAAGTTRPPPLQEINTNRI
        *********************************************:**************

R       VEEIHHHHFGGEKIKK-KEKEVRTFHSIGTGMGGGGGGGGGSHPPPPMSSGFGFETMKEK
S       VEEIHHHHFGGEKIKKKKEKEVRTFHSIGTGMGGGGGGGGGSHPPPPMSSGFGFETMKEK
        *************** ******************************************

R       APTVETRTDFARAGPATVMHMQQQNPEFSLVETDPPLAARRYRGFGGDKTSSTYDLVEQM
S       APTVETRTDFARAGPATVMHMQQQNPEFSLVETDPPLAARRYRGFGGDKTSSTYDLVEQM
        ************************************************************

R       HYLYVSVVKARDLPTMDVTGSLDPYVEVKLGNYKGVTKHVNKDQNPEWHQIFAFSKERVQ
S       HYLYVSVVKARDLPTMDVTGSLDPYVEVKLGNYKGVTKHVDKDQNPVWHQIFAFSKERVQ
        ************************************:***** *************

R       SNLLEVTVKDKDFTKDDIVGRLHFDLSEVPLCMPPDSPLAPQWYGLLDMHGNKVRGELML
S       SNLLEVTVKDKDFTKDDIVGRLHFDLSEVPLCMPPDSPLAPQWYGLLDMHGNKVRGELML
        ************************************************************

R       AVWMGTQADESFPDAWHSDAHDISHVNLATTRSKVYFSPKLYYLRVQILQAQDLVPWDRN
S       AVWVGTQADESFPDAWHSDAHDISHVNLATTRSKVYFSPKLYYLRVQILQAQDLVPWDRN
        ***:********************************************************

R       RPLDTYVKVQLGNQLRVSRPSQVHTINPVWNDDLMLVASEPFEDILVITVEDRVGPGKDE
S       RPLDTYVKVQLGNQLRVSRPSQVHTINPVWNDDLMLVASEPFEDILVISVEDRVGPGKDE
        ********************************************:***********
```

```
R       ILGRVILSVKDLPQRIDTHKLPEPIWFNLHKPSAAAEEETKRQKEKFSSKIHLRLCLDVG
S       ILGRVILSVKDLPQRIDTHKLPEPIWFNLHKPSAAAEEETKRQKEKFSSKIHLRLCLDVG
        ************************************************************

R       YHVLDESTHFSSDFQPSSRHLRKSGIGILELGILSAKKFPALKGNEGRTADAYCVAKYGN
S       YHVLDESTHFSSDFQPSSRHLRKSGIGILELGILSARKFPALKGNEGRTTDAYCVAKYGN
        ************************************:***********:**********

R       KWVRTRTLLDTLSPRWNEQYTWEVYDPCTVITIGVFDNCHTNGSREDSRDQRIGKVRIRL
S       KWVRTRTLLDTLSPRWNEQYTWEVYDPCTVITIGVFDNCHTNGSREDSRDKRIGKVRIRL
        ****************************************************:*********

R       STLEIHRVYTHFYPLLILTPSGLKKQGELQLALRFTCFAWVNMLAQYGRPLLPKMHYVQP
S       STLEIHRVYTHFYPLLILTPSGLKKQGELQLALRFTCFAWVNMLAQYGRPLLPKMHYVQP
        ************************************************************

R       IPIRHLDWLRHQAMQIVATRLARSEPPLRREIVEYMLDIDYHMFSMRRSKANFHRIMSVL
S       IPIRHLDWLRHQAMQIVATRLARSEPPLRREIVEYMLDIDYHMFSMRRSKANFHRIMSVL
        ************************************************************

R       SGVMTVCRWFNDICNWRNPITTCLVHILFVILVCYPELILPTIFLYLFVIGIWNYRLRPR
S       SGVMTVCRWFNDICNWRNPITTCLVHILFVILVCYPELILPTIFLYLFVIGIWNYRLRPR
        ************************************************************

R       HPLHMDARLSQAEVAHADELDEEFDSFPTGRPADIVRMRYDRLRSVAGRVQMVVGDLATQ
S       HPLHMDARLSQAEVAHADELDEEFDSFPTGRPADIVRMRYDRLRSVAGRVQMVVGDLATQ
        ************************************************************

R       GERAQALLSWRDPRATAIFIIFALIWAMLIYVTPFRLIAVLFGIYLLRHPRFRSKMHSAP
S       GERAQALLSWRDPRATAIFIIFALIWAMLIYVTPFRLIAVLFGIYLLRHPRFRSKMHSAP
        ************************************************************

R       VNFFKRLPSKSDMLLS
S       VNFFKRLPSKSDMLLS
        ****************
```

Additional material 4.4 Clustal O alignement of the nucleotidic sequences of the promoter of Quirky from GDDH13 (QKYGold), from the resistant chromosome of Florina (QKYFlo) and from the susceptible one (QKYSeqFLoS). Highlited in green the indel regions and in red the nucleotide difference among the Quirky promoter from resistant allele of Florina and the other two susceptible alleles. Is finally highlighted in yellow the region that is different between the two Florina sequences and the GDDH13 sequence (115 bp). Highlighted in dark green the Start codon of the Quirky gene.

```
QKYGold         ---------------------------------------------ATAAAATATCGATGA
QKYFlo          ---------------------------------------------ATAAAATATCGATAA
QKYSeqFloS      AAGCTGTTTCACGTTTGTCTCTTACCTCTCTCCCCACCCTACGGTATAAAATATCGATGA
                                                             *************.*

QKYGold         TATCGGAAATATCGATAGTTCAAAAATACGAAAATTTTGATGAAAATATCGAGATATTAT
QKYFlo          TATCGAAAATATCAGTAATTCAAAAATACAAAAATTTTGATAAAAATATTAAAATATTAT
QKYSeqFloS      TATCGGAAATATCGATAGTTCAAAAATACGAAAATTTTGATGAAAATATCGAGATATTAT
                *****.*******..**.***********.***********.******* .*.*******

QKYGold         CGATATCGATAGAAATTGAATAAAAACCACGAAAATTATAAGAA-AACTTGGAAATTTTT
QKYFlo          CGATATCGATAGAAATTGAATAAAAATCACGAAAATTGTAAGAAAAACTTGAAAATTTTT
QKYSeqFloS      CGATATCGATAGAAATTGAATAAAAACCACGAAAATTATAAGAA-AACTTGGAAATTTTT
                ************************** **********.****** ******.********

QKYGold         ATTGAAACTTTGCATGATGTTTATTTAGTCAATTATTTATTAATTTATCACAAAAAATTA
QKYFlo          ATTGAAACTTTGCAGGATGTTTATTTAATCAATTATCTATTAGTTAATCACAAAAAATTA
QKYSeqFloS      ATTGAAACTTTGCATGATGTTTATTTAGTCAATTATTTATTAATTTATCACAAAAAATTA
                ************** ************.******** *****.**:***************

QKYGold         GAAGTAAATGCATTGCATGATAGATATAACTGATTTAAGTTAATTATATAGCGAGCTGGC
QKYFlo          GAAGGAAATGCATTACATGATAGATATAATTGATTTAAGTTGATTATATAGCGAGCTGA-
QKYSeqFloS      GAAGTAAATGCATTGCATGATAGATATAACTGATTTAAGTTAATTATATAGCGAGCTGGC
                **** *********.************** ***********.****************.

QKYGold         AAACATTGTGAGTGTAGAAAATATGTAGTAATTAATAAAAGAAATTTAAACACAATATAA
QKYFlo          ----------GTATAGAAAATATGTAGTAATTAATGAAATTAGTTTAAACAGACTATAA
QKYSeqFloS      AAACATTGTGAGTGTAGAAAATATGTAGTAATTAATAAAAGAAATTTAAACACAATATAA
                          **.**********************.*** :*.******** *.******

QKYGold         TCATTTATATATAATGAATTAGTACAATATTTTACACTTTATACAT-G-----CAAGATA
QKYFlo          TCATATATATATAAAGAATTAGTATAATATTTTACACTTTATACATTGCATGGTAAGATA
QKYSeqFloS      TCATTTATATATAATGAATTAGTACAATATTTTACACTTTATACAT-G-----CAAGATA
                ****:*********:********* ******************** *     *******

QKYGold         CATGAGTGACTTAGTAAGGTCTAAAATATCGATGATATCGGAAATATCGGTAGTCCAAAA
QKYFlo          TGTGAATGATTGAGTAAGGTCTAAAATATTGATGATATCAAAAATATAGGTAGTCGAAAA
QKYSeqFloS      CATGAGTGACTTAGTAAGGTCTAAAATATCGATGATATCGGAAATATCTGTAGTCCAAAA
                 .***.*** * ***************** *********..******. ****** ****

QKYGold         ACACGGAAATTTCAATAAAAATATCAAGATATTATTGATATTTTAGACCATCCTCTGCCC
QKYFlo          ACATGAAAATTTCGATAGAAATATCTGAGATATTATGATATTTTAGACCATCCTCTAGCC
QKYSeqFloS      ACACGGAAATTTCAATAAAAATATCAAGATATTATTGATATTTTAGACCATCCTCTGCCC
                *** *.*******.***.*******:...::::*::********************. **
```

```
QKYGold         CACCACTTCTGTCTTTTTCTTCTTCACTCCCACTCCTTCTAGCGATCCTTCCATTTCACC
QKYFlo          CACCACTTCTGTCTTCTTCTTCTTCACTCCCACTCCTTCTAGCGATGCTTCCCTTTAACC
QKYSeqFloS      CACCACTTCTGTCTTTTTCTTCTTCACTCCCACTCCTTCTAGCGATCCTTCCATTTCACC
                ********************************************** *****.***.***

QKYGold         AGACAATATTCTCTGAGCTCTGA-------GCTCTCTCAACTCCGCCGAGCTAGCCTATG
QKYFlo          AGACAATATTCTCTGAGCTCTGAGCTCTGAGCTCTCTCAACTCCGCCGAGCTAGCC----
QKYSeqFloS      AGACAATATTCTCTGAGCTCTGA-------TCTCTCTCAACTGCGCCGAGCTAGCC----
                ************************        *********** *************

QKYGold         GGTCTCATTCCCTTATCCACTTGGTGATCAATTTTTTATGATCACTCTTTGTTAAGTTTG
QKYFlo          ------------------------------------------------------------
QKYSeqFloS      ------------------------------------------------------------

QKYGold         ATATCAAGAGTTGAAATTAAAACATTGAACACTTCTACTTATTCTTCACTCATGGCCAAT
QKYFlo          ----------------------------------------------------ATGGCCAAT
QKYSeqFloS      ----------------------------------------------------ATGGCCAAT
                                                                    *********

QKYGold         ACTAAACTTGTGGTGGAAGTTCACGACGCAAGCGACCTGATGCCGAAAGACGGCGACGGT
QKYFlo          ACTAAACTTGTGGTGGAAGTTCACGACGCAAGCGACCTGATGCCGAAA------------
QKYSeqFloS      ACTAAACTTGTGGTGGAAGTTCACGACGCA------------------------------
                ******************************

QKYGold         TTTGCG
QKYFlo          ------
QKYSeqFloS      ------
```

Additional material 4.5 Protocol of preparation of medium for rigeneration and micropropagation of Gala and Florina plants.

Regeneration media (1L)

4.3 g MS Duchefa® poudre
30 g Sucrose
3 g Phytagel
10 ml Vitamine solution (100X) (double concentration of Nicotinic Acid and Pyridoxine)
500 ul Thidiazuron (10 mg/ml) for a final concentration of 5 mg/L
200 ul Naphthylacetic acid (1mg/ml) for a final concentration of 0.2 mg/L

pH 5.75 and after autoclave add:
6 ml of Kanamycine 25 mg/ml for a final concentration of 150 mg/ml
4.5 ml of Cefotaxime 100 mg/ml for a final concentration of 450 mg/ml

Micropropagation media (1L)

4.3 g MS Duchefa® poudre
30 g Sucrose
6 g Plant Agar
10 ml Vitamine solution (100X) (double concentration of Nicotinic Acid and Pyridoxine)
250 ul Benzyl Aminopurine (1mg/ml) for a final concentration of 0.25 mg/L
500 ul Indole butyric acid (0.1 mg/ml) for a final concentration of 0.05 mg/L

pH 5.75 and after autoclave add :
6 ml of Kanamycine 25 mg/ml for a final concentration of 150 mg/ml
4.5 ml of Cefotaxime 100 mg/ml for a final concentration of 450 mg/ml

References

Alston FH & Briggs JB (1970). Inheritance of hypersensitivity to rosy apple aphid *Dysaphis plantaginea* in apple. Can J Genet Cytol 12:257-258.

Angeli G & Simoni S (2006). Apple cultivars acceptance by *Dysaphis plantaginea* Passerini (Homoptera: *Aphididae*). J of Pest Sci, 79:175-179.

Barbagallo, S. (1996). Afidi delle principali colture fruttifere. Bayer.

Blommers LHM, Helsen HHM & Vaal FWNM (2004). Life history data of the rosy apple aphid *Dysaphis plantaginea* (Pass.) (Homopt., *Aphididae*) on plantain and as migrant to apple. J Pest Sci, 77:155-163.

Brotman Y, Silberstein L, Kovalski I, Perin C, Dogimont C, Pitrat M, ... & Perl-Treves R. (2002). Resistance gene homologues in melon are linked to genetic loci conferring disease and pest resistance. TAG Theor Appl Genet, 104:1055-1063.

Cevik V & King GJ (2002). Resolving the aphid resistance locus *Sd-1* on a BAC contig within a sub-telomeric region of *Malus* linkage group 7. Gen 45:939-945.

Daccord N, Celton JM, Linsmith G, Becker C, Choisne N, Schijlen E, van de Geest H, Bianco L, Micheletti D, Velasco R, Di Pierro EA, Gouzy J, Rees DJG, Guérif P, Muranty H, Durel CE, Laurens F, Lespinasse Y, Gaillard S, Aubourg S, Quesneville H, Weigel D, van de Weg E, Troggio M, Bucher E (2017) High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. Nat Genet 49:1099-1106.

Dangl JL and Jones JD (2001). Plant pathogens and integrated defence responses to infection. Nat 411:826-833.

Dogimont C, Bendahmane A, Chovelon V & Boissot N (2010). Host plant resistance to aphids in cultivated crops: genetic and molecular bases, and interactions with aphid populations. Comptes rendus biol, 333:566-573.

Dogimont C, Chovelon V, Tual S, Boissot N, Rittener-Rüff V, Giovinazzo N, Bendahmane A, Pitrat M (2008). Molecular diversity at the *Vat*/*Pm-W* resistance locus in melon. Cucurbitaceae 2008 219-228.

Erdin N, Tartarini S, Broggini GA, Gennari F, Sansavini S, Gessler C & Patocchi A (2006). Mapping of the apple scab-resistance gene *Vb*. Gen 49:1238-1245.

Fidantsef AL, Stout MJ, Thaler JS, Duffey SS, Bostock RM (1999). Signal interactions in pathogen and insect attack: expression of lipoxygenase, proteinase inhibitor II, and pathogenesis-related protein P4 in the tomato, *Lycopersicon esculentum*. Physiol Mol Plant Pathol, 54:97-114.

Fulton L, Batoux M, Vaddepalli P, Yadav RK, Busch W, Andersen SU, Jeong S, Lohmann JU, Schneitz K. (2009). DETORQUEO, QUIRKY, and ZERZAUST represent novel components involved in organ development mediated by the receptor-like kinase STRUBBELIG in *Arabidopsis thaliana*. PLoS Genet 5(1), e1000355.

Gao LL, Anderson JP, Klingler JP, Nair RM, Edwards OR & Singh KB (2007). Involvement of the octadecanoid pathway in bluegreen aphid resistance in *Medicago truncatula*. Molec Plant-Microbe Interact, 20:82-93.

Gygax M, Gianfranceschi L, Liebhard R, Kellerhals M, Gessler C, Patocchi A (2004). Molecular markers linked to the apple scab resistance gene *Vbj* derived from *Malus baccata* jackii. Theor Appl Genet, 109:1702-1709.

Gururani MA, Venkatesh J, Upadhyaya CP, Nookaraju A, Pandey SK & Park SW (2012). Plant disease resistance genes: current status and future directions. Physiol Mol Plant Pathol 78:51-65.

Halitschke R & Baldwin IT (2003). Antisense LOX expression increases herbivore performance by decreasing defense responses and inhibiting growth-related transcriptional reorganization in *Nicotiana attenuata*. Plant J 36:794-807.

Iorio RA, Di Sandro A, Paris R, Pagliarani G, Tartarini S, Ricci G, ... & Del Duca S. (2012). Simulated environmental criticalities affect transglutaminase of *Malus* and *Corylus* pollens having different allergenic potential. Amino Acids, 42:1007-1024.

Faccioli G, Baronio P, Pasqualini E & Malavolta C (1985). *Dysaphis plantaginea*: ethology and injuriousness. Informatore Agrario, 41:61-63.

Jones JD, Dangl JL (2006) The plant immune system. Nat 444:323-329

Kessler A, Halitschke R & Baldwin IT (2004). Silencing the jasmonate cascade: induced plant defenses and insect populations. Sci, 305:665-668.

Klingler J, Creasy R, Gao L, Nair RM, Calix AS, Jacob HS, ... & Singh KB (2005). Aphid resistance in *Medicago truncatula* involves antixenosis and phloem-specific, inducible antibiosis, and maps to a single locus flanked by NBS-LRR resistance gene analogs. Plant Physiol, 137:1445-1455.

Kovach ME, Elzer PH, Hill DS, Robertson GT, Farris MA, Roop RM & Peterson KM (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene, 166:175-176.

Li C, Williams MM, Loh YT, Lee GI & Howe GA (2002). Resistance of cultivated tomato to cell content-feeding herbivores is regulated by the octadecanoid-signaling pathway. Plant Physiol, 130:494-503.

Li L, Zhao Y, McCaig BC, Wingerd BA, Wang J, Whalon ME, ... & Howe GA (2004). The tomato homolog of CORONATINE-INSENSITIVE1 is required for the maternal control of seed maturation, jasmonate-signaled defense responses, and glandular trichome development. Plant Cell, 16:126-143.

Marchetti E, Civolani S, Leis M, Chicca M, Tjallingii WF, Pasqualini E, Baronio P (2009). Tissue location of resistance in apple to the rosy apple aphid established by electrical penetration graphs. Bull Insect 62:203-208.

McConn M, Creelman RA, Bell E & Mullet JE (1997). Jasmonate is essential for insect defense in *Arabidopsis*. Proceed Nat Acad Sci, 94:5473-5477.

Miles PW (1999) Aphid saliva. Biol Rev Camb Philos Soc 74:41-85.

Mohase L & Van Der Westhuizen AJ (2002). Salicylic acid is involved in resistance responses in the Russian wheat aphid-wheat interaction. J Plant Physiol, 159:585-590.

Pagliarani G, Dapena E, Miñarro M, Denancé C, Lespinasse Y, Rat-Morris E et al. (2016). Fine mapping of the rosy apple aphid resistance locus *Dp-fl* on linkage group 8 of the apple cultivar 'Florina'. Tree Genet Genom, 12:1-12.

Pagliarani G, Paris R, Arens P, Tartarini S, Ricci G, Smulders MJ & van de Weg WE (2013). A qRT-PCR assay for the expression of all *Mal d 1* isoallergen genes. BMC Plant Biol, 13:51.

Pasqualini E, Antropoli A & Zecchini G (1996). Control perspectives against grey aphid (*Dysaphis plantaginea* Pass., Rhynchota Aphididae) on apple: fourth contribution. Informatore Fitopatologico, 46:50-55.

Radchuk VV & Korkhovoy VI (2005). The *rolB* gene promotes rooting in vitro and increases fresh root weight in vivo of transformed apple scion cultivar 'Florina'. Plant cell, tissue and organ culture, 81:203-212.

Rat-Morris E (1994) Analyse des relations entre *Dysaphis plantaginea* Pass. (Insecte Homoptère) et sa plante hôte *Malus* x *domestica* Borkh.: étude de la résistance du cultivar Florina. PhD thesis, Faculté des Sciences, Université de Tours, France (114 p).

Rat-Morris E, Lespinasse Y (1995). Pommier: la résistance au puceron cendré associée à la résistance à la tavelure. Une nouvelle source de résistance mise en évidence chez le cultivar Florina. Phytoma-La Défense des Végétaux, 471:15-17.

Ramakers C, Ruijter JM, Deprez RHL & Moorman AF (2003). Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci lett, 339:62-66.

Royo J, León J, Vancanneyt G, Albar JP, Rosahl S, Ortego F, ... & Sánchez-Serrano JJ (1999). Antisense-mediated depletion of a potato lipoxygenase reduces wound induction of proteinase inhibitors and increases weight gain of insect pests. Proceed Nat Acad Sci, 96:1146-1151.

Rossi M, Goggin FL, Milligan SB, Kaloshian I, Ullman DE, Williamson VM (1998). The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. Proc Natl Acad Sci, 95 :9750-9754.

Seah S, Sivasithamparam K, Karakousis A & Lagudah ES (1998). Cloning and characterisation of a family of disease resistance gene analogs from wheat and barley. TAG Theor Appl Genet, 97:937-945.

Soriano JM, Joshi SG, Van Kaauwen M, Noordijk Y, Groenwold R, Henken B, ... & Schouten HJ (2009). Identification and mapping of the novel apple scab resistance gene *Vd3*. Tree Genet Genom, 5:475-482.

Untergasser A (2006) Maxiprep - Alkaline Lysis Untergasser's Lab. Summer 2006. http://www.untergasser.de/lab/protocols/maxiprep_alkaline_lysis_v1_0.htm

Vaddepalli P, Herrmann A, Fulton L, Oelschner M, Hillmer S, Stratil TF, Fastner A, Hammes UZ, Robinson DG, Schneitz K (2014). The C2-domain protein QUIRKY and the receptor-like kinase STRUBBELIG localize to plasmodesmata and mediate tissue morphogenesis in *Arabidopsis thaliana*. Dev 141:4139-4148.

Van der Westhuizen AJ, Qian XM, Botha AM (1998) Differential induction of apoplastic peroxidase and chitinase activities in susceptible and resistant wheat cultivars by Russian wheat aphid infestation. Plant Cell Rep, 18:132-137.

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A et al. (2010). The genome of the domesticated apple (*Malus* [times] *domestica* Borkh.) Nat Genet 42:833-839.

Chapter 5 – *Malus floribunda* progenies

5.1 ABSTRACT

The principal aim of this work is to identify the gene conferring resistance against rosy apple aphid (RAA) *Dysaphis plantaginea* (Passerini) in *Malus domestica*. Till now, Florina is the best well studied cultivar for its resistance to RAA, characterized by tolerance, antibiosis and antixenosis. When attacked by *D. plantaginea*, this cultivar does not show the typical leaf and shoot deformations. In Florina has been identified Quirky as putative gene responsible for the RAA response, but preliminary analysis suggest that its role is more related to the susceptibility than to the resistance. The identification of two GPI individuals suggest not to exclude the presence of a resistance gene more at the bottom of chromosome 8, that may influence the function of the Quirky gene. New recombining populations from *Malus floribuda* #821, i.e. the wild progenitor of the Florina, will be studied in order to confirm the location of the *Dp-fl* locus. *Malus floribunda* #821, was mostly studied for its resistance to scab (Gessler and Pertot, 2012), but it is also resistant to fire blight (Durel et al. 2009) and to RAA (Pagliarani et al. 2016). Thanks to the availability of new recombinant plants and thanks to their phenotyping it will be possible to increase the precision of the mapping of the RAA resistance locus and possibly identify other genes involved in RAA resistance.

5.2 INTRODUCTION

Rosy apple aphid (RAA, *Dysaphis plantaginea*, Passerini) is one of the most damaging insects affecting cultivated apple *Malus* × *domestica* (Brown and Mathews 2007, Parisi et al. 2013). RAA is present in Europe, North America, North Africa and Asia (Aslan and Karaca 2005, Brown and Mathews 2007, Miñarro and Dapena 2007) and causes severe damage on shoots, leaves and fruits that remain smaller and deformed, leading to significant economic losses (De Berardinis et al. 1994).

Till now, Florina is the best well studied cultivar for its resistance to RAA, characterized by tolerance, antibiosis and antixenosis. When attacked by *D. plantaginea*, this cultivar does not show the typical leaf and shoot deformations. After feeding on Florina plants, RAA has also been shown to be less fecund and with a high mortality (Rat-Morris 1994). On Florina plants, RAA has been observed moving from the leaves to the stems, suggesting repellent compounds released by the leaves or a more difficult stylet penetration (Angeli and Simoni, 2006). Consistently, electrical penetration graphs demonstrated that RAA stays on Florina leaves without stylet penetration for a longer period before the first probe and with a reduced duration of sap ingestion, thus indicating mechanical, biochemical or vascular resistance (Marchetti et al. 2009).

Malus floribunda #821 is the wild progenitor of Florina (Chapter 1, Figure 1.5). Being a very well-know wild species also called Japanese flowering crabapple, this wild species was mostly studied for its resistance to scab (Gessler and Pertot 2012), but it is also resistant to fire blight (Durel et al. 2009) and to RAA (Pagliarani et al. 2016). It was the first species studied for the scab resistance, named *Vf* gene, where a group of closely linked genes were identified. This gene was recently renamed *Rvi6* (Bus et al. 2011). *Malus floribunda* #821 was used as progenitor in many breeding programs worldwide to introduce scab resistance gene in new resistant cultivars and thus producing also unintentionally some RAA resistant cultivar. The *Malus floribunda* #821 derivative F_2-26829-2-2 was certainly carrying the scab resistance gene (Gessler and Pertot 2012), but also the RAA resistance gene that was introgressed in various scab-resistant cultivars, including GoldRush, Galarina, Liberty and Golden Orange (Pagliarani et al. 2016).

Recently, using different segregating populations of Florina the *Dp-fl* locus was mapped at the bottom of the chromosome 8 in a region of only 56 Kb. In this process two genotype-phenotype incongruences (GPI) individuals were identified in the Perico x Florina progeny (Chapter 4, Figure 4.4). Excluding these two individuals with resistant phenotype from the list of recombinants used for the gene mapping, no individuals recombining in the *Dp-fl* region resulted with resistant phenotype. Unfortunately, the GPI individuals were no more available to further test their resistance against RAA. In the present work, the main objective is to confirm the location of the *Dp-fl* locus of Florina in new progenies deriving from *Malus floribunda* #821, i.e. the wild progenitor of the Florina cultivar. During this process new phenotypic tests were performed and new polymorphic markers were developed. The mapping of the resistance gene against the RAA was then mapped. In this work the locus mapped in the *M. floribunda* progenies will be named differently (*Dp-Mflo - Dysaphis plantaginea* resistance from *Malus floribunda* #821) from the locus studied in the Florina progenies, even if the resistance locus is assumed to be the same.

5.3 MATERIALS AND METHODS

5.3.1 Plant material

Different segregating progenies with *Malus floribunda* #821 as parent were bred at INRA, Angers, France, to characterize the *Vfh* scab resistance gene (Bénaouf and Parisi, 2000). In this work, we utilized the same plant material but with the final goal of identifying individuals recombining within the *Dp-fl* locus. A total of about 700 individuals derived from two different progenies of *Malus floribunda* #821 have been utilized in this work: a Gala x *Malus floribunda* progeny consisting of 188 individuals and a Golden Delicious x *Malus floribunda* progeny including 500 individuals.

5.3.2 Marker development and *Dp-Mflo* mapping

Genomic DNA of all the progeny plants was isolated from young leaf tissues using a cetyltrimethylammonium bromide (CTAB) protocol according to Aldrich and Cullis (1993) with minor modifications. DNA concentration was assessed with a NanoDrop spectrophotometer (Thermo Fischer) and adjusted to 5 ng/µL.

For the identification of recombinant individuals, the two simple sequence repeats (SSR) markers that were already available at the bottom of the chromosome 8 were initially utilized: SSR_C and C-13470. Also the marker CH01h10 was utilized on the left border of the window. Then, during the *Dp-Mflo* mapping, other SSR markers were used, some of them were developed in the *Dp-fl* region in Florina and other were developed outside of the *Dp-fl* region. The complete list of the markers utilized in this work are listed in Additional Material 5.1. All microsatellites were identified from the newly available genome sequence of the GDDH13 (Daccord et al. 2017) using the 'Tandem Repeats Finder' tool (https://tandem.bu.edu/trf/trf.html, setting the maximum period size as 3 and the minimum copy number at 12). Primers were designed with Primer3 (http://primer3.ut.ee, version 4.0.0,) using default parameters.

A first evaluation of the amplified fragments was performed according to Gianfranceschi et al. (1998) in a 5% denaturing polyacrylamide gel using a 100 bp ladder (Biotium) as size standard. The gel was further stained with the silver staining method and images were acquired using the 440 CF Kodak Image System. A confirmation of the obtained result was then performed by capillary sequencing analysis. PCR amplifications were performed on the available recombinant individuals in a Biorad DNA Engine® thermal cycler in a volume of 11 µL containing 1.1 × Qiagen Multiplex PCR Master Mix, 0.2 µM of forward and reverse primers and 10 ng DNA. Amplification was done with a touchdown program as follows: a first denaturation step at 94 °C for 15 min, four cycles of denaturation (94 °C for 30 s), annealing (57 °C for 1 min, decreasing by one degree after each cycle), and extension (72 °C for 1 min). The program then continues with 29 cycles of denaturation (94 °C for 30 s), annealing (50 °C for 1 min), and extension (72 °C for 1 min). A final annealing (50 °C for 15 min) and extension step (72 °C for 15 min) were added. Electrophoresis of PCR products was performed using a 4-capillary sequencer (ABI 3130; ANAN platform, INRA-Angers). Amplification products were diluted 30 times and mixed (2.5 µL) with formamide (9.35 µL) and Gene Scan 500 LIZ (PE Applied Biosystems) as standard (0.15 µL).

5.3.3 Phenotypic evaluation of RAA resistance

Malus floribunda progenies are available in field at INRA, Angers. Individuals that were identified as recombining in the *Dp-Mflo* region according to the genotyping described above were grafted on MM106 rootstocks for a total of 10 replicates for each recombinant.

RAA resistance tests were performed under controlled conditions in a greenhouse at INRA, Angers, France, as follows. All the individuals were grown in pots at a temperature of about 22°C. Ten replicates of Florina, *Malus floribunda* #821, Gala and Golden Delicious genotypes were also used as resistant and susceptible controls. Six other individuals that were not recombining in the region were also added to the phenotypic test to enlarge the number of resistant and susceptible controls: three individuals with fully resistant genotype and three with fully susceptible genotype. Plants were artificially infested with RAA in mid-June when the young shoots were about 20-30 cm. To avoid aphid movements between plants, pots were placed on petri dishes in a plate filled with water, paying attention to prevent shoots overlapping. Infestation was performed placing two young adult apterous virginiparous females on the youngest and well-expanded leaf of each plant with a paint brush. All aphids were derived from a clonal aphid line of RAA obtained from one founder collected in the field and reared on seedlings of Golden Delicious. Scoring of infested plants was done 21 days after infestation. The following scoring scale was used: 0 = no leaf distortion; 1= leaf very slightly curled; 2= leaf slightly curled and 3= typically rolled leaf (Rat-Morris and Lespinasse 1995).

The phenotypic tests were performed once in 2016 and three times in 2017, but not all the individuals were phenotyped in all the tests. Then, phenotypic data were classified in 4 classes that are reported in Figure 5.1: when the individual was phenotyped twice or thrice as resistant the 'R' phenotype is written in a dark-green box, when it was phenotyped as resistant only once the 'R' phenotype is written in a light-green box, otherwise when the individual was phenotyped twice or thrice as susceptible the 'S' phenotype is written in a dark-rosy box and when was phenotyped as susceptible only once the 'S' phenotype is written in a light-rosy box.

5.3.4 Fine mapping of the *Dp-Mflo* region

After marker scoring and phenotypic evaluation, the graphical genotypes (Young and Tanksley, 1989) were drawn by combining data. Genotypic data were shown as 'AA' for the allele linked to susceptibility and 'AB' for the 'resistant' allele. When the genotypic data was not available and it was possible to deduce the genotype from the two flanking markers results (considering that double recombination events are very unlikely to occur in), the resistant phase was written 'ab' and the susceptible 'aa'.

5.3.5 Identification of candidate genes in the *Dp-Mflo* region

The identified *Dp-Mflo* region was deeply investigated in the GDDH13 genome available on line for the identification of candidate genes putatively involved in the response against *D. plantaginea*. Both structural and functional annotation was performed in the Golden Delicious genome, in particular the functional annotation is referred to the homolog gene studied in *A. thaliana* (TAIR website hwww.arabidopsis.org).

5.4 RESULTS

5.4.1 Phenotypic results

A total of 24 recombinant individuals were identified recombining between SSR_C and SSR C13470. Only 8 individuals were phenotyped in 2016 (R13-A056, R14048, R14-A105, R14-A088, R13-A146, R14-A062, R13-A062, R16-A024), all the remaining individuals were phenotyped in 2017. Phenotypic results are shown in the graphical genotype Figure 5.1. From this figure it is possible to deduce that most of the individuals tested resulted with resistant phenotype: 17 recombinant individuals versus 5 susceptible. Only two recombinant individuals (R14-A105 and R13-A146) have a dubious phenotype, which are shown with a dot in a white box in the Figure 5.1. A total of 13 recombinant individuals were phenotyped at least two times as resistant while only 4 were phenotyped only once. Two recombinant individuals were clearly susceptible twice, while the other three individuals resulted susceptible only once. The non-recombinant individuals showed a good correlation with the expected phenotype apart from the individual R14-A043 that resulted with uncertain phenotype. Golden and Gala susceptible controls resulted clearly susceptible in all the tests and Florina and *Malus Floribunda* #821 resulted clearly resistant.

5.4.2 Mapping of the *Dp-Mflo* locus

Genotypic results obtained with the development of new SSR markers are shown in graphical genotype (Figure 5.1). Initially only two SSR markers were used to identify recombinant markers, the SSR_C and the marker C13470. These two markers are about 2.4 million bp apart. Other markers developed for the *Dp-fl* mapping in Florina progenies were then added to this: SSR_377, SSR_228 and SSR_4. With this genotypic data and the first phenotyping of 2016, it was already clear that it was not possible to locate the *Dp-Mflo* locus in the same position as the *Dp-fl*. Indeed, the *Dp-fl* locus was located between the SSR_377 and SSR_228, but locating the phenotypic results of the recombinant progenies of *M. floribunda* progenies obtained in 2016, two individuals were in a wrong position (R14-A048, R13-A088). Enlarging the number of recombinant individuals tested in 2017, it was clearer that

Dp-Mflo was located outside of the *Dp-fl* locus. Further markers were necessary to fine-map the *Dp-Mflo* region between the SSR_4 and the SSR C13470, so that six other markers were designed (SSR_322, SSR-396, SSR_580, SSR-666, SSR_1250, SSR_1600). Thus, it was possible to map the *Dp-Mflo* between the SSR_666 and the SSR_1250, in a region of about 600 Kb. Only one individual resulted with phenotype-genotype incongruence (GPI), but the phenotypic data result from a single test. Further phenotypic tests are necessary to validate this result. The position of the *Dp-Mflo* locus is about 800 Kb distant from the *Dp-fl* locus in Florina. The complete list of the markers used during this fine-mapping process are listed in Additional Material 5.1.

5.4.3 Candidate genes in the *Dp-Mflo* region

In the *Dp-Mflo* region of about 600 Kb of the GDDH13 genome is present a big gap of about 150 Kb where there is no sequence to refer. The remaining part of the sequence includes a total of 55 predicted genes (from the gene code number MD08G1224800 to MD08G1230300) (Table 5.1), including two non-coding genes (MD08G1228800 and MD08G1229800). Among the 53 predicted coding genes, 5 genes in *A. thaliana* have been described as involved in defense response (MD08G1228400, MD08G1229600, MD08G1229700, MD08G1230100 and MD08G1230200). In detail, the gene MD08G1228400 is predicted to encode for an auxin signaling F-box gene. The gene MD08G1229700 encodes for a protein that belongs to protein kinase family. The three other genes MD08G1229600, MD08G1230100 and MD08G1230200 are predicted to encode for putative genes that are involved in the defense response. Furthermore, the predicted function of four other genes could be related to the attack of the aphid *D. plantaginea*: MD08G1225600 is predicted to be involved into oxidative stress; MD08G1226500 is predicted to be involved in ethylene-signaling pathway and heat acclimation, MD08G1227000 is predicted to be involved in response to cold, to osmotic stress, to salt stress and to wounding, and finally MD08G1227500 is predicted to be expressed in response to jasmonic acid and to light stimulus. It seems that the more interesting part of the sequence is the last part, where four genes putatively involved in the defense response are closely located.

Plant	CH01h10	SSR C	SSR_377	SSR_228	SSR_4	SSR_322	SSR_396	SSR_580	SSR_666	*Dp-Mflo*	SSR_1250	SSR_1600	SSR C13470
R14-A178	AB	AB	AB	AB	AB	AB	AB	AB	AB	R	AB	AB	AA
R14-A113	AB	AB	AB	AB	AB	AB	AB	AB	AB	R	AB	AA	AA
R15-A182	AB	AB	AB	ab	ab	AB	AB	AB	AB	R	AA	AA	AA
R13-A136	AB	AB	AB	AB	AB	AB	AB	AB	AB	R	AA	AA	AA
R13-A056	AB	AB	AB	AB	AB	AB	AB	AB	AB	R	AA	AA	AA
R15-A067	AB	AB	ab	ab	AB	ab	AB	AB	AB	R	AA	AA	AA
R15-A156	AB	AB	AB	AB	AB	AB	AB	AB	AB	R	AA	AA	AA
R15-A184	AB	ab	AB	ab	ab	AB	AB	AB	AB	R	AA	AA	AA
R14-A105	AB	AB	AB	AB	AB	AB	AB	AB	AB	-	AA	AA	AA
R14-A048	AB	AB	AB	AB	AB	AB	AB	AB	AB	S	AA	aa	AA
R14-A120	AB	AB	AB	AB	AB	AB	AB	AB	AA	S	AA	AA	AA
R15-A130	AB	AB	AB	ab	ab	AB	AB	AB	AA	R	AA	AA	AA
R13-A142	AB	AB	AB	AB	AB	AB	AA	AA	AA	S	AA	AA	AA
R13-A060	AA	aa	AA	aa	AA	AA	aa	aa	AA	S	AB	ab	AB
R14-A062	AA	AA	AA	AA	AA	AA	AA	AA	AA	S	AB	AB	AB
R13-A146	AA	AA	AA	AA	AA	AA	AA	AA	AA	-	AB	AB	AB
R14-A088	AA	AA	AA	AA	AA	AA	AA	AA	AA	R	AB	AB	AB
R13-A196	AA	aa	AA	aa	AA	AA	aa	aa	AA	R	AB	ab	AB
R13-A064	AA	AA	AA	AA	AA	AA	AA	AA	AA	R	AB	AB	AB
R13-A162	AA	aa	AA	aa	AA	AA	aa	aa	AA	R	AB	ab	AB
R15-A166	AA	AA	AA	AA	AA	AA	AA	AA	AA	R	AB	AB	AB
R13-A202	AA	aa	aa	aa	AA	AA	-	-	AB	R	AB	ab	AB
R13-A004	AA	aa	AA	aa	AA	AA	-	-	AB	R	AB	ab	AB
R13-A220	AA	-	-	-	AB	AB	ab	ab	AB	R	ab	ab	AB
R15-A097	AB	AB	AB	AB	AB	AB	AB	AB	AB	R	AB	AB	AB
R13-A062	AB	AB	AB	AB	AB	AB	AB	AB	AB	R	AB	AB	AB
R16-A024	AB	AB	AB	AB	AB	AB	AB	AB	AB	R	AB	ab	AB
R14-A185	AA	AA	aa	AA	AA	AA	aa	AA	AA	S	AA	aa	aa
R16-A023	AA	AA	aa	aa	aa	aa	AA	AA	AA	S	AA	AA	AA
R14-A043	AA	AA	AA	AA	AA	AA	AA	AA	AA	-	AA	AA	AA

Figure 5.1 Graphical genotype from the *Malus floribunda* progenies. Individuals phenotyped in 2016 are highlighted in purple, non-recombinant plants are highlighted in blue and the GPI individual is highlighted in yellow. Phenotypic data are showed in the *Dp-Mflo* column: individuals phenotyped twice or thrice as resistant 'R' phenotype was written in a dark-green box, individuals phenotyped as resistant only once the 'R' phenotype is written in a light-green box; individuals phenotyped twice or thrice as susceptible 'S' phenotype is written in a dark-rosy box and individuals phenotyped as susceptible only once 'S' phenotype is written in a light-rosy box. Genotypic data are shown as 'AA' when susceptible and 'AB' when resistant; when the genotypic data is deduced from the two flanking markers the resistant phase was written 'ab' and the susceptible 'aa'. Missing data are a dot (-) in a white cell. SSR markers developed within the *Dp-fl* locus of Florina are written in light-blue boxes

Table 5.1 List of the genes included in the *Dp-Mflo* region of about 600 Kb obtained from the GDDH13 annotation free available on line at https://iris.angers.inra.fr/gddh13 (Daccord et al. 2017). In the first column are listed the code of the genes referred to the GDDH13 genome, in the second column there are listed the functions inferred to the genes in GDD13 and in the last column there are specified the biological function linked to the genes into the TAIR website (website hwww.arabidopsis.org).

Name	Gene	Biological Process
MD08G1224800	UDP-Glycosyltransferase superfamily protein	-
MD08G1224900	Plantacyanin	anther development, pollination
MD08G1225000	Lung seven transmembrane receptor family protein	biological_process
MD08G1225100	partner of Y14-MAGO	mRNA transport, nuclear-transcribed mRNA catabolic process, nonsense-mediated decay, positive regulation of gene expression, regulation of translation
MD08G1225200	Ribosomal protein L10 family protein	translation
MD08G1225300	Plant protein of unknown function (DUF247)	biological_process
MD08G1225400	homolog of separase	chromosome separation, cytokinesis, embryo development ending in seed dormancy, endosperm development, exocytosis, meiotic chromosome separation, positive regulation of sister chromatid cohesion, proteolysis, regulation of establishment of cell polarity, response to glucose, unidimensional cell growth, vesicle-mediated transport
MD08G1225500	Plant protein of unknown function (DUF247)	biological_process
MD08G1225600	gene	ER to Golgi vesicle-mediated transport, intracellular protein transport, response to oxidative stress, vesicle-mediated transport
MD08G1225700	Plant protein of unknown function (DUF247)	biological_process
MD08G1225800	Other Eukaryotes -0 (source: NCBI BLink).	phloem development
MD08G1225900	homolog of separase	chromosome separation, cytokinesis, embryo development ending in seed dormancy, endosperm development, exocytosis, meiotic chromosome separation, positive regulation of sister chromatid cohesion, proteolysis, regulation of establishment of cell polarity, response to glucose, unidimensional cell growth, vesicle-mediated transport
MD08G1226000	gene	protein deubiquitination, ubiquitin-dependent protein catabolic process
MD08G1226100	homolog of separase	chromosome separation, cytokinesis, embryo development ending in seed dormancy, endosperm development, exocytosis, meiotic chromosome separation, positive regulation of sister chromatid cohesion, proteolysis, regulation of establishment of cell polarity, response to glucose, unidimensional cell growth, vesicle-mediated transport
MD08G1226200	homolog of separase	chromosome separation, cytokinesis, embryo development ending in seed dormancy, endosperm development, exocytosis, meiotic chromosome separation, positive regulation of sister chromatid cohesion, proteolysis, regulation of establishment of cell polarity, response to glucose, unidimensional cell growth, vesicle-mediated transport
MD08G1226300	alpha/beta-Hydrolases superfamily protein	-
MD08G1226400	chaperone protein dnaJ-related	biological_process
MD08G1226500	gene	cell division, cellular heat acclimation, deadenylation-independent decapping of nuclear-transcribed mRNA, ethylene-activated signaling pathway, gravitropism, heat acclimation, mRNA catabolic process,mRNA processing, miRNA catabolic process, nuclear-transcribed mRNA catabolic process, exonucleolytic, positive regulation of mRNA catabolic process, regulation of gene expression, epigenetic, response to 1-aminocyclopropane-1-carboxylic acid, response to ethylene, unidimensional cell growth, vasculature development
MD08G1226600	homolog of separase	chromosome separation, cytokinesis, embryo development ending in seed dormancy, endosperm development, exocytosis, meiotic chromosome separation, positive regulation of sister chromatid cohesion, proteolysis, regulation of establishment of cell polarity,

		response to glucose, unidimensional cell growth, vesicle-mediated transport
MD08G1226700	homolog of separase	chromosome separation, cytokinesis, embryo development ending in seed dormancy, endosperm development, exocytosis, meiotic chromosome separation, positive regulation of sister chromatid cohesion, proteolysis, regulation of establishment of cell polarity, response to glucose, unidimensional cell growth, vesicle-mediated transport
MD08G1226800	Protein of unknown function (DUF1644)	biological_process
MD08G1226900	Other Eukaryotes -4868 (source: NCBI BLink).	-
MD08G1227000	MAPK/ERK kinase kinase 1	MAPK cascade, cold acclimation, innate immune response, protein autophosphorylation, regulation of mitotic cell cycle, response to L-glutamate, response to cadmium ion, response to cold, response to osmotic stress, response to salt stress, response to wounding, root meristem growth, root system development, stress-activated protein kinase signaling cascade
MD08G1227100	phosphotyrosyl phosphatase activator (PTPA) family protein	positive regulation of catalytic activity
MD08G1227200	WRKY DNA-binding protein 9	regulation of transcription, DNA-templated, transcription, DNA-templated
MD08G1227300	protein-protein interaction regulator family protein	biological_process
MD08G1227400	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family	biological_process
MD08G1227500	subunit CSN8	COP9 signalosome assembly, multicellular organism development, photomorphogenesis, protein deneddylation, red, far-red light phototransduction, response to jasmonic acid, response to light stimulus
MD08G1227600	RNA polymerase Rpb7 N-terminal domain-containing protein	RNA metabolic process, RNA splicing, transcription initiation from RNA polymerase III promoter, transcription, DNA-templated
MD08G1227700	gene	gluconeogenesis, glycolytic process, positive regulation of flower development, starch metabolic process
MD08G1227800	Plant protein of unknown function (DUF639)	-
MD08G1227900	GNS1/SUR4 membrane protein family	biological_process
MD08G1228000	Protein of unknown function (DUF3411)	biological_process
MD08G1228100	CAMV movement protein interacting protein 7	transport of virus in multicellular host, vesicle-mediated transport
MD08G1228200	TCP-1/cpn60 chaperonin family protein	'de novo' protein folding, chaperone-mediated protein folding, embryo development ending in seed dormancy, mitochondrion organization, protein refolding
MD08G1228300	RING/U-box superfamily protein	
MD08G1228400	auxin signaling F-box 3	cellular response to nitrate, defense response, lateral root development, pollen maturation, primary root development, protein ubiquitination, stamen development
MD08G1228500	Other Eukaryotes -0 (source: NCBI BLink).	biological_process
MD08G1228600	Other Eukaryotes -0 (source: NCBI BLink).	biological_process
MD08G1228700	F-box/RNI-like superfamily protein	biological_process
MD08G1228800	5S_rRNA	
MD08G1228900	emp24/gp25L/p24 family/GOLD family protein	intracellular protein transport, protein transport, vesicle-mediated transport
MD08G1229100	maternal effect embryo arrest 9	embryo development ending in seed dormancy, pollen development
MD08G1229200	Other Eukaryotes -7 (source: NCBI BLink).	biological_process
MD08G1229300	gene	biological_process
MD08G1229400	Other Eukaryotes -3081 (source: NCBI BLink).	biological_process
MD08G1229500	OB-fold-like protein	biological_process

MD08G1229600	putative	defense response, defense response to bacterium, defense response to bacterium, incompatible interaction, response to auxin, signal transduction, systemic acquired resistance, salicylic acid mediated signaling pathway
MD08G1229700	protein kinase family protein	defense response, phosphorylation, protein phosphorylation, regulation of transcription, DNA-templated, transcription, DNA-templated
MD08G1229800	gene	
MD08G1229900	Phosphatidylinositol-4-phosphate 5-kinase family protein	-
MD08G1230000	Phosphatidylinositol-4-phosphate 5-kinase family protein	-
MD08G1230100	putative	defense response, signal transduction
MD08G1230200	putative	defense response, signal transduction
MD08G1230300	Phosphatidylinositol-4-phosphate 5-kinase family protein	-

5.5 DISCUSSION

In order to distinguish both progeny materials (i.e. Florina progenies and *M. floribunda* progenies), *Dp-fl* and *Dp-Mflo* have been called differently, but the whole genomic region of *M. floribunda* #821 has been shown to be fully inherited in Florina thanks to the available SSR markers. Thus, the simplest genetic explanation is that *Dp-fl* and *Dp-Mflo* should correspond to a single and same locus and gene. Nevertheless, an alternative, and more complex, genetic model could be proposed with the presence of two nearby RAA resistance genes in the bottom part of LG8, one being identified in Florina progenies, and the second one, located downstream, being identified *in M. floribunda* #821 progenies. For that latter situation, a complex influence of the genetic background should be invoked to be able to explain the alternative detection of either gene according to the Florina versus *M. floribunda* #821 background. Thus, this second proposal is much less probable than the first hypothesis of a same single RAA resistance gene in both Florina and *M. floribunda* #821.

The previous location of *Dp-fl* in Florina progenies was somewhat limited by the phenotyping process, which was no more possible for some recombinant individuals which were unfortunately not maintained in the field. The presence of two GPI from the Perico x Florina population (PF_P001 and PF_X-9504-33) was thus a weak point for the accurate position of *Dp-fl*. The availability of *M. floribunda* #821 progenies made it possible to challenge the position of *Dp-fl* locus. The newly detected position of the so-called *Dp-Mflo* RAA resistance gene indicates a genomic position about 600 Kb downstream on the chromosome 8. Furthermore, even if the total number of individuals from many different progenies of Florina were taken into consideration, not many individuals recombining in the *Dp-fl* locus were identified. Counting all the populations taken in consideration in the Florina mapping (Perico x Florina – 79 individuals; Raxao x Florina – 41; Florina x Melrose – 28; De la Riega x Florina – 43; Royal Gala x Florina – 94; Perleberg x Florina – 92, GoldRush x Florina -181; Meana x Florina -320) a total of 878 individuals were taken in consideration, so a total of 8 recombinant plants were identified in the *Dp-fl* region of about 300 Kb. While, in the *M. floribunda* progenies of about 700 individuals a total of 16 recombinants were identified in a region of about 600 Kb. Resulting in a limited number of recombinant plants identified in the Florina progenies considering the corresponding genetic distance; while a higher number was found in *M. floribunda* progenies.

The sixteen individuals recombining in the *Dp-Mflo* region are a good starting point for the fine-mapping of the *Dp-Mflo* region, but further phenotypic tests and new polymorphic markers are needed. The genes that are within the region seems to have a role in the defense response, indicating that this region could be involved in this function. Anyway, this analysis performed on the GDDH13 genome is only a starting point. Indeed, the complete sequence of the *M. floribunda Dp-Mflo* region could reveal

new genes, such as NBS-LRR genes, present only in the resistant phase that could explain the difference between the resistant and the susceptible individuals.

5.6 CONCLUSION

In conclusion, a new position of the locus conferring resistance to RAA was identified by analyzing both genotypic and phenotypic results of recombinant plants from *Malus floribunda* progenies. The *Dp-Mflo* region is about 600 Kb and is located 800 Kb downstream from the *Dp-fl* locus identified in Florina. The whole genomic region of *M. floribunda* #821 has been shown to be inherited in Florina thanks to the available SSR markers, so *Dp-fl* and *Dp-Mflo* should correspond to a single and same locus and gene. Nevertheless, an alternative second hypothesis of genetic model could be proposed with the presence of two nearby RAA resistance genes in the bottom part of LG8, one being identified in Florina progenies, and the second one, located downstream, being identified *in M. floribunda* progenies. To identify the genes involved in the resistance to *D. plantaginea* it will be necessary to have the complete sequence of the *M. floribunda* resistant chromosome and further steps of the BAC library screening of its descendant Florina would be necessary. The fine-mapping of the locus and the complete sequence will open a new rout towards the final cloning of the *Dp-fl/Mflo* RAA resistance gene.

5.7 Additional material

Additional Material 5.1 List of the primers utilized for the *Dp-Mflo* locus.

Primer name	Sequence (5'-3')	Amplicon length (bp)	Analysis	References
Ch01h10 F	TGCAAAGATAGGTAGATATATGCCA	94-114	Polyacrylamide and Sequencer	Wang et al. 2012
Ch01h10 R	AGGAGGGATTGTTTGTGCAC			
SSR C F	TGGCGGTCTCCTTTTGTTCA	259	Polyacrylamide	*Dp-fl* region
SSR C R	ACCCATCAATCATCATCCTACCA			
SSR 377 F	TGAAGCTCTTGGCTTCAATGAC	238/270	Polyacrylamide and Sequencer	*Dp-fl* region
SSR 377 R	GGTTCTCAGCCCTCTTATCTTCC			
SSR 228 F	ACCTGGTTGTGTGAGCATCC	191/231	Polyacrylamide and Sequencer	*Dp-fl* region
SSR 228 R	ACCAACGCCAAGCTATCTCA			
SSR 4 F	CTCCACAATGATGCTTAGTTGGT	228/266	Polyacrylamide and Sequencer	*Dp-fl* region
SSR 4 R	GGAATGAAATCAGCCAGATGGG			
SSR 322 F	CTTCCCTCCCCACCTGATAA	379/435	Polyacrylamide and Sequencer	New
SSR 322 R	GCTGGTCCTCCTCTTTCGA			
SSR 396 F	TTGGGTCCCTAGCTAGCATT	242	Polyacrylamide	New
SSR 396 R	TCTAACTCTTCCTTGCAAGCA			
SSR 580 F	CTTTCCACCTAGCTGTAACGG	213	Polyacrylamide	New
SSR 580 R	TGGTTTTGGATTGACGTGCT			
SSR 666 F	TCCACTTGCACATAGGGTACA	246	Polyacrylamide and Sequencer	New
SSR 666 R	CCATGGAGGCCTATCATGC			
SSR 1250 F	TTATGTTTGTGTGGTCGCGG	357	Polyacrylamide	New
SSR 1250 R	TAGATCGGGTACAACGCAGG			
SSR 1600 F	CGCAAGTCGTCCCAAAGAG	241	Polyacrylamide	New
SSR 1600 R	AGTTAGGCATTGGGAGTGGT			
SSR C13470 F	TCGATTCCTCAATCTCTCTCA	230/280	Polyacrylamide and Sequencer	Wang et al. 2012
SSR C13470 R	ATCGGAGAAAACCCAAATCC			

References

Aldrich C & Cullis CA (1993). CTAB DNA extraction from plant tissues. Plant Mol Biol Rep, 11:128-141.

Angeli G & Simoni S (2006). Apple cultivars acceptance by *Dysaphis plantaginea* Passerini (Homoptera: Aphididae). J Pest Sci 79:175-179.

Aslan B, Karaca I (2005). Fruit tree aphids and their natural enemies in Isparta region, Turkey. J Pest Sci 78:227-229.

Bénaouf G & Parisi L (2000). Genetics of host-pathogen relationships between *Venturia inaequalis* races 6 and 7 and Malus species. Phytopathol 90(3), 236-242.

Brown MW, Mathews CR (2007) Conservation biological control of rosy apple aphid, *Dysaphis plantaginea* (Passerini), in Eastern North America. Environ Entomol, 36:1131-9.

Bus VG, Rikkerink EH, Caffier V, Durel CE & Plummer KM (2011). Revision of the nomenclature of the differential host-pathogen interactions of *Venturia inaequalis* and *Malus*. Annual Rev Phytopathol 49:391-413.

Daccord N, Celton JM, Linsmith G, Becker C, Choisne N, Schijlen E, van de Geest H, Bianco L, Micheletti D, Velasco R, Di Pierro EA, Gouzy J, Rees DJG, Guérif P, Muranty H, Durel CE, Laurens F, Lespinasse Y, Gaillard S, Aubourg S, Quesneville H, Weigel D, van de Weg E, Troggio M, Bucher E (2017) High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. Nat Genet 49:1099-1106.

De Berardinis E, Baronio P, Baumgartner J (1994) The effect of aphid *Dysaphis plantaginea* (Pass, Hom, Aphididae) feeding on apple fruit growth. Ecol Modell 72:115-127.

Durel CE, Denancé C & Brisset MN (2009). Two distinct major QTL for resistance to fire blight co-localize on linkage group 12 in apple genotypes 'Evereste'and *Malus floribunda* clone 821. Genom, 52:139-147.

Gessler C & Pertot I (2012). *Vf* scab resistance of *Malus*. Trees 26:95-108.

Gianfranceschi L, Seglias N, Tarchini R, Komjanc M, Gessler C (1998). Simple sequence repeats for the genetic analysis of apple. Theor Appl Genet 96:1069-1076.

Marchetti E, Civolani S, Leis M, Chicca M, Tjallingii WF, Pasqualini E, Baronio P (2009). Tissue location of resistance in apple to the rosy apple aphid established by electrical penetration graphs. Bull Insectology, 62:203-208.

Miñarro M, Dapena E (2007). Resistance of apple cultivars to *Dysaphis plantaginea* (Hemiptera: *Aphididae*): role of tree phenology in infestation avoidance. Environ Entomol, 36:206-1211.

Pagliarani G, Dapena E, Miñarro M, Denancé C, Lespinasse Y, Rat-Morris E et al. (2016). Fine mapping of the rosy apple aphid resistance locus *Dp-fl* on linkage group 8 of the apple cultivar 'Florina'. Tree Genet Genom, 12:1-12.

Parisi L, Gros C, Combe F, Parveaud CE, Gomez C, Brun L. (2013). Impact of a cultivar mixture on scab, powdery mildew and rosy aphid in an organic apple orchard. Crop Protection, 43, 207-212.

Rat-Morris E (1994) Analyse des relations entre *Dysaphis plantaginea* Pass. (Insecte Homoptère) et sa plante hôte *Malus* x *domestica* Borkh.: étude de la résistance du cultivar Florina. PhD thesis, Faculté des Sciences, Université de Tours, France (114 p).

Rat-Morris E & Lespinasse Y (1995). Pommier: la résistance au puceron cendré associée à la résistance à la tavelure. Une nouvelle source de résistance mise en évidence chez le cultivar Florina. Phytoma-La Défense des Végétaux, 471:15-17.

Young ND, Tanksley SD (1989). Restriction fragment length polymorphism maps and the concept of graphical genotypes. TAG Theor Appl Genet, 77:95-101.

Wang A, Aldwinckle H, Forsline P, Main D, Fazio G, Brown S & Xu K (2012). EST contig-based SSR linkage maps for *Malus* × *domestica* cv Royal Gala and an apple scab resistant accession of *M. sieversii*, the progenitor species of domestic apple. Molec breed, 29:379-397.

General conclusions

The principal aim of this work was to identify the gene conferring resistance against the rosy apple aphid for developing new apple varieties resistant to *D. plantaginea*.

This work started from the mapped region of about 330 Kb in Florina resistant cultivar (Pagliarani et al. 2016). The availability of the Florina BAC library, at the University of Bologna – Department of Agricultural Sciences, allowed to start a BAC library screening in order to find the minimum tiling path covering the *Dp-fl* region. Five BAC clones fully covering the 'resistant' locus of Florina were identified. Through the development of new polymorphic markers it was possible to better define the position of the *Dp-fl* locus between two markers (SSR_377 and SSR_4) in a region including three BAC clones (63M14, 88H21 and 47A15) that were finally chosen for the sequencing. Through the assembling of the three BAC clones resulted a single contig of about 279 Kb, the region was then completely annotated for the identification of candidate genes.

New polymorphic markers were developed within the new region and new segregating progenies, having Florina as resistant parent, were taken into consideration. Finally, the *Dp-fl* resistance locus resulted mapped in a specific region of about 56 Kb where seven coding sequences were predicted to encode for genes. Unexpectedly, none of them were coding for known resistance proteins, including genes of the NBS-LRR family that were already reported as involved in the aphid resistance mechanisms of tomato and melon (Rossi et al. 1998; Dogimont et al. 2008). Therefore, a different resistance mechanism for RAA was hypothesized in Florina. Among the seven candidate genes, four had a biological function that could be related to the RAA attack. Thanks to the mapping it was possible to identify a single candidate gene. This gene encodes for a protein belonging to the C2 calcium/lipid-binding plant phosphoribosyltransferase family. A member of this family in *A. thaliana*, the Quirky (QKY) protein, is involved in cell-to-cell communications that control cell patterning, organ shape and development. In *A. thaliana* It has been shown that *qky*-mutants have twisted petals and leaves and an aberrant floral phyllotaxis (Fulton et al. 2009). For that reason, it was hypothesized that the QKY gene would have been correlated to the symptoms, such as leaf deformations, due to the salivary secretion of RAA. Furthermore, a recent study co-locates QKY protein at the plasmodesmata with a LRR-receptor-like kinase protein called STRUBBELIG, but independently from one another. This positioning seems to play a central role in cell-to-cell communication and in the growth of plant cells (Vaddepalli et al. 2014). In vascular plants, such as apple, plasmodesmata allow the movement of sap through the sieve element. During probing and feeding, the aphids secrete saliva directly into the phloem (Miles 1999), causing leaf and shoot deformations that could be correlated with the QKY gene. Indeed, a modification occurring at the plasmodesmata level could negatively or positively affect

phloem transport and hence aphid feeding. For all these reasons, in this work has been deeply investigated the Quirky gene to better understand its role in the response to RAA.

First analysis performed was the comparison among the sequences obtained from both resistant and susceptible Quirky alleles of Florina compared with the sequence of GDDH13. In the gene sequence, no particular differences were highlighted, but looking at the promoter sequence a region of 115 bp before the start codon was completely lacking in both resistant and susceptible alleles of Florina, while it was present in GDDH13. A second analysis was performed by real time PCR. Analyzing the gene expression, it was possible to deduce that the Quirky gene was expressed only in Golden and after RAA infestation. This data could explain the difference among the promoter sequence, that can influence the expression differences. Finally, to validate the Quirky gene function, a genetic transformation of both Gala and Florina was started, but the first transformed plants have still to be analyzed. Considering the obtained results, it is possible to conclude that the Quirky gene could be more in relation to susceptibility than to the resistance. Indeed, the expression levels are higher in Golden Delicious than in Florina after 72 hours from the infestation. Furthermore, all the recombinant individuals that define the locus of the *Dp-fl* show a susceptible phenotype. A recent study co-locates QKY protein at the plasmodesmata with an LRR-receptor-like kinase protein called STRUBBELIG (Vaddepalli et al. 2014) and a further gene with a similar structure could be hypothesized to explain RAA resistance. Furthermore, a NBS- LRR receptor protein has been already reported to be involved in the aphid resistance mechanisms of tomato and melon (Rossi et al. 1998; Dogimont et al. 2008).

During the fine-mapping process two individuals belonging to the Perico x Florina population (PF_P001 and PF_X-9504-33) exhibits genotype/phenotype incongruences (GPI) but these two individuals were no more available to confirm their phenotype. Two new progenies having *Malus floribunda* #821 as resistant parents were added to the work to further confirm the identification of the resistance gene.

By the phenotypic and genotypic analysis of the *Malus floribunda* progenies, the presence of a further resistance gene, that works together with the Quirky gene, cannot be excluded. Despite the whole genomic region of *M. floribunda* #821 has been shown to be inherited by Florina, the *Dp-Mflo* locus was mapped about 800 Kb apart from the *Dp-fl* locus. This could suggest the presence of a second gene involved in RAA resistance. Thus, an alternative and more complex hypothesis was formulated: two nearby RAA resistance genes in the bottom part of LG8 was postulated. The Quirky gene is putatively involved in susceptibility; while a second gene identified *in M. floribunda* #821 progenies and located downstream the *Dp-fl* locus, could be hypothesized to be involved in the resistance against the *D. plataginea*. The identification in GDDH13 of five genes putatively involved in defence response to RAA, can support this hypothesis. Anyway, this analysis on the GDDH13 genome is only a starting point because the complete sequence of the *M. floribunda Dp-Mflo* region is required.

The presence of a two closely-linked genes involved in resistance is rather common since R genes are known to cluster where the genes the AKR, TRR and RAP1 genes of *M. truncatula* confers resistance to three distinct aphid species and are located within a region of about 40 cM (Kingler et al. 2009; Steward et al. 2009)

In conclusion, additional phenotyping and genotyping are requested to finely locate the new gene.