

Diss. ETH Nr. 13793

**MAP-BASED CLONING OF RESISTANCE GENE
HOMOLOGUES IN THE *Vf* -REGION OF THE APPLE (*Malus sp.*)**

A dissertation submitted to the

SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZÜRICH

For the degree of
Doctor of Natural Sciences

Presented by

Andrea Patocchi

Dipl. sc. nat. ETH
Born March 19th, 1971
Citizen of Switzerland

Accepted on the recommendation of
Prof. Dr. K. Apel, examiner
Dr. C. Gessler, co-examiner
Dr. L. Gianfranceschi, co-examiner

2000

INDEX

| | |
|--|----|
| ABSTRACT | 1 |
| RIASSUNTO | 3 |
| 1. INTRODUCTION | 5 |
| 2. MATERIAL AND METHODS | 22 |
| 2.1 Fine mapping of <i>Vf</i> | 22 |
| 2.1.1 Plant material | 22 |
| 2.1.2 DNA extraction for fine mapping | 22 |
| 2.1.3 Fine mapping of the <i>Vf</i> linked markers M18 and AL07 | 22 |
| 2.2 Physical mapping of the <i>Vf</i>-region | 23 |
| 2.2.1 Preparation and embedding of intact nuclei in agarose plugs | 23 |
| 2.2.2 Digestion of the embedded HMW-DNA | 24 |
| 2.2.3 Pulsed field gel electrophoresis, blotting and hybridisation | 24 |
| 2.3 Chromosome walking | 26 |
| 2.3.1 The apple BAC library | 26 |
| 2.3.2 Fine mapping of AL07 and AM19 | 27 |
| 2.3.3 Screening of the BAC library and analysis of putative positive clones | 27 |
| 2.3.4 BAC miniprep and insert-end isolation | 28 |
| 2.3.5 Determination of contig ends during assembly | 29 |
| 2.3.6 Mapping of insert end-sequences and preparation of probes for the succeeding step | 30 |
| 2.4 Identification of resistance gene homologues | 31 |
| 2.4.1 Starting material | 31 |
| 2.4.2 Characterisation of positive cDNAs | 32 |
| 2.4.3 Development of <i>fv</i> primers | 32 |
| 2.4.4 Cloning of PCR products | 33 |
| 2.4.5 RT-PCR | 33 |
| 2.4.6 RACE | 34 |
| 2.4.7 ORFs analysis | 34 |

| | |
|--|----|
| 3. RESULTS | 35 |
| 3.1 Fine mapping of <i>Vf</i> | 35 |
| 3.2 Physical mapping of the <i>Vf</i>-region | 38 |
| 3.2.1 Extraction of high molecular weight DNA (HMW-DNA) | 38 |
| 3.2.2 Southern blot | 38 |
| 3.3 Chromosome walking | 40 |
| 3.3.1 Fine mapping of AL07 and AM19 | 40 |
| 3.3.2 Identification of positive clones | 41 |
| 3.3.3 End rescue | 41 |
| 3.3.4 The chromosome walking towards the <i>Vf</i> gene starting from M18 | 42 |
| 3.3.5 Identification of candidate <i>Vf</i> -containing clones | 46 |
| 3.4 Identification of resistance gene homologues | 47 |
| 3.4.1 Identification of cDNAs homologues to the tomato <i>Cf</i> resistance genes | 47 |
| 3.4.2 Development of specific primers for the <i>HcrVf</i> gene family (fv primers) and identification of <i>HcrVf1</i> in the contig | 49 |
| 3.4.3 Identification of <i>HcrVf2</i> , <i>HcrVf3</i> and <i>HcrVf4</i> | 53 |
| 3.4.4 RACE | 57 |
| 3.4.5 Analysis and comparison of <i>HcrVf1</i> , <i>HcrVf2</i> and <i>HcrVf3</i> | 60 |
| 3.4.6 Hybridisation of <i>HcrVf2</i> to the <i>EcoRI</i> digested BAC clones of the <i>Vf</i> contig | 62 |
| 4. DISCUSSION | 64 |
| 4.1 Fine mapping of <i>Vf</i> | 64 |
| 4.2 Physical mapping of the <i>Vf</i>-region | 66 |
| 4.3 Chromosome walking | 68 |
| 4.4 Identification of resistance gene homologues in the <i>Vf</i>-region | 70 |
| 4.4.1 Identification of <i>HcrVf1</i> , <i>HcrVf2</i> and <i>HcrVf3</i> | 70 |
| 4.4.2 Homologies of the <i>HcrVfs</i> to the <i>Cf</i> resistance genes (and other R-genes) | 73 |
| 5. OUTLOOK | 76 |
| 6. CONCLUSIONS | 78 |
| 7. REFERENCES | 79 |
| CURRICULUM VITAE | 86 |
| RINGRAZIAMENTI | 87 |

ABSTRACT

Scab, caused by the fungal pathogen *Venturia inaequalis*, is the most common disease of the cultivated apple (*Malus × domestica* Borkh.). Resistance against scab is found in some wild *Malus* species. *Vf* resistance of the small-fruited apple species *Malus floribunda* has been introgressed into several apple varieties during many years of breeding. It segregates as a monogenic trait and tightly linked genetic markers (M18, AM19 and AL07) have been found in recent years.

A method that has been shown to be successful for the isolation of genes of which only the phenotype and the map position are known, is map-based cloning. To check the feasibility of such a method with the available material (*Vf* markers and apple BAC library) an estimation of the physical distance separating the *Vf* flanking markers M18 and AL07 was made. The two markers hybridised to a common 870 kb *NotI* restriction fragment indicating that this is the maximal physical distance between them. The two markers were tested on 1179 progeny plants of three crosses and M18 and AL07 were positioned at 0.2 cM and 1.1 cM from *Vf* respectively, for a total genetic distance between them of 1.3 cM. These results indicated that the ratio between physical and genetic distance in the *Vf*-region is about 670 kb/cM. This led us to the conclusion that a chromosome-walking project using the available apple BAC library was feasible. To identify the best starting points for the chromosome walk on the opposite site of M18, 892 more progeny plants of different crosses were analysed with the three molecular markers (M18, AL07 and AM19) to find recombinants that allowed us to determine which of the two co-segregating markers, AL07 and AM19, mapped closer to *Vf*. Two recombinant plants indicated that AM19 was closest to *Vf*. The chromosome walking was initiated simultaneously from M18 and AM19. Thirteen BAC clones spanning the region between the two markers were found in nine chromosome walking steps. The size of the resulting contig is approximately 550 kb. In order to reduce the region where *Vf* lies, the 29 recombinant plants of the *Vf*-region were tested with the polymorphic markers produced during the chromosome walking. This way, it was possible to restrict the possible location of the *Vf* gene to a chromosomal region that could be covered by only five BAC clones spanning approximately 350 kb.

The screening of a cDNA library derived from Florina leaves inoculated with a conidial suspension of *V. inaequalis*, using as a probe the inserts of the five BAC spanning the *Vf*-region, allowed the identification of a cluster of genes coding for receptor-like proteins. Those genes show high homology with members of the *Cf* resistance gene family of tomato. The predicted amino acid sequence contains an extracellular leucine-rich repeat domain and a putative transmembrane domain. Those genes have been named *HcrVf* (Homologues of *Cladosporium fulvum* resistance genes of the *Vf*-region). The gene cluster is present on the chromosomes in coupling with the resistance as well as on the other one in repulsion (Florina is heterozygous for the *Vf*-region). The cluster in cis with the *Vf* gene is composed of 5 to 14 *HcrVf* genes. Three of them were sequenced completely. For two of the three genes analysed it was possible to establish that they are constitutively expressed in leaves and that they are differentially expressed in *Vf* and non-*Vf* plants. From what has been presented here it is therefore probable that one of these genes is the *Vf* resistance gene. However, this last point still has to be demonstrated.

RIASSUNTO

La malattia più comune del melo coltivato (*Malus × domestica* Borkh) è la ticchiolatura causata dal patogeno fungino *Venturia inaequalis*. Fonti di resistenza alla ticchiolatura si trovano in alcune specie selvatiche di melo. La resistenza *Vf*, derivata da *Malus floribunda* una specie di melo con frutti di piccole dimensioni, è stata introdotta mediante programmi di miglioramento durati molti anni in diverse varietà di melo. La resistenza *Vf* sembra essere controllata da un singolo gene per il quale negli ultimi anni sono stati trovati alcuni marcatori strettamente associati (M18, AM19 e AL07).

Un metodo che si è dimostrato efficace per clonare geni di cui sia nota solo la loro posizione di mappa è il map-based cloning. Per verificare la possibilità di applicare questa strategia alla clonazione di *Vf*, utilizzando le informazioni e il materiale a disposizione (marcatori associati a *Vf* e genoteca BAC di melo) è stata effettuata una stima della distanza fisica esistente tra i marcatori molecolari M18 e AL07 fiancheggiando il gene. Si è potuto così verificare che entrambi i marcatori ibridano ad uno stesso frammento di restrizione *NotI* di 870 kb, dimostrando che questa è anche la massima distanza fisica che li separa. Questi due marcatori sono stati successivamente saggiati su 1179 semenzali provenienti da tre incroci diversi ed è risultato che M18 e AL07 distano rispettivamente 0.2 cM e 1.1 cM da *Vf*. Questi risultati indicano che il rapporto tra distanza fisica e distanza genetica nella regione di *Vf* è di circa 670 kb/cM, provando anche che in questa regione cromosomica non vi è alcuna soppressione della ricombinazione. Per questo motivo si è concluso che un chromosome-walking fosse fattibile. Per stabilire quale dei due marcatori co-segreganti, AL07 e AM19, fosse più vicino a *Vf* e di conseguenza costituisse il punto migliore per cominciare il chromosome-walking dalla parte opposta a M18, è stato necessario analizzare ulteriori 892 piante provenienti da altri incroci. L'identificazione di due piante in cui era avvenuta una ricombinazione tra i due marcatori ha provato che AM19 è più vicino a *Vf* di AL07. Si è così deciso di iniziare contemporaneamente il chromosome-walking da una parte partendo da M18 e dall'altra partendo da AM19. Nove screening della genoteca BAC sono stati necessari per identificare i 13 cloni che coprono la regione tra i due marcatori. Le dimensioni del contig (serie di cloni contigui) così ottenuto è di circa 550 kb. Per ridurre

ulteriormente le dimensioni della regione in cui è localizzato *Vf*, le 29 piante recombinanti nella regione di *Vf*, individuate tra le 2071 piante analizzate in precedenza, sono state saggiate con tutti i marcatori polimorfi prodotti durante il chromosome walk. In questo modo è stato possibile restringere la regione contenente *Vf* ad un intervallo di circa 350 kb, che può essere coperto con un minimo di cinque cloni BAC.

Lo screening di una libreria cDNA, ottenuta da foglie di Florina inoculate con una sospensione conidiale di *V. inaequalis*, usando come sonda gli inserti dei cinque cloni BAC ha permesso di identificare un cluster di geni con struttura simile a quella di geni codificanti per proteine con funzioni recettoriali. I geni identificati presentano infatti molte omologie con i geni *Cf* di pomodoro la cui sequenza aminoacidica contiene un dominio extracellulare ricco in leucine (LRR, Leucine Repeat) e un putativo dominio transmembrana. I nuovi geni identificati sono stati denominati *HcrVf* (Homologues of *Cladosporium fulvum* resistance genes of the *Vf*-region). L'aver utilizzato una cultivar come Florina, eterozigote per il gene *Vf*, per la costruzione della libreria BAC, ha permesso di provare che, sia sul cromosoma in cis con la resistenza, e quindi parzialmente derivante da *Malus floribunda*, sia sul cromosoma omologo, derivante da *Malus × domestica*, fosse presente un cluster di geni *HcrVf*.

Il cluster *HcrVf* in cis con la resistenza è stato analizzato più in dettaglio e si è visto essere composto da un numero variabile tra 5 e 14 membri della famiglia di geni *HcrVf*. Tre di essi sono stati completamente sequenziati e per due si è potuto dimostrare che sono espressi costitutivamente nelle foglie. Dato che gli *HcrVf* identificati sono simili a geni di resistenza già clonati in altre specie e visto che alcuni sono espressi differenzialmente tra cultivar con e senza *Vf*, c'è una buona probabilità che uno di essi sia il gene di resistenza *Vf*, sebbene questo ultimo punto debba ancora essere dimostrato.

1. INTRODUCTION

Apple scab, caused by the fungal pathogen *Venturia inaequalis* (Cooke.) is the most important disease spread over all apple growing areas with high spring and summer rainfall.

V. inaequalis overwinters mainly as pseudothecia in apple leaf litter but also, less frequently, as dormant mycelium pustules on shoots and budscales. In spring ascospores are ejected from the pseudothecia that have been wet by rain at daylight and are wind-dispersed, while the conidia, produced asexually on the pustules, are splash-dispersed. Ascospores, and in a small percentage conidia, compose the primary inoculum. To germinate and successfully produce an infection, ascospores and conidia need a water film and have to land on young apple leaves or fruits of a compatible cultivar. The germinated spores penetrate the cuticle by the germ tube. The hyphae grow between cuticle and epidermal cells. Depending on the temperature, stroma grows and will produce conidiophores and conidia emerging from the cuticle. The asexually produced conidia will be dispersed by rain splashing and initiate the secondary infections. After infected leaves fall from the trees in autumn, the mycelium invades the interior of the leaf, antheridium and ascogonium develop, and if they are of opposite mating types, pseudothecia develop.

Depending on the weather, up to eight asexual cycles might be completed in one growing season (MacHardy 1996). If young fruits are infected, they will develop abnormally with a consequent total loss of commercial value. To prevent economic losses farmers treat their apple orchards with fungicide up to 15 times each year. The widespread concern today over chemical residues in food and ecosystems, and in the past the elimination of risk of an epidemic, has always stimulated breeders to look for apple selections with natural resistance to scab and mildew.

Pioneer work to look for resistance sources was done by Aderhold (1902), Wallace (1913) and Rudloff and Schmidt (1934, 1935). Because at that time no commercial variety possessed satisfactory resistance to scab, they began to look for resistance in related small-fruited *Malus* species. Rudloff and Schmidt found that *M. atropurpurea*, *M. micromalus*, *M. spectabilis* and *M. spectabilis* Kaido were free of scab or showed only very light infection during field inoculation tests. Also Crandall (1926) collected

different small-fruited *Malus* species and among them Hough (1944) found 6 selections of Asiatic *Malus* species (*M. atrosanguinea* 804, *M. floribunda* 821, *M. prunifolia* 19651, *M. ringo*, *M. turingo* and *M. zumi*) that exhibited "field immunity" to scab.

The most successful breeding program for scab resistance was the PRI program initiated in 1945 by Hough and Shay in the USA (Crosby *et al.* 1992). Thanks to extremely favourable weather conditions for *V. inaequalis* infections in spring 1943 that led to defoliation of unsprayed susceptible trees, Hough and Shay recognised in a full sib cross between two F₁ selections of Crandall's cross *M. floribunda* 821 × Rome Beauty performed in 1914, a nearly 1:1 segregation for scab resistance. Besides the discovery that the resistance was very probably inherited as a monogenic trait (as was confirmed with additional crosses), their really big breakthrough was the discovery of two resistant F₂ selections (26829-2-2 and 26830-2-2) with the highest quality and largest fruits available at that time. In fact, the progenies of a cross between small-fruited *Malus* species with commercial apple cultivars result in almost 100 % of plants that produce small fruits, making the breeding of commercially interesting resistant cultivars difficult. The two resistant F₂ selections were the source of scab resistance used in the PRI program that released the first apple scab resistant cultivar Prima in 1970. Dayton and Williams (1968) called this resistance gene *Vf* (*V* = *Venturia*, *f* = *floribunda*) because it originally derives from *M. floribunda* 821.

In the following years five other loci conferring monogenic qualitative resistance were identified: *Vm* from *M. micromalus*, *Vr* from *M. pumila* R12740-7A, *Vbj* from *M. baccata jackii*, *Vb* from Hansen's *baccata* #2 and *Va* from Antonovka PI 172623 (Williams and Kuc 1969).

Today, several apple selections with high scab resistance are available. Most of them carry only one major known resistance gene that is usually *Vf*. Janick *et al.* (1996) reported the presence of the *Vf* gene in 70 different scab-resistant cultivars.

A major problem of those new selections, already reported by Alston (1989), is the low acceptance by the consumers due to the lack of important quality components such as taste, storage quality, etc. The development of new resistant selections with the desired high fruit quality is attempted by crossing resistant selections with commercial cultivars having these qualities. Because backcrosses are impossible, due to the self-incompatibility of the apple, at each new breeding step a different parental

plant has to be used. This, combined with the high level of heterozygosity of the apple, leads to the production of progeny plants with completely different characters than those present in the parents. It is easy to imagine that thousands of seedlings have to be produced and kept in the field for years to have a minimal chance to select a plant fulfilling the requests for yield, precocity of yield, resistance to foliar diseases, fruit size, flavour and colour, tree habit, skin finish and storage quality. In fact, even if the screening for scab resistance can be accomplished in the greenhouse at the seedling stage, shortly after germination, the evaluation of many other characteristics is delayed, due to long juvenile periods (4 to 5 years). Typical examples are yield and fruit quality that can be evaluated only after 5 to 6 years. [To produce the new cultivar Ariwa it was necessary to screen about 30,000 seedlings (Kellerhals, personal communication)].

To reduce the number of plants that have to be kept in the field waiting for the expression of the phenotype, the genotype of seedlings can be analysed with molecular markers associated with the gene of interest. This strategy is called Marker Assisted Selection (MAS). To increase selection efficiency, molecular markers tightly associated with the gene or genes of interest have to be found. Recently, molecular markers for scab resistance genes *Vf* (Gianfranceschi *et al.* 1996, Tartarini 1999), *Vm* (Cheng *et al.* 1998), and *Vbj* (Gygax personal communication), for mildew resistance genes *Pl1* (Markussen *et al.* 1997) and *Pl2* (Seglias and Gessler 1997a) and for aphid resistance *Sd₁* (Roche *et al.* 1997) have been published. The progenies of a cross between plants carrying these genes can be tested for the presence of the correct marker allele at the seedling stage, and only the plants carrying the desired allele combinations will be evaluated in the field and selected for other characteristics for which no molecular markers are available.

It has been proposed (Crute and Pink 1996) that for the breeding of new cultivars with durable resistance several resistance genes, with different resistance mechanisms, should be combined in the same plant. This strategy is called gene pyramiding. The absolute need for pyramiding several monogenic resistance genes in apple cultivars has been demonstrated by the recent breakdown of the *Vf* resistance (Parisi *et al.* 1993, B naouf and Parisi 2000). MAS is fundamental for gene pyramiding, in fact, in the absence of races of *V. inaequalis* able to overcome simultaneously more than one resistance gene it would be impossible, without the use of molecular markers, to distinguish between seedlings carrying one or more resistance genes.

To develop alternative breeding strategies to breed for durable resistance against apple scab, it is necessary to have a better knowledge of the source of resistance and of the life cycle of the pathogen, but also more information about the interaction between apple and *V. inaequalis* is required

Up to the present the progeny plants of a cross were usually referred to as either resistant or susceptible to apple scab but the separation of these two phenotypes is not easy. Resistance or susceptibility depends on the host and on the pathogen genotype and on the age of the leaves, due to the presence of ontogenic resistance in the apple. Chevalier *et al.* (1991) classified the different macroscopic symptoms observed in *Vf*-resistant plants into coded classes, completing what was previously proposed by Hough and co-workers (Hough *et al.* 1953). The histological and scanning-electron microscopy (SEM) analysis led to the following proposed classification:

- Class 1: small pits appearing 4 to 6 days after inoculation, which do not enlarge with time. By SEM it was observed that the "pin-point" symptom is a depression of 100-500 μm in diameter where the epidermal cells have collapsed. Class 1 is the typical hypersensitive reaction.
- Class 2: symptoms appear between 7 and 10 days after inoculation in the form of chlorotic lesions with irregular edges, with the lesion centre sometimes slightly necrotic. In any case no conidia are produced. Histological observations indicated that very often the upper epidermis is partially collapsed, cells are often plasmolysed and the vacuoles enclose dense granulation.
- Class 3: symptoms appear 7 to 10 days after inoculation. This class has been divided into 2 sub-classes. Class 3a includes necrotic lesions and some chlorotic ones with occasionally very slight sporulation. Usually, secondary stroma become extensive and epidermal cells collapse over a large area. Class 3b includes clearly sporulating chlorotic and necrotic lesions. Epidermal cells of chlorotic lesions do not show particular changes while epidermal cells of the necrotic lesions are destroyed.
- Class 4: symptoms appear 10 to 12 days after infection. Stroma forms a dense subcuticular network. The leaves are covered by thousands of conidophores and the abundance of annelides is a sign of their activity. This class is characterized by the absence of visible histological modifications of the different leaf tissues.

As previously reported the *Vf* gene is thought to be inherited as a monogenic character since in crosses between a heterozygous *Vf* plant and a susceptible cultivar a 1:1 ratio

of resistant / susceptible progeny plants is obtained. However this 1:1 ratio is obtained only if plants classified as 0 (no symptoms), 1, 2, 3a and 3b are considered as resistant and plants in class 4 are considered as susceptible and the test conditions (inoculum density and climatic conditions) are appropriate. Indications of the correctness of this separation can be seen in the histological observation of Chevalier that remarked a histological reaction of the seedling only in the classes 1 to 3b. The final confirmation of this separation was given by the use of the *Vf* molecular markers that indicate the presence of the allele in coupling with *Vf* in the classes 0 to 3b (resistant) and their absence in the class 4 (susceptible) type reaction.

Interestingly, in contrast to what was reported by Chevalier *et al.* (1991), the class 1 type reaction is normally found only in *M. floribunda* 821, the source of the *Vf* resistance, and not in any of the *Vf*-resistant cultivars produced so far. In fact, already the F2 selections 26829-2-2 and 26830-2, derived from Crandall's cross *M. floribunda* 821 × Rome Beauty, express a class 2 type resistance reaction, indicating that "something" was lost very early during the breeding program. Bénaouf and Parisi (2000) proved the presence of a second resistance gene in *M. floribunda* 821, which they called *Vfh* (*h* = hypersensitivity), responsible for the class 1 type of reaction.

Gessler *et al.* (1989) postulated that the genetic factors conditioning class 0 (no symptoms), 2, 3a, 3b and 4 reactions are modifier genes, present in any *Malus* cultivar resistant or susceptible to scab. Those genes have an identifiable effect on the pathogenesis only in combination with the main gene, *Vf*. Gessler's "minimal" model with two loci, each with two alleles, is already sufficient to explain this variation of symptoms. One allele of each locus would have an additive effect "+" while the other "-" would have no effect. In a cross with two heterozygous parents (+/- +/-) 5 different combinations of modifier alleles are possible. Assuming that the homozygous combination of "+" in one of the two loci has the same effect as the heterozygous combination in both loci, only the total number of alleles "+" will be decisive for a stronger or weaker resistance. Plants carrying *Vf* and the 4 alleles "+" will express a resistance reaction of type 0 (no symptoms), plants with 3 alleles "+" the class 2, plants with 2 alleles "+" the class 3a, plants with 1 allele "+" the class 3b and plants carrying *Vf* with only "-" alleles will be classified with plants without *Vf* as class 4.

Support to this theory was given by QTL analysis (Seglias 1997b). The author reported the identification of different loci "modifying" the expression of scab resistance that were effective only in the presence of the *Vf* gene (per definition modifiers), and of some other loci with an additive effect that were effective also in the absence of *Vf*. However, those findings could not be definitively statistically confirmed.

Supplementary support to this theory was provided by the discovery of the race 6 and by experiments published by Parisi *et al.* (1993). [Races of scab are defined by their virulence to certain breeding stocks (reviewed by Boone 1971)]. Race 6 was found for the first time at Ahrensburg (Germany). It is defined as the group of scab pathotype that can incite class 4 type symptoms on certain *Vf* plants but not on *M. floribunda* 821. The experiments performed by Parisi *et al.* (1993) consisted in inoculating the progenies of a cross between *Vf*-heterozygous and a susceptible cultivar with isolates of the race 6 and an inoculum without race 6 ("Angers inoculum"). As expected, when the progenies were inoculated with the "Angers inoculum" 50% of the plants were resistant, showing a range of symptoms distributed over all resistant reaction classes (0 to 3b). In contrast, when the race 6 inoculum was used, all the plants exhibited a class 4 type of reaction. Since the *Vf* gene was overcome and the modifiers did not play a role anymore, therefore no other classes were found. However, in this case also, the classification of all plants in the same class 4 type of reaction could be an artefact, due the lack of a precise differentiation in classes within the susceptible plants.

Benaouf and Parisi (1997) identified a seventh race, namely race 7 that is able to infect *M. floribunda* 821 but not Golden Delicious. By inoculating Golden Delicious with race 7 Benaouf and Parisi (2000) demonstrated that such "susceptible" cultivars carry a resistance gene that they called *Vg*.

The presence of "ephemeral" resistance genes (resistance genes overcome by a high percentage of the pathogen population) in so-called "susceptible" cultivars was already demonstrated by Sierotzki *et al.* (1994) and Koch (1998). They found that some isolates, collected from susceptible cultivars, were able to induce characteristic class 4 type symptoms mostly on the cultivar they originated from. In contrast on other "susceptible" cultivars resistance reactions were observed. Moreover Koch *et al.* (2000) reported that some isolates were able to produce class 4 type symptoms on several "susceptible" cultivars, demonstrating that several virulence alleles were

present in the same isolate. Those results should not be forgotten while developing strategies to prevent or reduce resistance breakdown. The results of Sierotzki *et al.* (1994) and Koch (1998) indicated the presence of unknown resistance genes in various "susceptible" cultivars and of avirulence/virulence alleles in the fungus conditioning compatible and incompatible reactions. Those points of evidence indicate that if the classification of *V. inaequalis* pathotypes would take into consideration also the "ephemeral" resistance genes normally present in *Malus* × *domestica*, the number of scab races should be much higher than seven. From those studies, which included a total of 13 "susceptible" cultivars, the authors could for example identify 13 new differential races.

In his work, Koch *et al.* (2000) could also prove that the frequencies of complex races (pathogens carrying more than one virulence) was lower if the isolates were isolated from an apple monoculture than if they were isolated from an orchard containing a cultivar mixture. In fact, in one orchard composed of 26 different cultivars, isolates carrying up to eight different virulences were found. This can probably indicate that in absence of selection pressure, isolates carrying a single virulence gene are favoured probably because complex races tend to have a lower fitness.

The possibility of finding pathotypes in the population of apple scab that can specifically attack only certain cultivars, indicate that Flor's "gene-for-gene" theory (Flor 1971) is valid and holds for the interaction between *V. inaequalis* and *Malus*, too. The theory assumes that for each resistance allele in the plant there exists a corresponding pathogen avirulence (Avr) allele. The interaction between the proteins coded by the resistance allele and its corresponding avirulence allele gives an incompatible reaction so that the pathogen cannot infect the plant. On the contrary, if one of the two factors is missing there is a compatible interaction and the pathogen can develop and infect the plant. The resistance gene plays a role in recognising the pathogen, and for this reason resistance genes are also called "recognition genes" or "R-genes". It is assumed, and in part also demonstrated, that after recognition of the corresponding Avr allele product, the R-gene product, directly or through the interaction with other proteins, activates different cascades of defence-reactions. The genes involved in the defence mechanism are the so-called "defence-response" genes and include phytoalexins, hydrolytic enzymes and PR-Proteins (pathogenesis-related proteins). However none of these proteins has been detected in apples up to now.

A method that proved to be successful for the isolation of genes for which only the phenotype and the map position are known (as is the case of *Vf*) is map-based cloning (Paterson and Wing 1993, Wing *et al.* 1994). This strategy was used for cloning several resistance genes like: *Pto* (Martin *et al.* 1993), *RPS2* (Bent *et al.* 1994), *Xa21* (Song *et al.* 1995), *RPM1* (Grant *et al.* 1995), *Cf-2* (Dixon *et al.* 1996), *Mlo* (Büschges *et al.* 1997), *Hsl^{pro-1}* (Cai *et al.* 1997), *RPP5* (Parker *et al.* 1997), *Mi* (Milligan *et al.* 1998) and *Pib* (Wang *et al.* 1999). This approach could be divided into three main steps:

1. The gene of interest has to be mapped in a large segregating population. The region containing the gene has to be saturated with markers so that tightly linked markers are identified. For this purpose different kinds of markers could be used:

- AFLP (Amplified Fragment Length Polymorphism) (Zebeau 1992)
- RAPD (Random Amplified Polymorphic DNAs) (Williams *et al.* 1994)
- RDA (Representational Difference Analysis) (Lisitsyn *et al.* 1993)
- RFLP (Restriction Fragment Length Polymorphism) (Burr *et al.* 1994)
- SSR (Simple Sequence Repeats) (Morgante *et al.* 1993)
- SNP (Single Nucleotide Polymorphism) (Wang *et al.* 1998)
- CAPS (Cleaved Amplified Polymorphic Sequence) (Konieczny *et al.* 1993)
- SCAR (Sequence Characterized Amplified Region) (Konieczny *et al.* 1993)

For the *Vf* resistance gene several RAPD markers were found (Yang *et al.* 1994, Koller *et al.* 1994, Tartarini *et al.* 1996, Gardiner *et al.* 1996). The three RAPD markers closest to the gene, M18, AL07 and AM19 were cloned and transformed in CAPS or SCAR markers (Gianfranceschi *et al.* 1996, Tartarini *et al.* 1999).

Although different research groups mapped those markers in the same chromosomal position and in the same order, the position of the *Vf* gene has always been debated. Gessler *et al.* (1995) and Tartarini *et al.* (1999) placed *Vf* between the molecular markers AL07 (cosegregating with AM19) and M18. Hemmat *et al.* (1998) and Gardiner *et al.* (1996) mapped *Vf* near AL07 and M18 respectively, but on the opposite site of the one proposed by Gessler *et al.* (1995) and Tartarini *et al.* (1999) (Fig. 3B). Tartarini *et al.* (1999) mapped *Vf* at a distance of 0.5 cM from M18, while AL07 and AM19 are both located at 0.9 cM from *Vf* on the opposite side.

2. The second step of a map-based cloning project is the estimation of the physical distance separating the two closest markers bracketing the gene (physical mapping).

This check is important, because it has been reported that markers that were genetically close were in fact physically very far apart, due to a suppression of recombination occurring at that specific chromosomal region (e.g. Ganai *et al.* 1989, Young and Tanksley 1989). It is therefore very important to verify the absence of suppression of recombination, especially when the region of interest has been introgressed from a wild species, as it is the case of *Vf*.

Once the maximal physical distance between the flanking markers has been determined, it is then possible to decide whether new markers, closer to the gene, have to be found or if the physical distance between the available markers could be covered using a large-insert genomic library. It is also possible to decide which kind of library is best suited to the purpose. Currently several large-insert genomic libraries can be constructed:

| | Average insert size | References |
|--|---------------------|---------------------------------|
| - cosmids | 40kb | (Meyerowitz <i>et al.</i> 1980) |
| - P1 phage | 70-20kb | (Pierce <i>et al.</i> 1992) |
| - BAC (Bacterial Artificial Chromosome) | 150kb | (Shizuya <i>et al.</i> 1992) |
| - PAC (P1-derived Artificial Chromosome) | 150kb | (Ioannou <i>et al.</i> 1994) |
| - YAC (Yeast Artificial Chromosome) | 500kb | (Burke <i>et al.</i> 1987) |

Vinatzer *et al.* (1998) published the construction of a BAC library with inserts of Florina, an apple cultivar heterozygous for the *Vf* gene, obtaining an average insert size of 120 kb. Since the library is composed by 36,864 clones and the size of the haploid *Malus × domestica* genome is about 770 Mbp (Arumuganathan and Earle 1991), the library represents approximately 5x the haploid genome equivalents. Once the library of choice has been constructed or is available, the next step of the map-based cloning is to identify the clones carrying the gene of interest. Depending on the insert size and on the physical distance separating the two flanking markers, a chromosome landing (Tanksley *et al.* 1995) or a chromosome walking will be performed. If the physical distance between the flanking markers is relatively small and it is comparable to that of the library insert size, it should be possible to identify, with only one library screening, a single clone containing

both markers (and therefore the gene of interest). On the contrary, if the physical distance is bigger and such a clone is not found or the clones identified with the two markers do not overlap, a chromosome walking has to be performed. In this case contigs of overlapping clones proceeding from the two flanking markers in direction of the gene have to be constructed until the chromosomal region co-segregating with the gene is completely covered.

It is essential for the map-based cloning of a gene to analyse large segregating populations to identify as many recombinant plants as possible in which a crossing-over event has occurred between the closest markers and the gene of interest. Those recombinants will be used for different purposes: a) to map the new markers produced during the chromosome walk in order to establish in which direction the contig has to be extended; b) to check the progress of the walk and c) to reduce as much as possible the region of the contig containing the gene.

The methodology of chromosome walking with BAC clones has been successfully applied in rice to isolate the genomic regions containing the *Pi-ta²* and *xa5* genes, conferring resistance to rice blast and bacterial blight respectively (Nakamura *et al.* 1997, Yang *et al.* 1998).

3. Once the minimal chromosomal region spanning the gene has been identified the gene itself has to be identified. Different strategies can be used depending on the size of the delimited region. If the region is relatively small the fastest way to identify candidate genes is to completely sequence the clones (or clones) to find candidate open reading frames (ORFs) showing high homology to known resistance genes. In contrast, if the region to analyse is still quite large it will be preferable to screen a cDNA library using the inserts of the BAC clones of the contig as probes to identify cDNAs present in the genomic region. Martin *et al.* (1994) used this strategy to isolate the *Pto* gene. Once ORF sequences or partial ORF sequences are found, the comparison of those sequences with the public databases will reduce the number of candidate genes. Among all cDNAs only a few will provide similarities with already cloned genes. The best candidates will then be used in complementation experiments to transform susceptible plants (in the case of cloning of plants resistance genes). Only the resistance gene will in fact be able to convert a susceptible cultivar into a resistant one, triggering the plant defence mechanisms. On the contrary, if no homology can be found among the cDNAs hybridising to the clones of the contig spanning the region of the gene,

to the already cloned resistance genes, all the genomic sequences containing the complete ORF of the cDNAs have to be used in complementation experiments. Again the one that will transform a susceptible cultivar into a resistant one will be the resistance gene.

To recognise candidate resistance gene among all the ORFs found it is necessary to increase our knowledge about the main features of the already cloned resistance genes and about the mechanism of plant resistance.

Plants possess two types of resistance. The first type is called horizontal resistance (or quantitative) and it is effective against all isolates of a pathogen. It is assumed that this resistance is due to the additive effect of resistance gene alleles. The second type of resistance is called vertical (or qualitative) and is effective against all pathotypes except those few pathotypes that carry the specific virulence allele that can overcome the specific resistance gene allele. This system is explained by the gene-for-gene hypothesis (Flor 1971).

Pryor (1987) proposed the following scenario to explain the development of such a system. The evolutionary ground state is considered to be a compatible interaction in which a pathogen has evolved to be virulent on a particular host plant. Selection favours the evolution and spread of host individuals that specifically recognise the pathogen and resist to the infection. For example, a receptor that evolved to activate a defence response to pathogens in general may be modified so that it specifically recognizes a particular pathogen product (an avirulence allele product). The pathogen responds by modifying or losing the avirulence allele by mutation. The host is now susceptible and again selection is brought to bear on new host R-gene allele specificity. Consequently, the evolution of gene-for-gene, actually an allele-for-allele interaction, can be seen as a continuing move-countermove process.

A typical example, not following Flor's gene-for-gene theory, comes from the first cloned resistance gene, the maize *Hm1* gene (Johal and Briggs 1992). *Hm1* encodes for a protein that inactivates the HC-toxin, a pathogenicity factor produced by the fungus *Cochliobolus carborum*. The genetics of interaction of this gene differ from those of the gene-for-gene because toxin-deficient strains also lose their ability to cause disease in maize cultivars that do not carry *Hm1*.

However, many of the resistance genes cloned so far follow the gene-for-gene theory. Such R-genes have been divided into five classes according to some features they have in common (De Wit 1997, Bent 1996):

- I: The first class is represented by the first cloned plant resistance gene that conforms to a classic gene-for-gene relation, the *Pto* gene of the tomato, conferring resistance against *Pseudomonas syringae* *pv.* *tomato* (*Pst*) (Martin *et al.* 1992). *Pto* encodes a cytoplasmic serine/threonine protein kinase, suggesting a role for *Pto* in signal transduction by protein phosphorylation. The modulation of the phosphorylation state is the most common mechanism that living organisms use to control protein activity. A model proposed to explain the mechanism of the signal transduction after physical interaction of *Pto* with the *AvrPto* protein of *P. syringae* *pv.* *tomato* has been formulated, thanks to the results of yeast two-hybrid experiments (not yet proven *in planta*) (Zhou *et al.* 1997). A type III secretion system secretes *AvrPto* protein into the plant cell where it interacts with the *Pto* protein. The recognition event, which may also involve *Prf* (Salmeron *et al.* 1995), a gene required for the hypersensitive response, activates the *Pto* kinase. *Pto* self-phosphorylates and becomes active. The *Pti1* protein, a serine/threonine kinase, interacts with the *Pto-AvrPto* complex and becomes phosphorylated, too. The phosphorylated *Pti1* is then involved in the activation of a cascade of reactions leading to the hypersensitive response. The *Pto-AvrPto* complex interacts with and also phosphorylates *Pit4*, *Pit5* and *Pit6* transcription factors involved in the regulation of many PR genes.
- II: The second class of resistance genes encodes for cytoplasmic proteins with leucine-rich repeats (LRR), leucine zippers (LZ) and nucleotide binding sites (NBS), suggesting a role for these molecules in intracellular signaling. LRR are multiple, serial repeats of a motif of about 24 amino acids in length. LRR contains leucines or other hydrophobic residues at regular intervals and can also contain regularly spaced prolines and asparagines. Specificity of LRR resides in the amino acids between the leucines of the repeat (Kobe and Deisenhofer 1994). In terms of function LRR have been shown to mediate protein-protein interaction (Kobe and Deisenhofer 1994). LRR may serve as binding domains for *Avr* proteins or may facilitate the interaction of R-gene proteins with other proteins that participate in the defence signal transduction (Dixon *et al.* 1996). NBS domains occur in diverse proteins with ATP or GTP binding activity. The function of NBS domains in the activation of the plant's defence is yet unknown. Finally the LZ motif, present in this class of proteins, seems to play a role in facilitating the protein-protein interaction. LZs are known for their role in homo-

and heterodimerization of eukaryotic transcription factors, although their role in R-gene function is only speculative.

Members of this class are *RPS2* of Arabidopsis against *Pseudomonas syringae tomato* (Bent *et al.* 1994), *RPM1* of Arabidopsis against *Pseudomonas syringae pv maculicola* (Grant *et al.* 1995), *I2* of the tomato against *Fusarium oxysporum lycopersici* (Simons *et al.* 1998), *Mi* of the tomato against *Meloidogyne spp.* (Milligan *et al.* 1998) and *Pib* of rice against *Magnaporthe grisea* (Wang *et al.* 1999).

- III: The third class is composed of genes very similar to those of the second class, but in addition to the LRR and NBS domains the proteins encoded by those genes contain an N-terminal region homologous to the *Drosophila* Toll protein and the mammalian interleukin-1 receptor (IL-R). This region in plant R-genes has been designated the TIR (Toll/Interleukin-1 Resistance). Due to the sequence similarities between R-genes and Toll and Interleukin-1, a working hypothesis is that these R-genes will trigger plant response by a similar mechanism (activation of transcription factors).

Members of this class are the *N* gene of tobacco against Tobacco Mosaic Virus (Whitman *et al.* 1994), *L6* of flax against *Melampsora lini* (Lawrence *et al.* 1995) and *RPP5* of Arabidopsis against *Peronospora parasitica* (Parker *et al.* 1997).

- IV: To the fourth class belong genes encoding extracytoplasmic LRR proteins with membrane anchors (transmembrane domain) and short C-terminal cytoplasmic tails. To this class belong the tomato genes *Cf-9* (Jones *et al.* 1994), *Cf-2* (Dixon *et al.* 1996), *Cf-4* (Thomas *et al.* 1997) and *Cf-5* (Dixon *et al.* 1998), conferring resistance to different races of *Cladosporium fulvum*, and *Hs1^{pro-1}* of sugar beet, conferring resistance to the nematode *Heterodera schachtii* (Cai *et al.* 1997).

One of the best-studied genes of this class is *Cf-9*. The N-terminus of the *Cf-9* protein includes a probable signal peptide necessary for its transportation across the cell membrane. This domain is followed by a domain composed of 28 LRRs with the consensus sequence LxxLxxLxxLDLSSNNLxGxIPsx (where x is any amino acid). The conserved glycine residue in the LRR is also found in other proteins that carry a putative extracellular LRR domain. Apart from some N-glycosylation sites there are few conserved amino acids internal to the LRR backbone. The C-terminus of the protein consists of a very acidic domain (10 of 18 amino acids are negatively charged), followed by a transmembrane domain

(37 hydrophobic amino acids), and a basic domain of 21 amino acids (8 of them positively charged). The C-terminus ends with the residues KKRY. In animals and yeast the KKxx motif functions as a signal for retrieval of membrane-bound proteins from the Golgi apparatus to the endoplasmic reticulum.

V: The fifth class is represented by the rice *Xa21* gene against *Xanthomonas campestris pv oryzae* (Song *et al.* 1995), a gene encoding for a protein with an extracellular LRR domain, a transmembrane domain and an intracellular serine/threonine kinase domain. Due to the presence of an extracellular LRR and a cytoplasmic kinase domain, the Xa21 protein appears to possess perceiving and signaling domains, so that only the interaction of the Xa21 protein with the AvrXa21 elicitor could be needed to start the resistance reaction.

An R-gene that cannot be placed into one of the five classes mentioned above is *Mlo* of barley against *Erysiphe graminis f. sp. hordei* (Büschges *et al.* 1997). This monogenic resistance is mediated by the recessive allele *mlo*. The resistance is effective against almost all tested isolates of the fungal pathogen, which means the resistance is not race-specific. Under pathogen-free conditions *mlo* plants exhibit a spontaneous leaf cell death phenotype and the appearance of multiple defence functions such as cell wall modification, accumulation of defence-related transcripts and phytoalexins. The deduced amino acid sequence of Mlo presents seven transmembrane helices and reveals no homologies to any other described plant resistance gene. Two hypothesis of the function of Mlo have been proposed by Büschges *et al.* (1997): 1) Mlo suppresses a default cell suicide program in foliar tissue and/or 2) Mlo has a specific negative regulatory function by down-regulating multiple defence functions. In any case Mlo does not seem to be involved in recognition of the pathogen.

It is very interesting that R-genes from different species with specificity to a wide variety of viral, nematode, bacterial and fungal pathogens often encode structurally similar proteins. This similarity suggests a high degree of conservation among the pathways that plants use to trigger defence responses. Another feature widespread in plants is that very often the members of an R-genes family are clustered in the genome. Different genes in these clusters often encode resistance against different races of the same pathogen species. One example is the *Cf-9* cluster that besides *Cf-9* contains *Cf-4* and three other *Cf-9* homologues (called *Hcr9s* for homologue of *Cladosporium* resistance gene *Cf-9*, Thomas *et al.* 1998).

The interaction between R-gene products and the pathogens Avr-gene products are only the initial steps of a complicated cascade of responses that collectively confer resistance on the plants. The following section reports a series of induced downstream reactions as reviewed by Hammond-Kosack and Jones (1996).

1. *Hypersensitive Response (HR)*. Classically HR is defined as the localised induced host cell death at the site of infection by a pathogen (Agrios 1997). In interactions with obligate biotrophic pathogens, cell death would deprive the pathogen of access to further nutrients. In the case of hemibiotrophic and necrotrophic pathogens, the cellular decompartmentalization may lead to the release of harmful preformed vacuolar substances. Alternatively, the level of induced phytoalexins, which normally are rapidly turned-over in plant cells, may accumulate to inhibitory concentrations because they are no longer metabolised.
2. *Reactive Oxygen Species (ROS)*. The production of ROS is often the first response activated in many incompatible interactions. ROS include superoxide (O_2^-), hydrogen peroxide (H_2O_2) and hydroxyl radical ($\cdot OH$). These highly reactive oxygen species are produced by enzymes involving e.g. NADPH oxidase of the plasma membrane of the host cell. The activated oxygen radicals trigger the hydroperoxidation of membrane phospholipids, producing mixtures of toxic lipid hydroperoxides that disrupt the plant's cell membranes. The ROS could also be involved in oxidation of phenolic compounds into more toxic quinones. H_2O_2 is essential for the formation of lignin polymer precursors via peroxidase activity. ROS are also responsible for the cross-linking of cell wall glycoproteins, thus rapidly making the plant cell walls more refractory to microbial penetration and enzymatic degradation. H_2O_2 is required to increase the production of enzymes required for the synthesis of salicylic acid (SA). The generation of ROS leads also to the alteration of the redox potential. It is known that in mammals many transcription factors are redox regulated. Changes of the redox potential may also regulate the stability of specific defence related mRNA transcripts in plants (Mehdy 1994)
3. *Cell wall fortification (CWF)*. The fortification of cell walls can increase resistance in various ways. CWF could impede leakage of nutrients for the pathogen. The diffusion of pathogen toxins and enzymes to the sensitive parts of the cells can be retarded. For biotrophic pathogens CWF could prevent entry of haustoria precluding parasitism. Cell wall fragments produced by hydrolyzing

enzymes may elicit additional defence response. One type of CWF is the formation of papillae that physically block fungal penetration. Rapid callose deposition in cell walls is also associated with pathogen incompatibility. The blockage of plasmodesmata with callose is essential to impede cell-to-cell movement of viruses.

4. *Salicylic Acid (SA)*. SA has been demonstrated to be involved in HR and in stimulation of systemic acquired resistance (SAR). SAR acts non-specifically throughout the plant and reduces the severity of disease caused by all classes of pathogens including those that are normally virulent. SAR can be induced by the infection with a pathogen or by treatment with certain chemical compounds (SA for example). The role of SA in defence is complex and may vary from species to species. It has been demonstrated for example that the lack of accumulation of SA in transgenic tobacco and Arabidopsis lines correlated with weakened local R-gene-mediated resistance responses (Delaney *et al.* 1994) and also with a blockage in the induction of various defence genes. In contrast, no reduction of the Cf-gene-mediated resistance was observed in the *C. fulvum* - tomato interaction in transgenic tomato plants with reduced SA production.
5. *Pathogenesis Related proteins (PR)*. PR proteins include intra- and extracellular proteins that are widely distributed in plants in trace amounts but are produced at much greater concentrations after pathogen attack or stress. PR proteins may show either strong antifungal or antibacterial activity and some are known to be chitinases, glucanases or chitin binding proteins.
6. *Phytoalexins*. Phytoalexins are toxic antimicrobial substances produced in appreciable amounts after stimulation by infections with various types of phytopathogenic microorganisms or by chemical compounds or by mechanical injury (Agrios 1997). Phytoalexins are produced by healthy cells adjacent to damaged and necrotic cells in response to substances diffusing from the damaged cells. Phytoalexins are not produced during compatible biotrophic infections. They accumulate around both resistant and susceptible necrotic tissues. Resistance occurs when one or more phytoalexins reach a concentration sufficient to restrict pathogen development. Most of the phytoalexins are produced in response to infection by fungi, but a few bacteria, viruses and nematodes have also been shown to induce production of phytoalexins. Most phytoalexin elicitors are high molecular weight substances like glucans, chitosan, glycoproteins and

polysaccharides, which are released from the fungal cell wall by plant enzymes. Most of those elicitors are not specific, they are present in both compatible and incompatible races of the pathogen and induce phytoalexins accumulation irrespective of the plant cultivar. Species or races of fungi that are pathogenic to a particular plant species seem to stimulate production of lower concentrations of phytoalexins than non-pathogens (Agrios 1997).

This short introduction summarises the present knowledge about the mechanism of resistance in plants. The positive results obtained in other species and good background information about apple scab encouraged the initiation of a map-based project aimed at cloning the apple resistance gene *Vf*. The cloning of *Vf* is important also because, in contrast to almost all resistance genes cloned so far, *Vf* does not show a hypersensitive response. The resistance is characterised by symptoms ranging from no symptoms to necrotic lesions with little sporulation, while lesions with abundant sporulation are observed in plants not carrying *Vf*. This information about the *Vf* phenotype, a genetic map of the *Vf* region (Gessler *et al.* 1995) and three *Vf* flanking markers (M18, AL07 and AM19) were the starting points of the map-based cloning project of the *Vf* gene.

This thesis reports the following steps:

- fine mapping of the *Vf* gene to establish its position on the genetic maps
- physical mapping to estimate the maximal distance between the *Vf* flanking markers
- chromosome walking using BAC clones to span the *Vf*-region with overlapping clones
- analysis of the coding sequences of the *Vf*-spanning contig
- identification of a cluster of genes with homologies to the *Cf* resistance genes in tomato.

2. MATERIAL AND METHODS

2.1 FINE MAPPING *Vf*

2.1.1 *Plant material*

For the fine mapping of *Vf*, young leaves from the crosses Braeburn x FAW 7167 (*Vf*) and Fuji x Ariwa (*Vf*), 317 and 441 plants respectively, were collected in nurseries of the FAW (Swiss Federal Research Station Wädenswil), frozen in liquid nitrogen, and freeze-dried.

Scab resistance was evaluated after greenhouse tests as described in Gianfranceschi *et al.* (1996). Plants were classified for resistance according to Chevalier *et al.* (1991).

2.1.2 *DNA extraction for fine mapping*

DNA was extracted following the protocol of Aldrich and Cullis (1993), with minor modifications. The extraction buffer was modified by adding 2 % PVP40 (w/v) and reducing the final concentration of β -mercaptoethanol to 2 % (v/v). The DNA was purified from RNA using RNaseA (Boehringer Mannheim). All the volumes were adapted to the small-scale extraction.

2.1.3 *Fine mapping of the *Vf* linked markers M18 and AL07*

The primers of the CAPS marker M18 (Gianfranceschi *et al.* 1996) were modified in order to increase the annealing temperature to 60 °C. The new sequences of the primers are: CH-M18for ACCACCAACTCCACCCTACA, CII-M18rev CTTT CACATGTCTAACAATTTTGG. Specific PCR amplifications of M18 and AL07 (Gessler *et al.* 1995) markers were performed in a volume of 15 μ l containing 5 ng of genomic DNA, 1 U Taq DNA Polymerase (Amersham-Pharmacia) 1X provided reaction buffer (50 mM KCl, 1.5 mM MgCl₂ and 10 mM Tris-HCl), 0.1 mM dNTP (Boehringer Mannheim) and 0.2 μ M of each primer. Amplifications were performed in a Perkin Elmer Cetus Gene Amp PCR System 9600. The following temperature cycles were used: 2 min 30 sec at 94 °C, then 35 cycles of 30 sec at 94 °C, 30 sec at

60 °C and 1 min at 72 °C, and 1 cycle of 10 min at 72 °C. Samples were kept at 4 °C after amplification. M18 amplification products were then cleaved with 3 U *TaqI* restriction enzyme (Boehringer Mannheim).

The results of the screening of the two segregating populations were added to the previously screened Florina x Nova Easygro population (491 plants, Gianfranceschi *et al.* 1996).

JoinMap v. 1.4 (Stam 1993) was used to construct genetic maps since this program allows the construction of a linkage map using data obtained from different segregating populations.

2.2 PHYSICAL MAPPING OF THE VF-REGION

2.2.1 Plant material

The cultivar Florina was the source of DNA for physical mapping experiments. The plants were grown in a greenhouse and young leaves were collected, immediately frozen in liquid nitrogen and either directly used to prepare HMW-DNA or stored at -80°C.

2.2.2 Preparation and embedding of intact nuclei in agarose plugs

Apple nuclei were isolated from 20 g of fresh or frozen leaves as described by Zhang *et al.* (1995) with the following modifications.

Anti-oxidants were added to the extraction buffer as reported by Paterson *et al.* (1993), and the time of the slow centrifugation (57 g) was extended to 15 min. To better remove leaf debris, the solution containing nuclei was centrifuged a second time at 57 g at 4 °C for 15 min before the last washing step. Nuclei were embedded in agarose plugs by mixing them with an equal volume (for 20 g leaves about 1 ml) of liquid 1.6 % low melting point agarose maintained at 45 °C. The mixture was then poured into ice cold plug moulds and kept on ice for 15 min.

The plugs (20 mm, 10 mm, 1 mm) were subsequently incubated in the lysis buffer (0.5 M EDTA pH 9.0-9.3, 1 % sodium lauryl sarcosine, 0.1 mg/ml proteinase K) for 36-48 h at 50 °C with gentle shaking.

2.2.3 Digestion of the embedded HMW-DNA

After incubation in the lysis buffer the plugs were washed as described by Zhang *et al.* (1995). The plugs were then cut into quarters and separately equilibrated for 2 h on ice with 1 ml of 1x restriction enzyme reaction buffer and 2 mM spermidine (Fluka Chemie, Buchs, Switzerland). After one hour the buffer was replaced and the plugs were incubated for a second hour. After equilibration, the plugs were incubated for another hour on ice in 200µl of 1x restriction enzyme buffer, 2 mM spermidine and 60 U of the respective restriction enzyme (Boehringer Mannheim) in order to allow the enzyme to access the DNA within the agarose plugs. Digestion was performed overnight at the recommended temperature.

The following 16 rare cutting restriction enzymes have been tested: *Cla*I, *Mlu*I, *Nae*I, *Not*I, *Sal*I, *Sfi*I, *Sma*I, *Xho*I, *Apa*I, *Bbr*PI, *Ksp*I, *Nar*I, *Pvu*I, *Spe*I, *Swa*I, *Xba*I (Boehringer Mannheim).

2.2.4 Pulsed field gel electrophoresis, blotting and hybridisation

Contour-clamped homogeneous electric field (CHEF, Chu *et al.* 1986) electrophoresis was performed in a CHEF DR II (BioRad) system. Pulsed field gel electrophoresis (PFGE) was performed in 1 % agarose in 0.5x TBE. For the separation of the DNA fragments shown in figure 2, the gel was run for 22 h with an initial pulse time of 60 s and a final pulse time of 120 s at 200 V. Gels were stained with ethidium bromide (1µg/ml) for 15 min, de-stained for 10 min in ddH₂O and photographed.

Alkali blot of the DNA was performed on Amersham Hybond-N⁺ nylon membranes using LKB 2016 vacugene vacuum blotting (Pharmacia), following the manufacturer's instructions with the following modifications: incubation time in 0.25 N HCl was extended to 40 min and blotting time was increased to 90 min.

The marker AM19 was not used as a probe because in Southern blot experiments it hybridised to repetitive sequences. About 50 ng of the PCR-amplified M18 and AL07

fragments were purified (NucleotraP[®]CR, Macherey-Nagel) and labelled using the Pharmacia oligo-labeling kit, following the manufacturer's instructions. No further purification was performed after labelling. Overnight pre-hybridisation was performed at 65 °C in 30 ml of the following solution: 10 % dextran sulphate, 1 % SDS, 50 mM Tris-HCl pH 7.5, 1 M NaCl and 600 µg of sheared and boiled salmon-sperm DNA. Hybridisation was performed by adding the denatured probe to the pre-hybridisation solution and incubating at 65 °C over night. Membranes were then washed once in 2x SSC for 10 min at room temperature, once in 0.5x SSC + 0.1 % SDS at 65 °C for 20 min and once in 0.1x SSC + 0.1 % SDS at 65 °C for 20 min. Membranes were exposed for 2-3 days at -80 °C to Kodak X-Omat AR autoradiographic films. Before re-using the membranes for hybridisation they were stripped in boiling 0.1% SDS for 5 min. Probe removal was checked by exposing the membranes at -80 °C on Kodak X-Omat AR autoradiographic film for 4 days.

2.3 CHROMOSOME WALKING

2.3.1 *The apple BAC library*

The apple BAC library (Vinatzer *et al.* 1998) used for the chromosome walk (kindly provided by H-B Zhang, Texas A&M University) has the following characteristics:

- Cloning vector: pECBAC1
- Inserts: partially digested high molecular weight DNA of 'Florina', cultivar heterozygous for the *Vf* resistance gene. Partial digestion was obtained by varying ratios of *EcoRI* to *EcoRI* methylase
- Total number of clones: 36,864 (96 384-wells plates)
- Average insert size: 120 kb
- Number of apple haploid-genome equivalents: five

Apple BAC library filters (provided by the BAC Center of the Texas A&M University): the 96 384-wells plates were high-density spotted on 24 filters (four plates on each filter). On each filter the four plates were spotted two times.

This part of the work was performed in collaboration with the group of Professor Silvano Sansavini, department of woody fruit trees (DCA), University of Bologna, Italy. These results have been published in Molecular and General Genetics (1999) 262: 884-891 with the title: Construction of a 550 kb BAC contig spanning the genomic region containing the apple scab resistance gene Vf.

Only the material and methods of the results that I obtained personally or part of the results obtained by the two groups that can not be separated is reported here.

2.3.2 Fine mapping of AL07 and AM19

To furthermore confirm the map position of *Vf* and to establish if the chromosome walking had to be started from AL07 or AM19 (on the opposite site from M18), a total of 2071 plants (from ten different populations) were analysed with the three closest molecular markers (M18, AL07 and AM19). PCR conditions for all three markers were the same as previously described. The primer sequences of the marker AM19 (Tartarini *et al.* 1999) are: AM19for CGTAGAACGGAATTT GACAGTG and AM19rev GACAAAGGGCTTAAGTGC. The results of the screening of our three mapping population ['Florina' × 'NovaEasyGro', 'Braeburn' × FAW 7167 (*Vf*) and 'Fuji' × 'Ariwa' (*Vf*)] were combined with the results of the screening of seven populations segregating for *Vf* of the DCA Bologna ['Primiera' × 'Self Cox's', 'Royal Gala' × 'Sel DCA 81.417.002', 'Primiera' × 'Sel DCA 79.410.002', 'GoldRush' × 'Sel DCA 79.410.002', 'Primiera' × 'Abbondanza rossa', 'Primiera' × 'Pink Lady' and 'Primiera' × 'Gold Chief'].

JoinMap v.1.4 (Stam 1993) was used to construct the linkage map of all 2071 plants.

2.3.3 Screening of the BAC library and analysis of putative positive clones

The apple BAC library (Vinatzer *et al.* 1998) has been screened by a radioactive method as follows: nylon membranes were pre-hybridised overnight at 65°C in hybridisation buffer (0.5M sodium phosphate pH 7.2, 7 % SDS, 1 % BSA, 1 mM EDTA and 0.2 mg/ml denatured salmon sperm). Hybridisation was performed overnight by adding 50 ng radioactively labelled (Pharmacia oligo-labelling kit) purified PCR product (NuclcotraP®CR, Macherey-Nagel). 1 ng of lambda DNA was mixed with the probe before labelling to visualise the grid of all clones spotted on the filters this simplifies the reading of the co-ordinates of the hybridising positive clones. Filters were then rinsed with 2×SSC and washed twice in 0.5×SSC, 0.1 % SDS for 25 min at 65°C and once in 0.1×SSC, 0.1 % SDS for 25 min at 65°C. The filters were exposed for 2-3 days to X-ray film (Kodak X-Omat AR). Putative positive clones were picked from the library and checked by colony PCR. BAC DNA was extracted (procedure see next paragraph) from those clones that yielded a PCR product of the expected length. Each sample (about 300 ng plasmid) was digested

with *EcoRI*, loaded on 0.8 % agarose gels, and run in 1×TBE or 1×TAE. Gels were stained with ethidium bromide, photographed and blotted onto Hybond-N⁺ membranes (Amersham). Hybridisations were performed as described above, with the same marker used as a probe for the screening.

2.3.4 BAC miniprep and insert-end isolation

Standard alkaline minipreps were performed to extract BAC plasmids. To produce the cleared lysate of 3 ml overnight cultures (LB medium with 12.5 µg/ml chloramphenicol), 250 µl of resuspension, 250 µl of lysis and 350 µl of neutralisation solution of the Promega Wizard Plus SV Miniprep DNA Purification System were used. Cleared lysate was precipitated with two volumes of ethanol and resuspended in 100 µl TE pH 8.0. No further purification was needed.

The end of the insert flanking the sequence of the T7 universal primer on pECBAC1 (vector used for the construction of the BAC library) was called T7-end, and the end of the insert close to the Sp6 universal primer sequence was called Sp6-end.

For insert-end isolation, 4 µl aliquots of BAC DNA were digested separately with 5U of various restriction enzymes (*HindIII*, *SphI* and *BamHI* for the T7-ends and enzymes *NcoI*, *EclXI*, and *BseAI* for the Sp6-ends) in a final reaction volume of 40 µl. After restriction digestion the DNA was precipitated and resuspended in 150 µl TE pH 8.0. The DNA in 70 µl aliquot was self-ligated in a total volume of 100 µl using 20 units of T4 ligase (Pharmacia) at 16 °C overnight. Five µl of the ligation products were used for end isolation by inverse PCR (IPCR) (Cai *et al.* 1995).

PCR reactions were performed in a Perkin-Elmer Gene Amp PCR System 9600 in a volume of 50 µl: 10mM Tris-HCl pH 9.0, 1.5 mM MgCl₂, 50 mM KCl, 0.1 mM dNTPs, 0.2 µM primers and 0.07 U/µl Taq polymerase (Pharmacia). PCR was carried out for 35 cycles at 94 °C for 30 sec, 50 °C for 30 sec and 72 °C for 1 min with a final cycle of 10 min at 72 °C.

T7 insert-ends were amplified by using T7 and Sp6 universal primers while three primers compatible with the universal primer Sp6 were developed to perform IPCRs of the Sp6-ends (*BseAI*for: AACACTATCCCATATCAC, *EclXI*for: GCTAACTC ACATTATTG or *NcoI*for: ATTCATTAAGCATTCTG). The primer *EclXI*for can be used also with the restriction enzyme *EaeI*. To isolate the Sp6-ends, the BAC1/BAC2

primers (Woo *et al.* 1994) were also used in combination with the ligation products of the *Hind*III digestion.

PCR products longer than 350 bp were purified (QIAquick PCR Purification Kit, Qiagen), quantified, and about 50 ng were used as a template for cycle sequencing. Reactions were run on an ABI 373 DNA sequencing system (PE Applied Biosystem, Foster City CA, USA). If more than one restriction enzyme produced a PCR fragment, the smaller fragment was sequenced from both sides (if necessary); the larger fragments were sequenced only with the Sp6 primer for the T7 end-rescue, or with either the BseAIfor, EclXIfor or NcoIfor primer for the Sp6 end-rescue. The sequences were then assembled using GCG (Genetics Computer Group, Madison, WI, USA) software package, version 8. The program Primer version 0.5 (Whitehead Institute for Biomedical Research) was used to develop specific primers on the assembled sequences of the ends of each insert.

2.3.5 Determination of contig ends during assembly

Two approaches were used to identify the most external BAC ends of the contig. The first one was based on PCR. All specific BAC end primer pairs were tested on all clones found by screening the library, including the clone from which the probe was generated. The primer pair that amplified a fragment only from the clone from which it was generated was considered as the new end of the contig. The second approach consisted of hybridising Southern blots carrying *Eco*RI digestions of all clones of the contig with labelled PCR-amplified ligation products of end-rescues or PCR fragments amplified with end sequence-specific primers (procedure as described before). Again, the BAC end hybridising only to the clone from which it was derived or hybridising to an *Eco*RI restriction fragment not present in the other clones (in the case that the corresponding sequence was repeated in the region), was considered the new end of the contig.

2.3.6 Mapping of insert end-sequences and preparation of probes for the succeeding step

To produce a polymorphic marker on the new end of the contig it was first determined whether the primers developed there showed a length polymorphism between plants known to be either homozygous- heterozygous-resistant or homozygous-susceptible. If no length polymorphism was detected, the PCR products were digested with restriction enzymes in order to identify one that reveals a polymorphism. The enzymes were chosen from those that cut the end sequence between the two primers, as determined with the MAPSORT function of GCG software package (Genetics Computer Group, Madison WI, USA).

To determine if a marker was suitable for used as a probe for the next screening of the library, the PCR product of the marker was used as probe on a Southern blot of 'Florina' DNA. If bands could be distinguished, i.e. the probe didn't produce a smear, the marker was used for the next screening.

2.4 IDENTIFICATION OF RESISTANCE GENE HOMOLOGUES

2.4.1 Starting material

The 101 cDNAs found to hybridise to the *Vf*-region and the *Eco*RI and *Sau*3AI BAC subclone libraries of the five minimal BAC clones required to span the *Vf*-region (M18-2, M18-5, M18-6, M18-7 and M18-8), kindly provided by B. Vinatzer (DCA Bologna, Italy), were the starting material of this part of the thesis.

cDNA library characteristics:

- Plant material: leaves of the apple cultivar 'Florina' (heterozygous for *Vf*) harvested 24 h and 48 h after inoculation with a conidial suspension of *Venturia inaequalis* (Gianfranceschi *et al.* 1996)
- Kit used: Stratagene cDNA Synthesis Kit
- Probe used for cDNA library screening: ³²P labelled inserts of the five minimal BAC clones required to span the *Vf*-region
- Total number of cDNA clones screened: about 750,000
- Total number of cDNAs hybridising to the probe: 101 cDNAs

*Eco*RI and *Sau*3AI BAC sublibraries:

- Cloning vector: pGEM-4z (Promega)
- High density filters were prepared using the Beckman 2000 Automated Workstation after Zhang *et al.* (1996)

This part of the work was also performed in collaboration with the group of Professor Silviero Sansavini, department of woody fruit trees (DCA), University of Bologna, Italy.

Only the material and methods of the results that I obtained personally or part of the results obtained by the two groups and that cannot be separated is reported here.

2.4.2 Characterisation of positive cDNAs

The inserts of the 101 cDNAs, hybridising to the inserts of the five minimal BAC clones required to span the *Vf*-region, were amplified using the primers M13-forward and M13-reverse to determine their insert size. To verify that they were real positives, Southern blots were performed after running the PCR products on 1% agarose gels using the same probe as for the cDNA library screening. The plasmids of the positive cDNAs were extracted (Wizard Plus SV miniprep kit, Promega), gel quantified and sequenced from the poly-A tail with the M13-forward primer. cDNA sequences were compared with each other to identify multiples using the function Fasta of the GCG software package (Genetics Computer Group, Madison WI, USA). The longest cDNA of each group was also sequenced with the M13-reverse primer. In order to identify candidate *Vf* genes, BlastX and Fasta searches were performed on all sequences using the www.ddbj.nig.ac.jp website to analyse the cDNAs for homologies to cloned resistance genes.

2.4.3 Development of *fv* primers

During the screening of the cDNA library three cDNAs with homology to the *Cf* resistance gene family of tomato (Jones *et al.* 1994, Dixon *et al.* 1996, Thomas *et al.* 1997, Dixon *et al.* 1998) were identified. The three cDNAs (8-1, 10-5 and 15-4) were compared with each other and the region spanning the stop codon was found to be highly similar in all of them. In this conserved region the *fv*-rev primer (AGCACTATCTATTCAGCAAYYSTG) was developed. Comparing the sequences of the BAC subclones (hybridising to the cDNA 8-1) with the *Cf* resistance genes sequences, it was found that some subclones presented homology to the 5' region of the *Cf* genes (within the LRR domain). On one of those subclone sequences (63Sau2H13m13r), the primer *fv*-for was developed (CAATGCCTTACGTGGTG AAA) (Fig. 7). PCR conditions were the same as for the markers AL07, CH-M18 and AM19 (see previous section).

2.4.4 Cloning of PCR products

PCR products obtained from the BAC clones spanning the *Vf*-region using the *fv*-for and *fv*-rev primers (*fv*-PCR products) were gel purified (Qiagen Gel extraction kit), gel quantified and ligated into the pGEMT cloning vector (Promega), following manufacturer's instructions. To distinguish transformants that had inserts of the same size, the clones were analysed by RFLP-PCR, amplifying the inserts using the *fv* primers and digesting the PCR products with *Sau3AI*, *HaeIII*, *AluI* and *TaqI* separately. The different PCR products that were identified were completely sequenced by primer walking.

2.4.5 RT-PCR

RT-PCR was performed to check the expression of candidate resistance genes. Total RNA was extracted (Promega SV Total RNA Isolation System following manufacturer's instructions) from about 25 mg of leaves of three cultivars, 'Florina' (*Vf*), 'Prima' (*Vf*) and 'Golden delicious' inoculated with *V. inaequalis*. Inoculation of young leaves was performed as described by Gianfranceschi *et al.* (1996). The leaves were harvested at 0 h, 12 h, 24 h, 48 h, 72 h and 96 h after inoculation. For infection times longer than 48 h the plants were transferred from the growth chamber (100 % humidity, 18 °C) to the greenhouse (80 % humidity, 18 °C). To perform RT-PCR the QIAGEN OneStep RT-PCR kit was used, following the manufacturer's instructions. As template for RT-PCR three bulks, one for each cultivar, were set up by mixing the same volumes of 1:50 dilutions of extracted total RNA of leaves harvested at different times, without previous quantification. To check the quality of the extracted RNA (absence of DNA), control reactions were performed by inhibiting reverse-transcriptase (15 min at 95 °C). To perform RT-PCR of *HcrVf2*, the primers *fv*-for (CAATGCCTTACGTGGTGAAA) and 6210rev (CAGGGATTCCAGCCAATCTA) were used. These primers were also used for the mapping of *HcrVf2*. RT-PCR of *HcrVf3* was performed with the primers ORF3for (TCACGGTTCAAAGACCATCC) and ORF3rev (GCAGGGATCCCAGATATTGC). The annealing temperature of all the primers is 60 °C. Single RT-PCR products were purified (QIAGEN QIAquick PCR Purification Kit), gel quantified and sequenced without cloning. RT-PCR reactions producing several bands were cloned in pGEM[®]-T vector using the Promega

pGEM[®]-T Vector Systems (cloning of RT-PCR products obtained with the fv primers).

2.4.6 RACE

RACE experiments were performed, using the Roche 5'/3' RACE kit, to confirm that the transcription start of the *HvrVf* ORF's was reached. The reactions were performed using the Roche Expand[™] High Fidelity PCR System. The 'Florina' RNA bulk, used for the RT-PCR experiments, was used as template. All three antisense primers, necessary for the amplification of the 5' gene end, were developed using the WWW Primer Picking program (Primer 3.0 from the Whitehead Institute for Biomedical Research) in common regions of *HcrVf1* and *HcrVf2* genomic sequences. *HcrVf3* and *HcrVf4* were not considered (*HcrVf3* is not expressed and *HcrVf4* is not completely characterised). For cDNA synthesis (as primer SP1), the primer RT-FD5 rev (CAACCAGTCAGATGCTTTG) was used. For the following two PCR reactions the primers 6210IPCRrev (TGCTTTAAACTGAGCAAAGAAGG) and the primer RACE Sp3 FD5-6210 (TATTAAGGTGCAGCTCGTGG) were used (as SP2 and SP3 primers respectively). Because more than one band was amplified all the amplified bands were cloned using the Promega pGEM[®]-T Vector Systems. Bands of the expected size were excised from the gel and blunt-end cloned using the Roche PCR cloning kit. Two to three clones for each band were sequenced and the sequences compared to the genomic sequences of the *HcrVf* genes.

2.4.7 ORFs analysis

Open reading frames were predicted using the ORF FINDER at the web site www.ncbi.nlm.nih.gov/gorf/gorf.html. Putative signal peptide cleavage sites and putative transmembrane domains of the *HcrVf* genes were predicted using the CBS prediction server on the web site www.cbs.dtu.dk/services/). N-glycosylation sites were predicted according to Alberts *et al.* 1990 each time the amino acids sequence Asn-X-(Thr/Ser) was found, with the X being any amino acid.

3. RESULTS

3.1 FINE MAPPING OF *Vf*

The order of *Vf* and the markers M18 and AL07 on the genetic map was already determined in the Florina (*Vf*) x Nova EasyGro (*Vf*) population (491 plants), generated for molecular analysis in 1994 (Gessler *et al.* 1995, Gessler *et al.* 1997). On that population the scab resistance scoring was done accurately and plants with dubious scoring have been double-checked for scab resistance in the field over several years. In that population it was proven that *Vf* maps between the two molecular markers M18 and AL07.

To increase the resolution of the genetic map around *Vf* and to find more recombinant plants necessary for the chromosome walking, the resistant plants of two other populations, Braeburn x FAW 7167 (*Vf*) (317 plants) and Fuji x Ariwa (*Vf*) (441 plants), were screened with the molecular markers M18 and AL07. The plants were classified into resistance classes according to Chevalier *et al.* (1991). Plants in classes 0 to 3b were considered to be resistant and plants in class 4 to be susceptible (the latter were not available). Those populations were generated for breeding purposes; therefore, their scab resistance scoring was not done as carefully as in the previous population. In fact, about 9 % (70 out of 758) of the plants considered to be phenotypically resistant were found to be carrying both molecular marker alleles in repulsion with *Vf* (Tab. 1).

Table 1: Scab resistance classification of the resistant plants in populations Braeburn x FAW 7167 (*Vf*) and Fuji x Ariwa (*Vf*) and distribution of the 70 plants with the alleles of M18 and AL07 in repulsion with *Vf* and their relative percentage in each class. The susceptible plants (class 4) of the two crosses were not available for mapping.

| Resistance classes | 0 | 2 | 3a | 3b | Total |
|---|------------|-----------|--------------|---------------|--------------|
| Total number of plants in each class | 33 | 172 | 422 | 131 | 758 |
| Number of plants with the marker alleles in repulsion with <i>Vf</i> divided into classes (and %) | 4 (12%) | 7 (4%) | 27 (6.4%) | 32 (24.4%) | 70 (9.2%) |

These results have been published in Theoretical and Applied Genetics 99 (6): 1012-1017 with the title: Towards the map-based cloning of Vf: fine and physical mapping of the Vf-region.

According to the order of the markers established in the Florina x Nova EasyGro, all those plants should be considered as double recombinants or as incorrectly classified for resistance. The latter is the most probable hypothesis, as supported by the presence of a higher percentage of these plants in the classes 0 and 3b, the two classes that are most prone to errors. Class 0 contains plants showing no symptom of infection and thus includes, besides the highly resistant plants, those plants that “escaped” the attack of the pathogen. Class 3b includes plants showing chlorotic and necrotic spots with little sporulation. Plants of this class are considered to be resistant, however it is not always easy to distinguish class 3b from class 4, which contains the heavily infected plants that are considered to be susceptible.

The 70 plants were re-checked in the field and 10 out of these were confirmed to be resistant, although the molecular analysis excluded the presence of the *Vf* gene. The greenhouse resistance classification of these 10 plants was the following: 1 plant was classified as class 2, 2 plants were classified as 3a and 7 were classified as class 3b. The remaining 60 plants were grouped as follows:

- 34 plants were found to be incorrectly classified as clear sporulating lesions were observed, so that the classification was changed to class 4 (susceptible plants).
- 19 plants were dwarf. Due to the slow and reduced growth, the leaves of dwarf plants usually present ontogenic resistance. Once transferred to the field some dwarf plants showed clear sporulating lesions, confirming this hypothesis.
- 5 plants showed a reduced growth. For this group of plants the same hypothesis as for the dwarf plants can be proposed. Also in this case a plant presenting sporulating lesions could be found in the field.
- 2 plants died during the winter and could not be re-analysed.

Although the check in the field allowed correcting the classification of more than half of these plants, all 70 plants were not used for the construction of the linkage map. The obtained genetic distances from M18 to *Vf* and AL07 to *Vf* are 0.2 cM and 1.1 cM respectively, giving a total of 1.3 cM between the markers M18 and AL07 (Fig. 1).

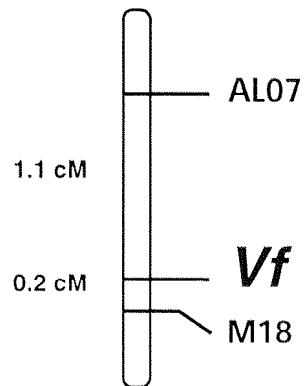


Figure 1: Linkage map of the *Vf* genomic region. The map was calculated with the program Joinmap 1.4 (Stam 1993) using the segregation data of three populations, Florina x Nova Easygro (491 plants), Fuji x Ariwa (409 plants) and Braeburn x FAW 7167 (279 plants), for a total of 1179 plants analysed.

3.2 PHYSICAL MAPPING OF THE *Vf*-REGION

3.2.1 Extraction of high molecular weight DNA (HMW-DNA)

White pellets of nuclei were obtained using the modified protocol of Zhang (1995). After the fixation of the nuclei into agarose plugs and the following incubation in lysis buffer, the plugs of some extractions turned to light brown while others remained white. This change of colour is probably due to the oxidation of a little cell debris containing polyphenols, but it does not seem to influence the digestion of the DNA. Pulsed-field gel electrophoresis (PFGE) of undigested and digested HMW-DNA proved that the extracted DNA is of megabase size and accessible to restriction enzymes (data not shown).

3.2.2 Southern blot

To estimate the physical distance occurring between the *Vf*-flanking markers M18 and AL07, 16 rare cutting restriction enzymes were tested. Only *NotI* produced a fragment that hybridised to both *Vf* flanking markers, AL07 and M18 (Fig. 2). The size of this fragment is about 870 kb. Since the genetic distance between AL07 and M18 is 1.3 cM (Fig.1), the maximum specific ratio between physical and genetic distance for the *Vf*-region was calculated to be 670 kb/cM.

Other restriction enzymes, such as *MluI*, *SfiI*, *SmaI* and *NaeI*, produced fragments hybridising to either marker ranging from 200 kb to 550 kb, but no other restriction fragment on which both markers could hybridise was found. Restriction enzymes *ApaI*, *KspI*, *PvuI* and *ClaI* did not produce fragments in the resolution range of the chosen PFGE settings. The HMW-DNA fragments of those restriction enzymes hybridising to the probes were in the compression band, indicating that their size is bigger than 1.6 Mbp. All other restriction enzymes tested produced fragments hybridising to the probes smaller than 100 kb (Table 2).

In figure 2, in the lanes of the restriction enzymes *MluI* and *NotI*, two bands appear. This can be explained by a restriction polymorphism (Florina is heterozygous for *Vf*) which could be caused not only by point mutations modifying the recognition site, but also by a different methylation of the site.

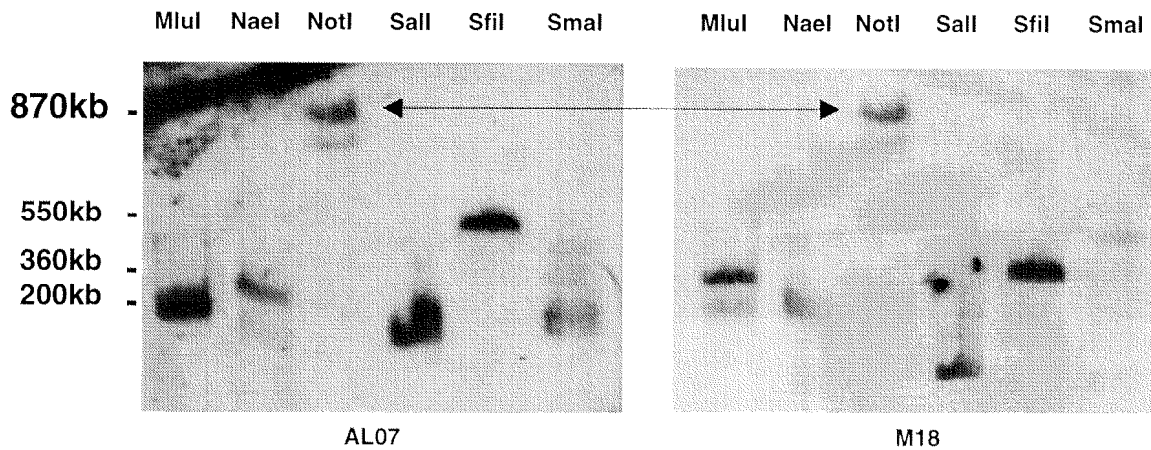


Figure 2: Hybridisation of a pulsed field gel electrophoresis (PFGE) gel with the two *Vf* flanking markers AL07 and M18. On the left panel the probe used was AL07, on the right one M18. The arrow indicates the 870 kb DNA fragment hybridising to both markers. PFGE conditions: 1 % agarose, 0.5x TBE, 14 °C, 22 h with an initial pulse time of 60 s and a final pulse time of 120 s at 200 V. Yeast chromosomal DNA strain YNN295 (BioRad) was used as size standard.

Table 2: Size of PFGE-separated restriction fragments (in kb) hybridising to the markers AL07 and M18

| Fragment | Probe | |
|--|-----------|-----------|
| | AL07 | M18 |
| <i>MluI</i> | 200 kb | 300 kb |
| <i>NaeI</i> | 230 kb | 170 kb |
| <i>NotI</i> | 870 kb | 870 kb |
| <i>SfiI</i> | 550 kb | 360 kb |
| <i>SmaI</i> | 150 kb | 400 kb |
| <i>ApaI, ClaI, KspI and PvuI</i> | >1,600 kb | >1,600 kb |
| <i>BbrPI, NarI, Sall, SpeI, SwaI, XbaI, and XhoI</i> | <100 kb | <100 kb |

3.3 CHROMOSOME WALKING

3.3.1 Fine mapping of AL07 and AM19

The *Vf* marker AM19, mapped with 109 plants, has been found to be co-segregating with AL07 (Tartarini 1996) and could therefore be an alternative starting point to AL07 for the chromosome walking. Analysing the 1179 plants of the previous section, no informative recombinant plants were found. To definitively determine which marker is closer to *Vf*, it was necessary to analyse seven more crosses to reach a total of 2071 plants. Two plants with a recombination between AM19 and AL07 were finally found. These two plants indicated that AM19 is closer to *Vf* than AL07.

The mapping of 2071 plants with all three markers (M18, AL07 and AM19) did not modify the order of the markers, the position of the *Vf* gene and the genetic distances also varied only slightly (Fig. 3A).

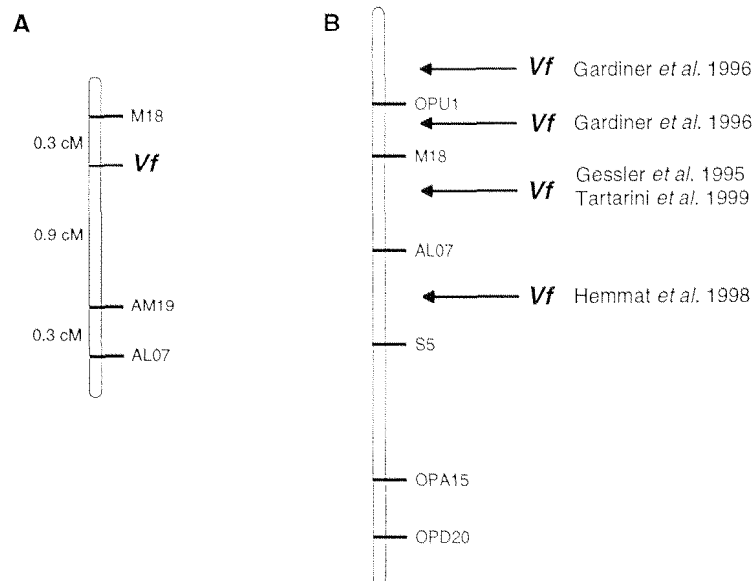


Figure 3: **A** Linkage map of the *Vf* genomic region. The map was calculated with the program Joinmap 1.4 (Stam 1993) using the segregation data of 10 populations, for a total of 2071 plants analysed. **B** Different *Vf* positions in the literature.

This part of the work has been performed in collaboration with the group of Professor Silvano Sansavini, department of woody fruit trees, University of Bologna, Italy. These results have been published in Molecular and General Genetics (1999) 262: 884-891 with the title: Construction of a 550 kb BAC contig spanning the genomic region containing the apple scab resistance gene Vf. Only the results that I personally obtained are reported here, along with inseparable parts of the results obtained jointly.

3.3.2 Identification of positive clones

Each clone hybridising to the probes used in the screenings was double tested in order to confirm their belonging to the contigs. The first check consisted in performing a PCR reaction using the specific primers of the marker used for the screening, in order to determine if the clones were real positives or contained only a sequence homologous to the marker. Since polymorphic markers were used for most screenings of the library, it was usually possible to determine immediately whether the newly identified clones were derived from the chromosome carrying the resistance gene or from its homologue. The second check consisted in the comparison of the *Eco*RI restriction patterns of the newly identified clones with those of the clones of the pre-existing contig. This test usually gave a clear confirmation of the overlap; furthermore, it permitted the estimation of the length of the overlap and the length of the concluded step. Probing the Southern blots bearing *Eco*RI digests with the BAC end used for the screening and with the BAC ends derived from the newly identified clones, also gave an additional confirmation of the correctness of the contig.

3.3.3 End rescue

End rescue IPCR always produced, at least with one of the restriction enzymes, a PCR fragment that could be sequenced. For most of the ends, PCR products from more than one restriction enzyme were obtained. These PCR products were usually of different lengths and enabled the design of two pairs of primers per BAC end; one immediately at the end of the BAC clone, and the other in a more internal position (approximately at 500 to 2000bp from the end). This option was used when a highly repeated sequence was present at the end of the clone, or when the first marker produced on the end of the clone was not polymorphic.

3.3.4 *The chromosome walk towards the Vf gene starting from M18*

The screening of the apple BAC library with M18 allowed the identification of three positive clones (M18S-1, M18S-2 and M18-1). The clone M18-1 contains DNA derived from the chromosome carrying *Vf*, which was called “resistant”, while the other two clones M18S-1 and M18S-2 contain DNA from the homologous chromosome not carrying *Vf*, which was called “susceptible”. All three BAC inserts contained a *NotI* restriction site.

The overlap of the clones M18S-1 and M18S-2 was determined by PCR. To determine in which direction the contig had to be extended to walk towards *Vf*, polymorphic markers on the protruding ends of the contigs were produced. PCR amplification of genomic DNA, using primers developed at the T7-end of M18S-1 (marker QU97), produced a polymorphic band, QU97L, that appeared only in the resistant plants, and one band of the expected size, QU97S, that appeared in both susceptible and resistant plants. QU97 was mapped and found to co-segregate with *Vf*. When M18-1, the only “resistant” clone found during the first library screening, was tested with QU97, only QU97S (and not QU97L) was amplified, indicating that we had found a new marker near *Vf* that was not on the contig in construction. For this reason it was decided to use QU97L as a probe in the next screening of the library. Unfortunately, when QU97L was hybridised to genomic DNA, it was found to contain a highly repeated sequence. Comparison of the sequences of QU97L and QU97S, revealed that the only difference between them was an insertion of about 160 bases in QU97L. Southern analysis indicated that the highly repeated sequence present in QU97L was this 160 bp insertion. However, since QU97S and QU97L share non-repetitive sequences, it was possible to use QU97S as a probe to find the clones containing QU97L. Thirty-seven new positive clones were found. Seven of them were confirmed as positives after PCR. Four clones were derived from the “resistant” chromosome; i.e., M18-2, M18-3, M18-4 and M18-5, and three from the “susceptible” one; i.e., M18S-3, M18S-4 and M18S-5. Interestingly, the clones M18-4 and M18-5 contained only the QU97L band, while M18-2 and M18-3 contained both QU97S and QU97L bands filling the gap between M18-1, which contains only QU97S, and M18-4 and M18-5. The fact that M18-2 and M18-3 contained both QU97S and QU97L indicates a duplication of this sequence; consequently, QU97S

and QU97L are not alleles of the same locus. Therefore, the screening of the library with QU97S enabled performing two steps of chromosome walking in one (Fig. 4).

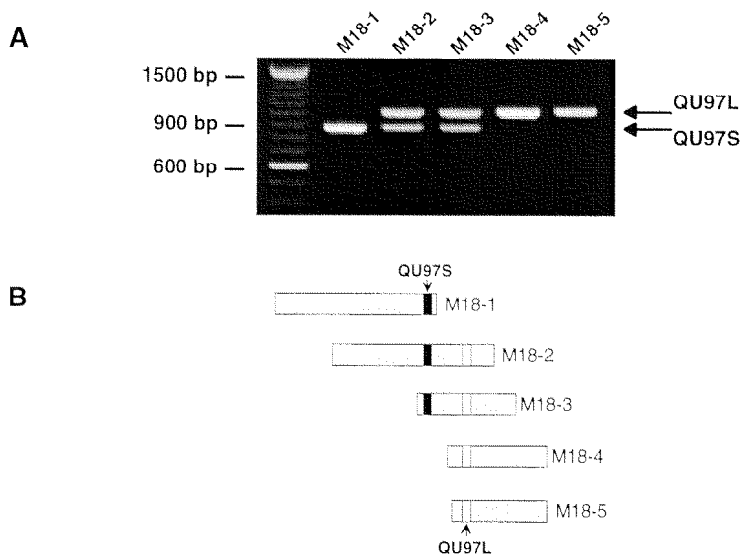


Figure 4: **A** PCR amplification with the marker QU97 of the first five clones derived from the “resistant” chromosome found during the chromosome walking started from M18. QU97 amplifies two bands on the “resistant” chromosome, QU97S and QU97L. The only difference between QU97S and QU97L is an insertion of 160bp in QU97L. **B** Overlaps of the five clones derived from the “resistant” chromosome containing the two QU97 loci deduced from the PCR pattern shown in Fig. 4A.

A similar observation was made with the “susceptible” contig. In this case, QU97S identified the clone M18S-5 that did not overlap directly with the clones found with the marker M18 (M18S-1 and 2); but the clones M18S-3 and M18S-4, also found with QU97S, covered the gap between them.

The restriction patterns of the clones M18-4 and M18-5 were very similar, and comparison of the Sp6-end sequences showed that they have one end in common. Hybridisation determined that the Sp6-end of M18-5 (and M18-4) was the new end of the “resistant” contig. The marker developed on the new end (M5S) was mapped and found to co-segregate with *Vf*.

During the next three steps of the chromosome walk, only one positive clone was identified in the library each time: M18-6, M18-7 and M18-8 respectively. These clones were confirmed and assigned to the “resistant” contig. The clone M18-8 was also identified during the chromosome walk that started from AM19, suggesting that the whole region between M18 and AM19 had been covered. On the M18-8 Sp6 sequence (i.e. the end of the M18 contig) a polymorphic marker (M8S) was developed. The analysis with the marker M8S of the plants that showed a

recombination in the region between M18 and AM19 demonstrated that this marker does not co-segregate with *Vf*, thus indicating that the genomic region spanning *Vf* had been completely covered (Fig. 5 and Fig. 6A).

The “susceptible” contig was not extended further because the Sp6-end of the last clone M18S-5 contained a highly repeated sequence that could not be used to screen the library. Furthermore, all but one marker (M7T) that were produced on the “resistant” contig were dominant, making it impossible to identify the clones from the “susceptible” chromosome among those identified during the screening of the library. In the case of M7T no “susceptible” clone was found, indicating a probable gap in the library for this region of the “susceptible” chromosome (Fig. 6A).

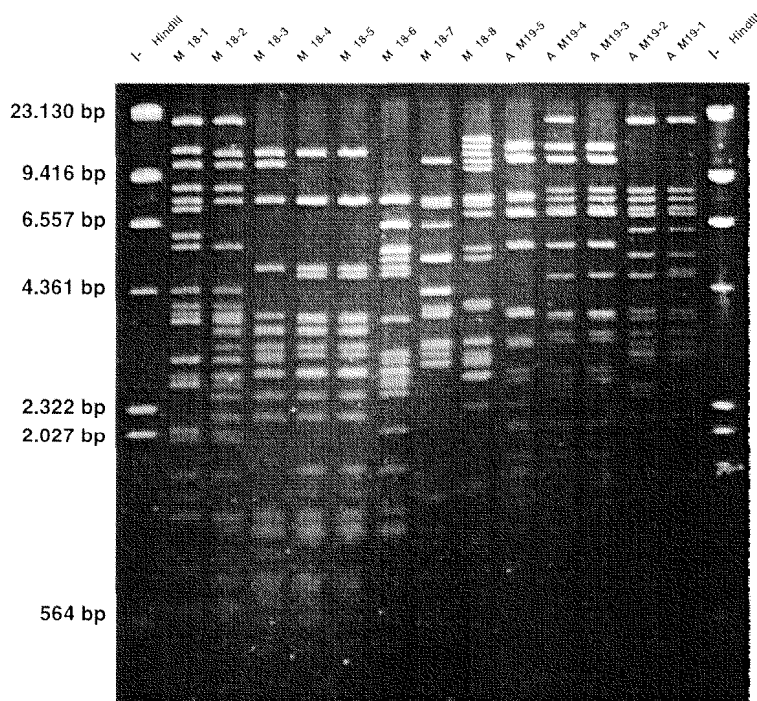


Figure 5: *EcoRI* restriction digests of the 13 BAC clones spanning the chromosomal region between M18 and AM19 that includes *Vf*.

Information about the chromosome walking started from AM19 and AL07 is published in Molecular and General Genetics (1999) 262: 884-891 with the title: Construction of a 550 kb BAC contig spanning the genomic region containing the apple scab resistance gene Vf.

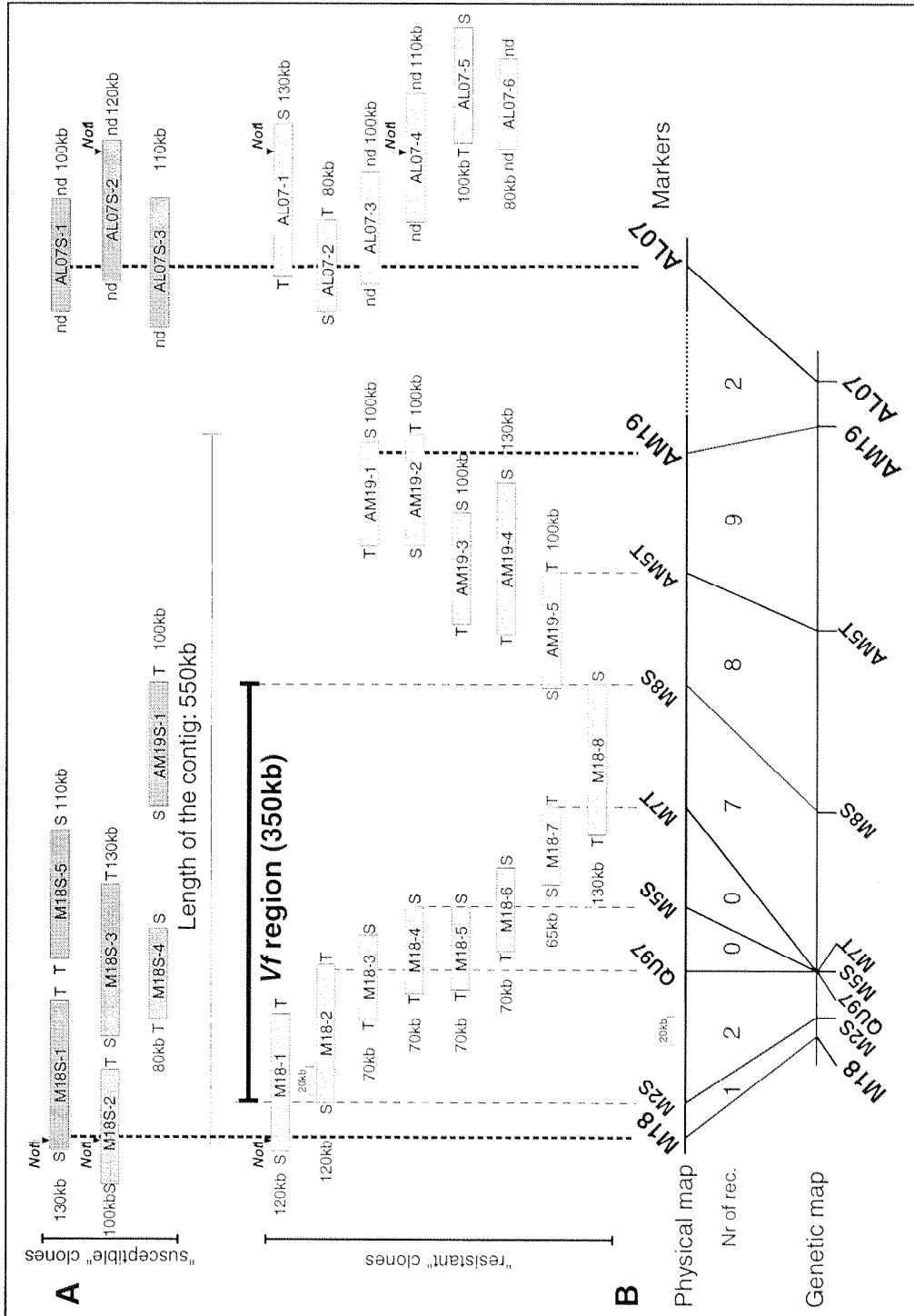


Figure 6: **A** BAC contig spanning the *Vf*-region. BAC clones derived from the "resistant" chromosome, carrying *Vf*, are indicated with light grey horizontal bars. BAC clones derived from the homologous "susceptible" chromosome are indicated with grey horizontal bars. Numbers beside the clones indicate their insert size. T and S indicate the T7 and Sp6 ends of the insert, respectively, and 'nd' indicates that the kind of end was not determined. The vertical arrows indicate the two *NotI* restriction sites flanking the *Vf*-region. **B** Relationship between physical and genetic distance. The numbers reported between the markers indicate the number of recombination events (Nr of rec.) found in 2071 plants analysed. M18, AM19 and AL07 were the starting points of the chromosome walk. The others markers were produced during the chromosome walk.

3.3.5 Identification of candidate *Vf*-containing clones

The contig spanning the region between the M18 and AM19 markers is about 550 kb long. This region has been covered with 13 clones. To reduce the size of the interval in which the *Vf* gene could be located, the 27 recombinants of the M18 – AM19 interval were analysed with the polymorphic markers produced during the chromosome walk. Since the physical distance between M18 and AM19 was found to be 550 kb, the average distance between the crossing-overs in this interval is 20 kb. By mapping the markers of the interval it was found that the distribution of recombination events in the interval is very irregular (Fig. 6B). While three crossing-overs are located between M18 and QU97, an interval that is approximately 150 kb long, and 24 recombinants are located between M7T and AM19, which define a stretch of approximately 300 kb, no crossing-overs were identified between QU97 and M7T, the markers co-segregating with *Vf*. The distance between these last two markers is about 140 kb. However, the possible location of *Vf* is not restricted to this interval, but to the region delimited by the two closest crossing-over events flanking it. The sites of these crossing-overs are on the BAC clones M18-2 and M18-8. Therefore, *Vf* is located between the M2S marker, developed on the Sp6-end of clone M18-2, and the M8S marker, developed on the Sp6-end of clone M18-8. These are the closest markers to *Vf*, which do not co-segregate with it. The distance between these two markers is 350 kb. This region is covered by a minimum of five clones: M18-2, M18-5, M18-6, M18-7 and M18-8.

3.4 IDENTIFICATION OF RESISTANCE GENE HOMOLOGUES

3.4.1 Identification of cDNAs homologous to the tomato *Cf* resistance genes

750,000 clones of a cDNA library of leaves of Florina inoculated with *V. inaequalis* were screened with the inserts of the five minimal BAC clones that span the *Vf*-region (M18-2, M18-5, M18-6, M18-7 and M18-8, Fig. 6A). To these five BAC inserts hybridised 101 cDNAs. After sequencing from the poly-A tail, the sequences were compared. Fifty unique genes were found. These cDNAs were then sequenced from the 5'-end and both 5' and 3' ends were compared to public sequence databases. Three cDNAs (cDNAs 8-1, 10-5 and 15-4) were similar to each other and showed similarities to *Cf* resistance genes of tomato (Jones *et al.* 1994, Dixon *et al.* 1996, Thomas *et al.* 1997, Dixon *et al.* 1998). Those cDNAs were completely sequenced but unfortunately they were not full length. Using BlastX, their deduced amino acid sequences were about 35% identical to the proteins encoded by the *Cf* resistance genes. The three cDNAs proved to be members of a gene family coding for proteins with a putative transmembrane domain and extra-cellular leucine rich repeat (Tab. 3). The longest of the three cDNAs, cDNA 8-1, was used to probe a Southern blot of *Eco*RI digested BAC clones of the 550 kb BAC contig of the *Vf*-region (including clones from the "susceptible" as well as from the "resistant" chromosomes). The probe hybridised to multiple bands of DNA fragments of the contig comprising clones co-segregating with resistance in cis and in trans. The high number of bands indicated the presence of a gene cluster of many *Cf* resistance gene homologues. The members of the cluster were called *HcrVf* genes (Homologues to *Cladosporium fulvum* resistance genes of the *Vf*-region, in analogy to Parniske *et al.* 1999).

To identify the BAC fragments hybridising to cDNA 8-1, the *Eco*RI and *Sau*3AI sub-libraries of the five BAC clones were screened with the cDNA. The positive subclones were partially sequenced using the universal primers M13-forward and M13-reverse.

This part of the work has also been performed in collaboration with the group of Professor Silvano Sansavini, department of woody fruit trees (DCA), University of Bologna, Italy.

The construction of the cDNA library of simulated leaves of Florina, the screening of the cDNA library with the 5 minimal BAC clones required to span the Vf-region, the subcloning of the 5 BAC clones with two restriction enzymes EcoRI and Sau3AI and the production of high-density filters carrying the subclones has been performed by B. Vinatzer (DCA Bologna). The 101 cDNAs hybridising to the five BAC clones that he found, the two sub-libraries of the five BAC clones and their high-density filters were the starting material to produce the results reported here.

Table 3: Comparison of the amino acid sequences of the three cDNAs of *Florina* with similarities to the proteins encoded by the *Cf* resistance genes of tomato. In *Cf-9* the domain A is a signal peptide, B is the presumed NH₂-terminus of the mature protein, C is a leucine-rich repeat (LRR), D has an unknown function, E is an acidic domain, F is an hydrophobic domain (transmembrane) and G is the C-terminal basic domain. The three cDNAs show at their C-terminal regions the same domains as the *Cf* resistance genes. The cDNA 10-5 shows in the domain C two internal stop codon (marked with a star). The amino acids of the LRR consensus are indicated in red and blue.

| <i>Cf-9</i> | cDNA 8-1 | cDNA 10-5 | cDNA 15-4 |
|---|--|--|--|
| <p>LSXXLGGXGLXXLXSLXSLXGXILP</p> <p>A</p> <p>MDCVLVEFLMLLFLFQLALSSS LPHKCPDQALSLDQ KMFETINRNASDYCYDIRTYVDIY PRTLWNKSTSCSMDGVMHCDFTGQ</p> <p>B</p> <p>VLAIDLRCSQIQKFKGSUS LPLSKKREDKSENETKSLSPK FGPSNETHDLSHSEFVLIIEE LCHSKSLVIRLGDQVGLSLVYFNEELL IKMLTQRRNEESVMISSYIIS NFSSHLTPTLQSLCHLHGLLEH VHSSDQSHLSVAPELIVRPPTEK WNSASDPTVYVDSVLIARLEKS FSSHTSHEHZMGKRLSPTIEK LWETNIVPHTGDNAEYPIESH PTLRKAKLSIVAMRFDGGLER LSFWTQERLDSNLSQTPLESN ISGQNRKELTSSPHHMSALSW LPSLSEVHDISSMHPKRIQEF KSKTSAVHLKQKIKGRIEMS LAKKAKGLLASHNLSHLSA LCHLCTLLLDLGHNLKSTHLEQCV VRRBYLSDIKKRRKSTTNTT PSVGNLRAVLSHGKELTKVRS MINKYALVLDGANNMADTFHW YGYLFOKILSRGNKHCPIKSSGN TNLWGLQVDSNNGFQNLBRI LGNLQWKEIDESTGFPFETSDPY DLYZYVPTI STKGDYDSVRI LDSMRLINLSKRFEGCHFSI IGDVGARTNLSHVALECHAFAS PQM-SVLSLELSSMKI SGEI PQ LAS-ITP-LEVANLSHHLVCELEKQ</p> <p>C</p> <p>VIAICQMDPCHGVKTKLFWISF *Y*WFLAVSELDVGLRFTCSINPCVIW LICLLIFCHCCLVSEVYELPANI WNUISKVVIWVHSHNQLYQIQ PTFGAQTIVVDLSSRRFTALFI VETSQIWMIDSSNSPSSVPHFFCDR PEEPKQHLVHGHULLSKVDFDC MMNMQGVVHLENHTVIVVYVS LGYLQQLDSHGRNQLCVLEFHS LQNTSLSVVDIGNGFTOSLIMMI GKS-SKLOILNLSRNEPKDIPYE VCYLKSLALLDLARNKLSLTERCFHN LSAMNLSSESSSTRPQWGSFGIF SEVENAILVTKGTEMY SKI-LGLVGMIDLSCHFMVYGEIPEE LSS-LIAGOSLINSNNRFTGGIFSK LGNMVAQESIDPSMNIIDERTIPS MTTLEPESHINLSYNNLTERLEE</p> <p>D</p> <p>NSAPCALADFSQIESTTSFWMGVES GHTENAILLVTKGTEMY TKILGFVAGMDLSCHFMVYGEIPEE LTIGLLELQSVLINSNNRFTGGIFSK IGSMVAQESIDPSMNIIDERTIPS MVKLTFLESHINLSYNNLTERLEES</p> | <p>LSXXLGGXGLXXLXSLXSLXGXILP</p> <p>A</p> <p>MDCVLVEFLMLLFLFQLALSSS LPHKCPDQALSLDQ KMFETINRNASDYCYDIRTYVDIY PRTLWNKSTSCSMDGVMHCDFTGQ</p> <p>B</p> <p>VIAIDLRCSQIQKFKGSUS LPLSKKREDKSENETKSLSPK FGPSNETHDLSHSEFVLIIEE LCHSKSLVIRLGDQVGLSLVYFNEELL IKMLTQRRNEESVMISSYIIS NFSSHLTPTLQSLCHLHGLLEH VHSSDQSHLSVAPELIVRPPTEK WNSASDPTVYVDSVLIARLEKS FSSHTSHEHZMGKRLSPTIEK LWETNIVPHTGDNAEYPIESH PTLRKAKLSIVAMRFDGGLER LSFWTQERLDSNLSQTPLESN ISGQNRKELTSSPHHMSALSW LPSLSEVHDISSMHPKRIQEF KSKTSAVHLKQKIKGRIEMS LAKKAKGLLASHNLSHLSA LCHLCTLLLDLGHNLKSTHLEQCV VRRBYLSDIKKRRKSTTNTT PSVGNLRAVLSHGKELTKVRS MINKYALVLDGANNMADTFHW YGYLFOKILSRGNKHCPIKSSGN TNLWGLQVDSNNGFQNLBRI LGNLQWKEIDESTGFPFETSDPY DLYZYVPTI STKGDYDSVRI LDSMRLINLSKRFEGCHFSI IGDVGARTNLSHVALECHAFAS PQM-SVLSLELSSMKI SGEI PQ LAS-ITP-LEVANLSHHLVCELEKQ</p> <p>C</p> <p>NSAPCALADFSQIESTTSFWMGVES GHTENAILLVTKGTEMY TKILGFVAGMDLSCHFMVYGEIPEE LTIGLLELQSVLINSNNRFTGGIFSK IGSMVAQESIDPSMNIIDERTIPS MVKLTFLESHINLSYNNLTERLEES</p> | <p>LSXXLGGXGLXXLXSLXSLXGXILP</p> <p>A</p> <p>MDCVLVEFLMLLFLFQLALSSS LPHKCPDQALSLDQ KMFETINRNASDYCYDIRTYVDIY PRTLWNKSTSCSMDGVMHCDFTGQ</p> <p>B</p> <p>VIAIDLRCSQIQKFKGSUS LPLSKKREDKSENETKSLSPK FGPSNETHDLSHSEFVLIIEE LCHSKSLVIRLGDQVGLSLVYFNEELL IKMLTQRRNEESVMISSYIIS NFSSHLTPTLQSLCHLHGLLEH VHSSDQSHLSVAPELIVRPPTEK WNSASDPTVYVDSVLIARLEKS FSSHTSHEHZMGKRLSPTIEK LWETNIVPHTGDNAEYPIESH PTLRKAKLSIVAMRFDGGLER LSFWTQERLDSNLSQTPLESN ISGQNRKELTSSPHHMSALSW LPSLSEVHDISSMHPKRIQEF KSKTSAVHLKQKIKGRIEMS LAKKAKGLLASHNLSHLSA LCHLCTLLLDLGHNLKSTHLEQCV VRRBYLSDIKKRRKSTTNTT PSVGNLRAVLSHGKELTKVRS MINKYALVLDGANNMADTFHW YGYLFOKILSRGNKHCPIKSSGN TNLWGLQVDSNNGFQNLBRI LGNLQWKEIDESTGFPFETSDPY DLYZYVPTI STKGDYDSVRI LDSMRLINLSKRFEGCHFSI IGDVGARTNLSHVALECHAFAS PQM-SVLSLELSSMKI SGEI PQ LAS-ITP-LEVANLSHHLVCELEKQ</p> <p>C</p> <p>NSAPCALADFSQIESTTSFWMGVES GHTENAILLVTKGTEMY TKILGFVAGMDLSCHFMVYGEIPEE LTIGLLELQSVLINSNNRFTGGIFSK IGSMVAQESIDPSMNIIDERTIPS MVKLTFLESHINLSYNNLTERLEES</p> | <p>LSXXLGGXGLXXLXSLXSLXGXILP</p> <p>A</p> <p>MDCVLVEFLMLLFLFQLALSSS LPHKCPDQALSLDQ KMFETINRNASDYCYDIRTYVDIY PRTLWNKSTSCSMDGVMHCDFTGQ</p> <p>B</p> <p>VIAIDLRCSQIQKFKGSUS LPLSKKREDKSENETKSLSPK FGPSNETHDLSHSEFVLIIEE LCHSKSLVIRLGDQVGLSLVYFNEELL IKMLTQRRNEESVMISSYIIS NFSSHLTPTLQSLCHLHGLLEH VHSSDQSHLSVAPELIVRPPTEK WNSASDPTVYVDSVLIARLEKS FSSHTSHEHZMGKRLSPTIEK LWETNIVPHTGDNAEYPIESH PTLRKAKLSIVAMRFDGGLER LSFWTQERLDSNLSQTPLESN ISGQNRKELTSSPHHMSALSW LPSLSEVHDISSMHPKRIQEF KSKTSAVHLKQKIKGRIEMS LAKKAKGLLASHNLSHLSA LCHLCTLLLDLGHNLKSTHLEQCV VRRBYLSDIKKRRKSTTNTT PSVGNLRAVLSHGKELTKVRS MINKYALVLDGANNMADTFHW YGYLFOKILSRGNKHCPIKSSGN TNLWGLQVDSNNGFQNLBRI LGNLQWKEIDESTGFPFETSDPY DLYZYVPTI STKGDYDSVRI LDSMRLINLSKRFEGCHFSI IGDVGARTNLSHVALECHAFAS PQM-SVLSLELSSMKI SGEI PQ LAS-ITP-LEVANLSHHLVCELEKQ</p> <p>C</p> <p>NSAPCALADFSQIESTTSFWMGVES GHTENAILLVTKGTEMY TKILGFVAGMDLSCHFMVYGEIPEE LTIGLLELQSVLINSNNRFTGGIFSK IGSMVAQESIDPSMNIIDERTIPS MVKLTFLESHINLSYNNLTERLEES</p> |
| <p>D</p> <p>KQDFDFGNISYQGNLGLRGLFSLKICGG EDQVTTVAELDQEEEEED SFMISWQGLVGYGCGGLVIGLSVI YIMWSTQTPAMFS RMDLKLREHLLITPKKHKHKRY</p> <p>E</p> <p>BDQVTTVAELDQEEEEED</p> <p>F</p> <p>SFMISWQGLVGYGCGGLVIGLSVI YIMWSTQTPAMFS</p> <p>G</p> <p>RMDLKLREHLLITPKKHKHKRY</p> | <p>D</p> <p>STQLQSLNQSSFEVGNELCGAPLANKCS ANGVVPPPTVEQDGGGGYCLLEDE WFYVSLGVGFTTGFWIVLGSLLVN MWSHLLSGLLN</p> <p>E</p> <p>ANGVVPPPTVEQDGGGGYCLLEDE</p> <p>F</p> <p>WFYVSLGVGFTTGFWIVLGSLLVN MWSHLLSGLLN</p> <p>G</p> <p>RIVLKMIVIVEYV*</p> | <p>D</p> <p>STQLQSLNQSSFEVGNELCGAPLANKCS ANGVVPPPTVEQDGGGGYCLLEDE WFYVSLGVGFTTGFWIVLGSLLVN MWSHLLSGLLN</p> <p>E</p> <p>ANGVVPPPTVEQDGGGGYCLLEDE</p> <p>F</p> <p>WFYVSLGVGFTTGFWIVLGSLLVN MWSHLLSGLLN</p> <p>G</p> <p>RIVLKMIVIVEYV*</p> | <p>D</p> <p>TQLQSLDQSSFEVGNELCGAPLANKCS TNGVTPPTVEQDGGGGYRLEDE WFYVSLGVGFTTGFWIVLGSLLVN MWSHLLSGLLN</p> <p>E</p> <p>TNGVTPPTVEQDGGGGYRLEDE</p> <p>F</p> <p>WFYVSLGVGFTTGFWIVLGSLLVN MWSHLLSGLLN</p> <p>G</p> <p>RIVLKMIVIVEYV*</p> |

The comparison of the subclone sequences with those of the three cDNAs revealed that some subclone sequences were homologous, but not identical to the three cDNAs, thus indicating that the cDNAs were not from the *Vf*-region in cis with the resistance. The comparison of the positive subclone sequences in public sequence databases revealed that some sequences showed the same homologies to the *Cf* genes as the three cDNAs.

3.4.2 Development of specific primers for the *HcrVf* gene family (*fv* primers) and identification of *HcrVf1* in the contig

Due to the high sequence homology and, in most cases insufficient overlap of the *EcoRI* and *Sau3AI* BAC subclones hybridising to the cDNA 8-1, it was almost impossible to assemble the subclones to find complete ORFs. For a correct assembling of the subclones it was necessary to develop a pair of primers that could amplify the largest possible part of the *HcrVf* genes. The comparison of the sequence of this PCR product to the sequences of the subclones would allow the correct aligning of the subclones. Comparing the sequences of three cDNAs (cDNAs 8-1, 10-5 and 15-4) to each other, it was found that the most conserved part of the sequence was located just before the stop codon. In this region the *fv*-rev primer was developed. The *fv*-for primer was designed on the genomic sequence of a BAC subclone that encoded for a LRR showing high homology with tomato *Cf* genes (Fig. 7).

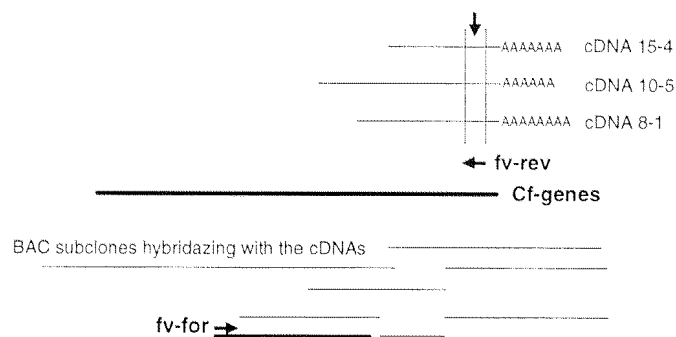


Figure 7: Development of *fv* primers. The *fv*-rev primer was developed in the most conserved sequence (region of the stop codon, indicated with a vertical arrow) among the three cDNAs hybridising in the *Vf*-region, presenting similarity with the *Cf* resistance genes. The second specific primer (*fv*-for) was designed on the sequence of a BAC subclone with homology in the 5' end of the *Cf* genes (LRR domain).

The fv primers were first used on the five BAC clones spanning the *Vf*-region. A single band of 1800 bp was obtained from M18-2 and M18-5, and two bands (1900 and 2700 bp long) were obtained from M18-6 (M18-6A and M18-6B). The four bands were cloned and 12 transformants per ligation were re-amplified, using the fv primers. The PCR products were digested with four restriction enzymes (*TaqI*, *Sau3AI*, *AluI* and *HeaIII*). Three different fv-PCR products were found. The bands obtained from M18-2 and M18-5 proved to be identical and for this reason it was assumed that they derived from the overlapping part of clone M18-2 and M18-5 (this band was called M18-2/5A) (Fig. 8). As expected, the comparison of the three fv-PCR product sequences to public databases revealed the same homologies to the *Cf* resistance genes found with the three cDNAs.

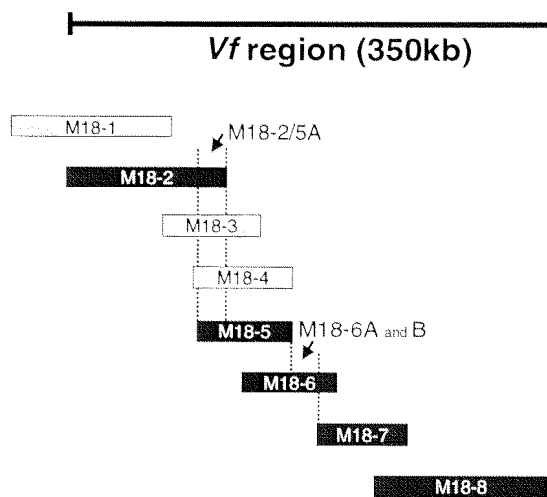


Figure 8: Relative position of the three fv-PCR products. The eight boxes represent the inserts of the eight BAC clones that span the *Vf*-region. The names of the BAC clones are written in the boxes. Black boxes represent the five minimal BAC clones necessary to cover the *Vf*-region. M18-2/5A, M18-6A and M18-6B are the three fv-PCR products obtained from BAC clone M18-2 (and M18-5) and M18-6 respectively, performing a PCR with the fv primers. Their positions on the contig are indicated with arrows.

The fv primers were then used in RT-PCR experiments to check if they would amplify expressed genes. RT-PCR experiments were performed using three bulks (one for each cultivar) of total RNA extracted from leaves of Florina (*Vf*), Prima (*Vf*) and Golden Delicious harvested at different times after inoculation with *V. inaequalis*.

Vf candidate genes are expected in the two *Vf* cultivars, Florina and Prima, and they should be absent in the susceptible cultivar Golden Delicious.

The RT-PCR produced two bands (named L and S) present in all cultivars and a third band (M band) that was observed only in the *Vf* resistant cultivars (Fig. 9A). All three bands have been cloned and the digestion of the amplified cloned fragments with several restriction enzymes allowed the identification of three different clones containing the band M (two from Prima and one from Florina). The same test showed that in all three cultivars a mixture of PCR products of the same size composed the S band. In contrast the L band was composed of a single PCR product identical in all three apple cultivars. Sequence comparison of the three clones containing the M band showed that they were almost identical (> 99 % identity, with only a few nucleotides that were different). A consensus sequence was generated and its comparison to partial sequences of BAC subclones hybridising to the cDNA 8-1 allowed the identification of two *Eco*RI subclones of M18-2 (about 3000 bp and 6000 bp long) with a 100 % match. The complete sequencing of the 3000 bp subclone, the extension of the partial sequence of the 6000 bp subclone and the assembly of the sequences led to the identification of the first ORF, *HcrVf1* (Fig. 9B and Tab. 4), possessing all the domains present in *Cf* resistance genes. Since the consensus sequence was generated from cloned RT-PCR bands, the expression of *HcrVf1* was automatically proven.

Putative *Vf* resistance genes have to fulfil three requirements: a) should map in the *Vf*-region, b) should be expressed and possibly similar to an already cloned resistant gene, and c) should be differentially expressed between *Vf* and non-*Vf* resistant cultivars. *HcrVf1* fulfilled the last two requirements, but the map position remained to be clarified. Several attempts were performed using different sets of *HcrVf1* specific primers, but although primers were developed in regions absent or different in the two *HcrVfs* found later, no polymorphic marker was found; thus it was not possible to map *HcrVf1* more precisely than its position on BAC clone M18-2 (Fig. 9C).

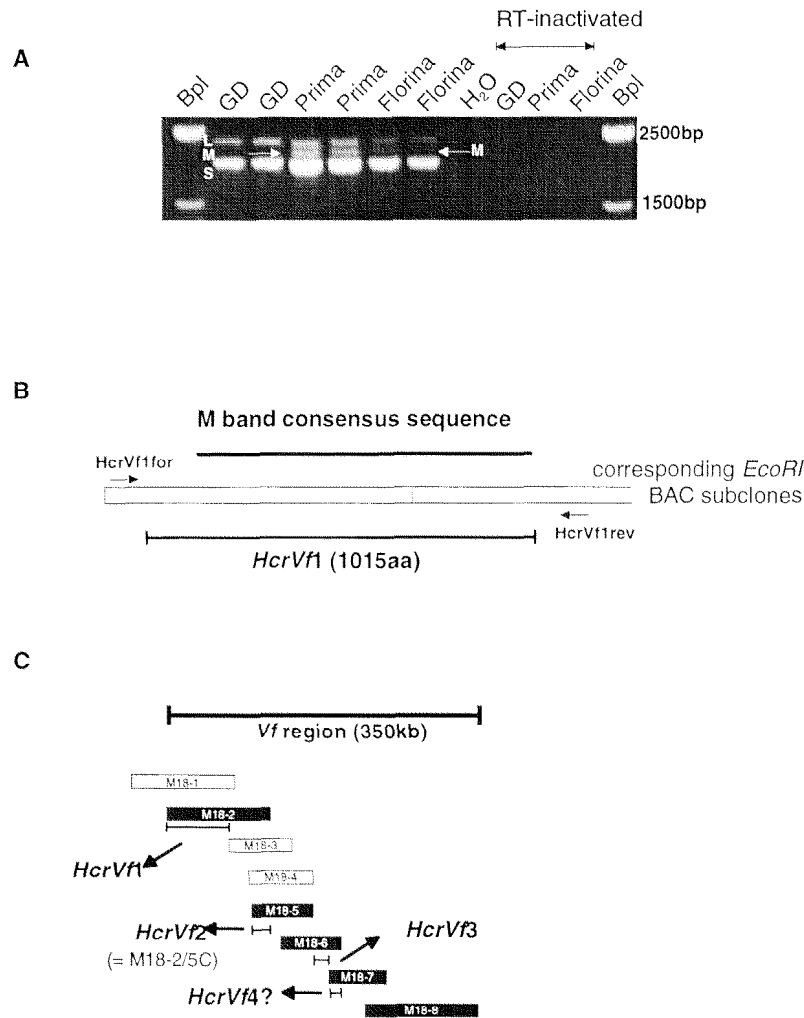


Figure 9: **A** RT-PCR bands obtained using the fv primers (specific primers for the *HcrVf* genes) with three mixtures of total RNA extracted from leaves, inoculated with *V. inaequalis* conidia, harvested at different times. In all the three cultivars tested; Golden Delicious (GD), Prima and Florina, the bands L and S were amplified, but in Prima and Florina (*Vf* resistant) a supplementary band (M) was amplified. To check the absence of DNA contamination in the extracted RNA, three separate RT-PCR reactions, one for each cultivar total RNA mixture, were performed with inactivated RT. No amplification was obtained indicating the absence of DNA in the extracted total RNA

B Assembling of *HcrVf1* full genomic sequence. The band M was cloned and a consensus sequence was generated. The comparison of the consensus sequence M with the BAC subclones hybridising to the cDNA 8-1 allowed the identification of two subclones with 100 % match. The translation of the assembled sequences of the two subclones permitted the identification of a first ORF (*HcrVf1*). A pair of primers, HcrVf1for and HcrVf1rev, was developed up-stream and down-stream from the ORF of *HcrVf1*. These primers permitted the amplification of the complete ORF of *HcrVf2* and *HcrVf3* from the BAC clones M18-5 and M18-6 respectively.

C Relative position of *HcrVf1*, *HcrVf2*, *HcrVf3* and *HcrVf4* on the contig spanning the *Vf*-region. The horizontal lines indicate the approximate position of the four *HcrVfs*. The position of *HcrVf4* on the contig is indicated with a question mark because it is based only on preliminary results.

N.B. *HcrVf3* is different from M18-6A and B obtained with fv primers (Fig. 8).

3.4.3 Identification of *HcrVf2*, *HcrVf3* and *HcrVf4*

The sequencing of the three different cloned fv-PCR products; M18-2/5A, M18-6A and M18-6B, obtained with the fv primers from the BAC clones M18-2 and M18-6 and the following translation led to the identification of a second putative ORF (*HcrVf2*). *HcrVf2* was found translating the sequence obtained from M18-2/5A. The other two sequences obtained from M18-6A and M18-6B were similar to *HcrVf1* and *HcrVf2* but presented a truncated ORF.

The comparison of *HcrVf2* partial sequence (1800bp) with the sequences of the subclones hybridising to the cDNA 8-1 allowed the identification of four BAC *EcoRI* subclones and four BAC *Sau3AI* subclones originating from M18-2 and M18-5, confirming the position of *HcrVf2* in the overlapping region of these two BAC clones. Thanks to the known sequence of the cloned fv-PCR product M18-2/5A and the overlapping of some *EcoRI* and *Sau3AI* BAC subclones, the assembling of the corresponding genomic sequence was possible (Fig. 10).

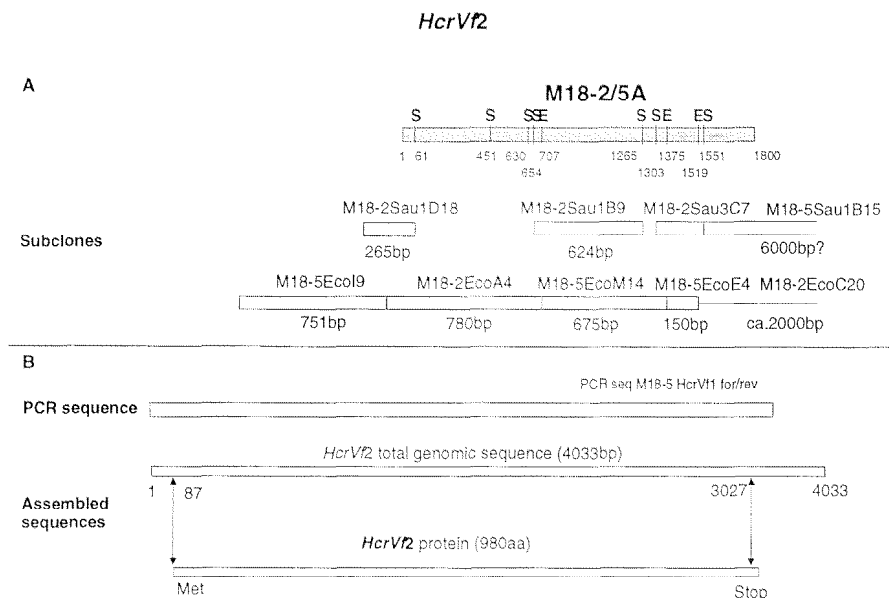


Figure 10: Assembling of *HcrVf2* sequence. **A** Relative position of *EcoRI* and *Sau3AI* subclones composing *HcrVf2*. The *EcoRI* and *Sau3AI* subclones corresponding to *HcrVf2* are indicated with horizontal white bars. Their names and their insert sizes are written above and below the bars respectively. The fv-PCR product (M18-2/5A) that allowed the alignment of the subclones is indicated with a grey bar. **B** Result of the assembling of BAC subclone sequences and of the PCR sequence obtained with the *HcrVf1*for/rev primers with the BAC clone M18-5. The assembled sequence of 4033bp contains an ORF of 980 amino acids.

To further extend the genomic sequence at the 5' end of the assembled sequence, a probe in the overhanging part of the *Sau3AI* BAC subclone insert (M18-2*Sau1D18*) was used to screen the M18-2 and M18-5 *EcoRI* BAC subclone library. One *EcoRI* subclone was found (M18-5*EcoI9*), thus allowing the extension of the genomic sequence. The same strategy was applied to find the corresponding *Sau3AI* BAC subclone with no success.

To identify the putative start of *HcrVf2*, another strategy was used. Comparing *HcrVf1* genomic and *HcrVf2* partial genomic sequences, it was noticed that the two genes were not only similar in the ORFs but also down-stream from the stop codon. In the 150 nucleotides down-stream from the stop codon, the two ORFs differed by just two nucleotides. Speculating that also up-stream from the presumed transcription start of *HcrVf1* a sequence homology between the two ORFs would be present, a pair of primers was developed up- and down-stream from the *HcrVf1* coding sequence (*HcrVf1for* and *rev*, Fig. 9B). The primers were tested on the BAC clones of the contig and amplifications were obtained from M18-2, M18-5 and M18-6 (one band each). *EcoRI* restriction digestion of the PCR product obtained from M18-2 confirmed the amplification of *HcrVf1* but indicated also the amplification of a second product, which showed some of the *EcoRI* fragments expected for *HcrVf2*. On the contrary, the *EcoRI* digested PCR product, obtained from M18-5, showed only the fragments expected from *HcrVf2*. The sequencing of this latter cloned PCR product confirmed that this product was *HcrVf2*, that the BAC subclones were assembled correctly and allowed obtaining the presumed complete ORF of *HcrVf2*.

The sequence analysis of the cloned PCR product obtained from M18-6 allowed the identification of a third ORF that was called *HcrVf3* (Fig. 9C).

The comparisons of the amino acid sequences of *HcrVf1*, *HcrVf2* and *HcrVf3* with those of *Cf-9* indicated that the *HcrVf* ORFs should be complete, because all seven domains present in *Cf-9* protein were also identified in the predicted *HcrVf* proteins (Tab. 4).

To investigate the expression of *HcrVf2* and *HcrVf3* (the expression of *HcrVf1* had already been demonstrated) RT-PCR experiments were performed as described for *HcrVf1*. The comparison of the genomic sequences of *HcrVf1*, *HcrVf2* and *HcrVf3* permitted the finding of regions specific to *HcrVf2* and *HcrVf3* where specific primers were developed. The result of the RT-PCRs for both *HcrVfs* was a single band present in Florina and Prima that was absent in Golden Delicious. The RT-PCR sequences of the *HcrVf2* specific products obtained from Florina and Prima were identical. The comparison of the RT-PCR sequences with the genomic sequence of *HcrVf2* showed a 100% match. It was therefore proven that the *HcrVf2* is expressed only in *Vf* cultivars and is absent in "non-*Vf*" plants. The mapping of *HcrVf2* proved its co-segregation with the resistance.

As was the case for *HcrVf2*, the sequences of the RT-PCR products of Florina and Prima obtained with *HcrVf3* specific primers were identical. The comparison of these RT-PCR sequences with the *HcrVf3* genomic sequence revealed that the sequences were very similar (97 % identity) but not exactly the same. Therefore, it was not possible to prove that *HcrVf3* is expressed. However, the results of this last RT-PCR experiment led to identification of a fourth ORF that was called *HcrVf4*. The comparison of this new RT-PCR sequence with the partial sequences of the BAC subclones permitted the identification of an *EcoRI* subclone of about 4500bp that spans almost the full ORF. *HcrVf3* and *HcrVf4* were not mapped because they lay on the BAC clone M18-6, which co-segregates with *Vf*.

The constitutive expression of *HcrVf2* and *HcrVf4* could be proven by RT-PCR performed on total RNA extracted from leaves of Florina and Prima harvested just after inoculation (0 h), performed using *HcrVf2* and *HcrVf3* specific primers. The same experiment repeated using the *fv* primers demonstrated the constitutive expression of *HcrVf1*.

3.4.4 RACE (Rapid Amplification of cDNA Ends)

The 5'-RACE methodology was used to find the 5'-end of the ORFs previously identified including the transcription and translation starting points. First strand cDNA was synthesised from total RNA using an antisense gene-specific primer (SP1) and AMV reverse transcriptase. Terminal transferase was then used to add a homopolymeric A-tail to the 3'-end of the cDNA. Tailed cDNA was amplified by PCR using an antisense gene-specific primer (SP2) and an oligo dT-anchor primer. The obtained cDNA was further amplified by a second PCR using a nested, antisense gene-specific primer (SP3) and the PCR anchor primer. The result of this procedure is the amplification of the 5' gene extremity.

The three gene specific primers were developed at 829 bases (SP1), at 346 bases (SP2) and at 273 bases (SP3) from the putative translation start. All three primers were developed in a common region of *HcrVf1* and *HcrVf2*. *HcrVf3* and *HcrVf4* were not considered (*HcrVf3* is not expressed and *HcrVf4* is not completely characterised). The first PCR amplification did not produce any visible band while after the second one, nine bands of different length appeared. The PCR product mix was cloned and the transformants were analysed by PCR re-amplification of the cloned bands and digestion of the PCR products with *Bam*HI (comparing *HcrVf1* and *HcrVf2* genomic sequences a *Bam*HI restriction polymorphism was found).

Clones whose re-amplified insert was cut by *Bam*HI were candidates to be the 5' end of *HcrVf2*; all the others were candidates to be *HcrVf1*. The comparison of the sequences of the cloned RACE bands with *HcrVf1* genomic sequences revealed no match for *HcrVf1*. On the contrary clones C2 and D2 gave a perfect match with 280 nucleotides of *HcrVf2* genomic sequence (Fig. 11).

A perfect match was observed in the first 244 bases of the ORF and in the 36 bases up-stream from the putative translation start, indicating that the leader sequence of the gene might be 36 bases long. After these 36 bases a 26 bases long CT rich sequence was found. The presence of these 26 bases, not matching the genomic sequence of *HcrVf2*, could be explained in different ways. It is possible that even if the sequences of the two RACE clones present a perfect match with the genomic sequence of *HcrVf2* over 280 bases, these clones might not derive from the transcription of *HcrVf2*, but they could derive from another *HcrVf* gene with high homology to *HcrVf2*. A second hypothesis to explain such a sequence mismatch could be the

presence of an intron in the leader sequence. However, the absence of the AG nucleotides maintained at the 3'-end of all dicotyledonous introns does not seem to support this hypothesis. According to Roche customer support it is also possible that such a CT rich sequence could be a PCR artefact due to the slippage of the poly-T primer during the amplification. From the data presently available it is not possible to decide which of the hypotheses is the correct one. Although, the presence of a 280 perfect match between the transcript and *HcrVf2* genomic sequence strongly supports the idea that C2 and D2 are in fact the 5'-end of the *HcrVf2* transcript.

The analysis of the genomic sequence of *HcrVf2* up-stream from the presumed translation start permitted the identification of two transcription regulatory sequences. A putative TATA-box (sense strand sequence ATATATA) was found 26 bases from the presumed transcription start and a CAAT-box (sense strand sequence GTTTA) was found 22 bases from the TATA-box. The distance of 26 bases between the putative TATA-box and the transcription start (i.e. beginning of the match between the RACE sequence and the genomic sequence) is in good agreement with average distances observed in higher organisms. All these data indicate that very probably *HcrVf2* is complete and through homology also *HcrVf1* and *HcrVf3*.

```

                                                    CAAT-Box
HcrVf2 (sense strand)                GCGGTCTCATGCCGTAGAGGATGGTTTAGAAA

                                     26 bases long CT rich sequence
C2          -----CTTCTTCTCTCTTCTTCCCTGTTCTTGGCTCATAAC
D2          -----CTTCTTCTCTCTTCTTCCCTGTTCTTGGCTCATAAC
HcrVf2     AGGTGAGGCTCCCTAGTATATATAGGTAGCACCTCCTCTGGATTGTTTCATGCTCATAAC
                                     TATA-BOX   * * * * *
                                     +1 Met1      Met2

C2          ATCACCTGCTTCATAACACTTGAACCTTTATGGGAGAGAACCATGAGAGTTGTTATACTACTA
D2          ATCACCTGCTTCATAACACTTGAACCTTTATGGGAGAGAACCCGTGAGAGTTGTTATACTACTA
HcrVf2     ATCACCTGCTTCATAACACTTGAACCTTTATGGGAGAGAACCATGAGAGTTGTTATACTACTA
                                     *****
                                     +1 Met1      Met2

C2          ACCAGGTTTCTAGCCATTGCAACCATTACTTTTCAGTATTGGTTTATGCAATGGAAATCCC
D2          ACCAGGTTTCTAGCCATTGCAACCATTACTTTTCAGTATTGGTTTATGCAATGGAAATCCC
HcrVf2     ACCAGGTTTCTAGCCATTGCAACCATTACTTTTCAGTATTGGTTTATGCAATGGAAATCCC
                                     *****

C2          GGTGGCCTCCACTTTTGCAAAGTAAGCGAAAGACGAGCACTTCTGATGTTCAAGCAAGAT
D2          GGTGGCCTCCNCTTTTGCAAAGTAAGCGAAAGACGAGCACTTCTGATGTTCAAGCAAGAT
HcrVf2     GGTGGCCTCCACTTTTGCAAAGTAAGCGAAAGACGAGCACTTCTGATGTTCAAGCAAGAT
                                     *****

C2          CTCAAGGATCCTGTCATCGGGCTTGCATCGTGGGTTGCAGAAGAAGATTCAGACTGTTGC
D2          CTCAAGGATCCTGTCATCGGGCTTGCATCGTGGGTTGCAGAAGAAGATTCAGACTGTTGC
HcrVf2     CTCAAGGATCCTGTCATCGGGCTTGCATCGTGGGTTGCAGAAGAAGATTCAGACTGTTGC
                                     *****
                                     BamHI

C2          AGTTGGACAGGAGTTGTTTGTGATCATGTAA-----
D2          AGTTGGACAGGAGTTGTTTGTGATCATGTAA-----
HcrVf2     AGTTGGACAGGAGTTGTTTGTGATCATGTAAACCGGCCACATCCACGAGCTGCACCTTAA
                                     *****

```

Figure 11: Alignment of *HcrVf2* genomic sequence (sense strand) with two sequences of cloned RACE products (C2 and D2). The perfect match between two cloned RACE products and the genomic sequence of *HcrVf2* over 244 bp in the presumed ORF and 36 bp up-stream from the presumed first methionine indicate that most probably the transcription start of the gene has been reached (Met1 or Met2 indicated in bold). Putative TATA-box and CAAT box are indicated in bold. The *Bam*HI restriction polymorphism between *HcrVf1* and *HcrVf2* is indicated.

3.4.5 Analysis and comparison of *HcrVf1*, *HcrVf2* and *HcrVf3* ORFs

From the region of the *HcrVf1*, *HcrVf2* and *HcrVf3* it was possible to sequence 4362 bp, 3124 bp and 2937 bp respectively. The three regions contain a single major open reading frame (ORF) encoding polypeptides of 1015, 980 and 916 amino acids (Tab. 4). The *HcrVfs* predicted proteins show similar structures to those observed for the *Cf* resistance genes (Jones *et al.* 1994, Dixon *et al.* 1996, Thomas *et al.* 1997, Dixon *et al.* 1998).

According to sequence data a signal peptide and its cleavage site were predicted to be present at the N-terminus of the three proteins (domain A). The B domain would thus be the N-terminus of the mature protein, once the signal peptide has been cleaved. The three C domains of the *HcrVfs* (1, 2 and 3) contain 30, 29 and 26 imperfect LRR with an average length of 24 amino acids each, in good agreement with the consensus sequence LxxLxxLxLSxNxLxGxIP. As for the *Cf* proteins, no function could be assigned to the D domain, composed of 27 amino acids. The F domain, composed of 36 hydrophobic amino acids, is predicted to be the transmembrane domain of the protein. E and G domains, flanking the hydrophobic domain, are rich in polar amino acids, consistent with the predicted role of orientation and anchoring of the protein to the cell membrane. 26, 23 and 23 putative N-glycosylation sites are present in the domains B to D of *HcrVf1*, *HcrVf2* and *HcrVf3* respectively. As for the *Cf* resistance genes, these genes most probably encode for extracytoplasmic glycoproteins anchored to the cell membrane with the majority of the extracytoplasmic domain constituted of LRRs motifs (Tab. 4).

At the amino acid level the three complete ORFs, compared two by two, show on average about 87 % identity (GCG software package, function BESTFIT with gap penalty of 8 and gap extension of 2). The most conserved region among the three ORFs is the A domain and the region from D to G, presenting on average 95 % and 96 % identity respectively. The more variable domains are the domains B and C, presenting on average 84 % and 86 % identity respectively. Furthermore, it is possible to see that within the C domain the variability is not equally distributed. The C1 and C2 sub-regions are shown in Table 5. When compared, the three *HcrVfs* present an average identity of 89 % in the sub-domain C1, while the identity decreases to 82 % in C2. The lower level of identity in C2 it is not only due to difference in the number of LRRs (with reference to *HcrVf1*, *HcrVf2* lacks two, and *HcrVf3* lacks five LRRs),

3.4.6 Hybridisation of *HcrVf2* to the *EcoRI* digested BAC clones of the *Vf* contig

The hybridisation of *HcrVf2* to the complete *EcoRI* digestions of the "resistant" and "susceptible" BAC clones of the *Vf*-region allowed the identification of other loci with homologies to the sequenced *HcrVfs*. Knowing the reciprocal overlapping of the BAC clones (Fig. 6A), the fragments corresponding to *HcrVf1* and *HcrVf2*, the number of fragments expected for *HcrVf3* and for *HcrVf4* and assuming an average *HcrVfs* gene size of 3 kb, it is possible to estimate the minimal and the maximal number of *HcrVf* loci present in the region. For the minimal number of loci it is assumed that only one *HcrVf* is present in each fragment larger than 3 kb that hybridises to *HcrVf2*. For the estimation of the maximal number of loci the maximum possible number of genes of 3 kb for each fragment that hybridised to the probe are calculated. For example it is assumed that a 7 kb fragment will carry two loci. In this way in the contig part between M18-2 to M18-8 (350 kb) a minimum of 5 and a maximum of 14 loci can be postulated. On the BAC clones M18S-1, M18S-3, M18S-4 and M18S-5 (BAC clones from the homologous "susceptible" chromosome) a number of loci between seven and 15 can be hypothesised (Fig. 12). Another two *HcrVf* loci were found on adjacent BAC clones AM19-1 to 4 in cis with the resistance (data not shown).

In summary, in the BAC clones of the contig in cis and trans with the resistance, spanning 550 kb and 270 kb respectively, at least seven loci with homologies to *HcrVf2* can be hypothesised for each contig. In total, in the Florina *Vf*-region, a minimum number of 14 *HcrVf* loci can be presumed.

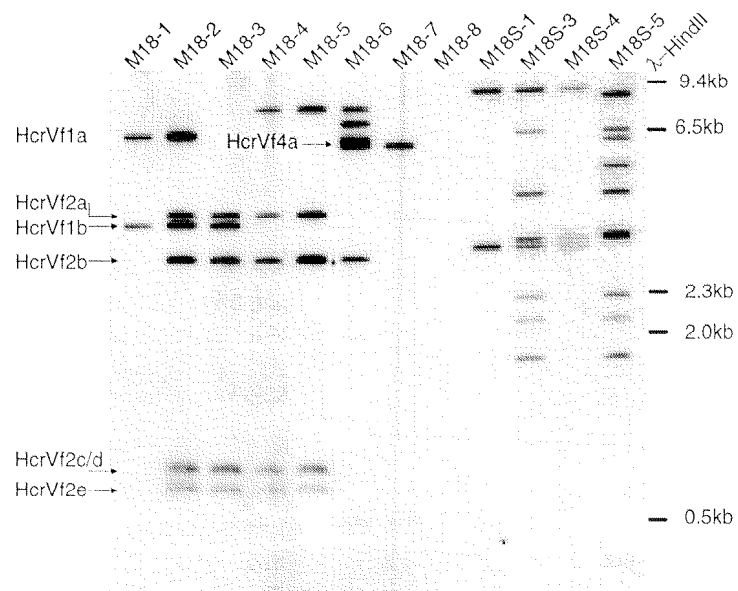


Figure 12: Hybridisation of *HcrVf2* with *EcoRI* digested BAC clones spanning the *Vf*-region. The BAC clones M18-1 to 8 contain DNA inserts from the chromosome carrying in cis with *Vf*, while the BAC clones M18S-1 to 5 contain inserts in trans with the resistance. The arrows indicate the fragments corresponding to the *HcrVf1*, *HcrVf2* and *HcrVf4*. The fragments of *HcrVf3* are not known, but from its restriction map it is known that two hybridising bands of M18-6 should contain its sequence. Of *HcrVf4* only one fragment is known.

4. DISCUSSION

4.1 FINE MAPPING OF *Vf*

The *Vf* gene, derived from *Malus floribunda* 821, is the most widely used source of resistance in apple breeding programs. *Vf* is considered to be a single dominant gene, but minor genes influence its resistance reaction. In fact, plants carrying *Vf* show very different phenotypes ranging from a complete absence of symptoms to symptoms such as chlorotic spots and even limited sporulation of *V. inaequalis* (Gardiner *et al.* 1996). Conversely, plants on which *V. inaequalis* sporulate heavily are considered not to carry the *Vf* gene. Due to the lack of a clear difference between the phenotypes of plants with or without *Vf*, the classification of the genotypes of some plants, in populations segregating for *Vf*, cannot be unambiguously identified. The uncertain resistance scores of these plants led to problems in mapping the gene. In fact, although *Vf* has been mapped by four different groups (Gessler *et al.* 1995, Gardiner *et al.* 1996, Hemmat *et al.* 1998, Tartarini *et al.* 1999) to the same chromosomal region, there is as yet no agreed map position (Fig. 3B).

During the analysis of the populations Braeburn x FAW 7167 (*Vf*) and Fuji x Ariwa (*Vf*), about 9% of plants classified as resistant amplified the alleles of M18 and AL07 markers in repulsion with *Vf*. After a second check in the field, it was possible to correct the classification (from resistant to susceptible plants) of 34 of these 70 plants. Twenty-four plants were dwarfs or had a reduced growth characteristic. Due to the slow and reduced growth, the leaves of dwarf plants present mostly ontogenic resistance and can not be successfully infected during the greenhouse test. A second explanation is that seedlings with reduced growth habits are shielded by the normal growing seedling and are not inoculated with the sprayed conidial suspension during the greenhouse test. In fact, in the field some dwarf plants showed clear sporulating lesions, confirming this hypothesis. However, without repeating the inoculation test on the dwarf plants which did not show clear sporulation and on the remaining 10 plants (two plants died), it is impossible to correct the classification of these plants. Nevertheless, for three reasons these plants are probably classified incorrectly. If these plants are classified correctly they must be double recombinants, but this is quite unlikely. Given a genetic distance of 0.5 cM between M18 and *Vf* and a genetic

distance of 0.9 cM between AL07 and *Vf* as reported by Tartarini *et al.* (1999), the probability of double crossing-over events cutting out the resistance gene should be 9×10^{-5} . The value of 4.5 % (36 plants out of 758) observed in our populations is significantly different (about 500 times greater) from the expected. The second indirect validation of this hypothesis comes also from the fact that if the scoring of *Vf* of these plants would be considered correct the markers AL07 and M18 should map very far from *Vf*. This would be in disagreement with our Florina x Nova Easygro data and also with works previously published by King *et al.* (1998), Maliepaard *et al.* (1998), Tartarini *et al.* (1996), Gardiner *et al.* (1996) and Hemmat *et al.* (1998). The third indication that the plants are very probably incorrectly classified for the resistance is given by the analysis of the 70 plants with a set of polymorph markers (internal to the *Vf*-region) developed during the chromosome walk (QU97, M5S, M7T and M8S, Fig. 6). All 70 plants amplified the marker alleles in repulsion with the resistance with all the markers tested, thus indicating that if these plants are considered double recombinants, the two recombination events of all the plants must be in the region between two marker next to each other. Because the physical distances between two markers range from about 50 to 100 kb this is highly improbable. This check seems to exclude the possibility that these plants can carry the *Vf* resistance.

If a mistake in the classification of the resistance was not made and therefore those plants do not contain *Vf*, the only other explanation that can be proposed is the presence of other resistance genes. The existence of such resistance genes with additive effect has been reported by Seglias (1997) with QTL mapping experiments. If this is the case, because the resistance phenotype of these plants is not due to the *Vf* gene, it is correct and necessary not to consider these plants for the mapping of *Vf*.

A similar problem has already been reported by Gessler *et al.* (1997). In that publication it was reported that five out 491 plants of the cross presenting the two marker alleles of M18 and AL07 in coupling with *Vf*, were classified as susceptible. The possibility that those plants were double recombinants was excluded because they were found at a too high frequency, 1 %, while the expected frequency was 9.6×10^{-4} . It was therefore assumed that the plants carried the *Vf* gene but were susceptible. It was proposed that the expression of the *Vf* gene was modified by the genetic background (modifiers) to such an extent that individuals carrying the resistance gene could become susceptible.

Compared to Gardiner *et al.* (1996) and Hemmat *et al.* (1998), who have used mapping populations of 258 and 73 plants respectively, our maps include 1179 and 2071 plants respectively. Due to the large number of individuals analysed, it was possible to determine the map position of *Vf* with a LOD score greater than 500 and to find out that AM19 maps closer to *Vf* than AL07.

The location of the *Vf* gene has been confirmed by three commercial scab resistant cultivars carrying *Vf* (Trent, Witos and Rewena). These three cultivars have undergone a recombination event between M18 (amplification of the allele in repulsion) and QU97 (amplification of the allele in coupling), indicating a different position of *Vf* than the one assigned by Gardiner *et al.* (1996) (Fig. 3B). The same line of reasoning can be applied to the position of *Vf* assigned by Hemmat *et al.* (1998). In fact, among the 2071 plants analysed, 27 recombinant plants indicate that *Vf* must be located between M18 and AM19, and not between AL07 and S5 as published in Hemmat *et al.* (1998).

For all these reasons I am confident that the position assigned to the *Vf* gene is correct and that for the cloning of the *Vf* gene it is correct to construct a contig spanning the chromosomal region between the M18 and AM19 molecular markers.

4.2 PHYSICAL MAPPING OF THE *Vf*-REGION

The haploid genome of *Malus x domestica* is estimated to be between 743-796 Mb (Arumuganathan and Earle 1991). The published linkage maps of apple range from 692 cM (Conner *et al.* 1997) to 984 cM (Maliepaard *et al.* 1998), by assuming an average haploid genome size of 770 Mbp, one cM would correspond for the largest linkage map to about 780 kb and for the smallest to 1110 kb.

The maximal physical distance of 870 kb between M18 and AL07 and their genetic distance of 1.3 cM calculated with the data of 1179 plants, or 1.5 cM with 2071 plants, lead to a specific ratio between genetic and physical distance for the *Vf*-region of 670 kb/cM and 580 kb/cM respectively. This value is not very different from the estimated value obtained with the largest linkage map of Maliepaard (780 kb/cM), indicating the absence of suppression of recombination in the region. In fact, where suppression of recombination was found, as in the Tm-2a region of the tomato, the

two values diverged by a factor of 7 (Ganal *et al.* 1989). The fact that recombination suppression was not observed could indicate that the introgressed region from the wild apple species is not substantially different from the one present in *Malus x domestica*.

The screening of the apple BAC library with the markers M18 permitted the identification of three BAC clones whose inserts were cleaved by *NotI*. The screening with AL07 gave six positive clones; two of them were also cleaved by *NotI* (one contained the allele of AL07 in coupling with *Vf* and the other the allele in repulsion, Vinatzer personal communication). The fact that these clones were the only ones found during the chromosome walk containing internal *NotI* restriction sites, indicate that the 870 kb *NotI* band (hybridising to the two markers) is not an artefact due to the presence of two *NotI* fragments of the same size, one carrying the M18 and the other carrying the AL07 marker. Considering that the physical distance between M18 and AM19 in the BAC contig is 550 kb and that the AM19-AL07 interval has a gap, the physical distance between the two *NotI* sites bracketing the *Vf*-region could be in fact 870 kb long, as estimated in the physical mapping experiment.

The aim of the physical mapping, performed before the chromosome walking, was to test the feasibility of the map-based cloning of *Vf* with the available material (apple BAC library and flanking markers M18 and AM19). The results that were available before the start of the chromosome walk were the estimation of the maximal physical distance between M18 and AL07 (870 kb), as well as the estimation of the specific ratio of the genetic/physical distance for this interval. Already this limited information permitted the estimation of the distance between M18 and *Vf* to be between 130 kb ($0.2 \text{ cM} \times 670 \text{ kb/cM}$) and 170 kb ($0.3 \text{ cM} \times 580 \text{ kb/cM}$), predicting the feasibility and good chance of success of a chromosome walking with the available material.

4.3 CHROMOSOME WALKING

The work presented here reports an improved methodology for chromosome walking with BAC clones. The IPCR technique proposed by Cai *et al.* (1995) for the isolation of T7-ends of BAC inserts has been successfully modified and applied also for Sp6-ends. This technique involves a three-step procedure consisting of BAC DNA digestion followed by ligation and PCR. The purified PCR products can be used directly for sequencing and as RFLP probes. These probes were used in Southern blot hybridisations of the *EcoRI* digested clones of the already established contig and of the new clones from which the probes had been derived. This allowed identification of the protruding ends of the contig. By using the probes on genomic Southern blots of recombinant plants and of 'Florina', the progress of the walk and the presence of repeated DNA in the probes could be assessed at the same time. In this way, it was possible to proceed immediately to the next library screening, i.e., the next step of chromosome walking. In the meantime, a new PCR-based marker was developed based on the sequence of the new end of the contig, allowing the identification of the real positives among the clones identified by hybridisation. At the same time, co-dominant markers allowed the discrimination among the positive BAC clones of the clones derived from the "resistant" chromosome from those derived from the "susceptible" one. The established procedure minimises the time between one chromosome walking step and the next.

The resulting contig containing *Vf*, encompassing the whole chromosomal region between M18 and AM19, measures about 550 kb. The M18-AM19 interval is completely covered by 13 resistant BAC clones. Nine "susceptible" clones were also found from the M18-AL07 interval, but they cover the region only partially (Fig. 6A). Clones connecting the M18-AM19 "resistant" contig with the BAC clones identified for AL07 were not found, thus indicating that the chromosome walking has been performed with a BAC library with a critical number of haploid genome equivalents. In fact, although the apple BAC library represents five haploid genome equivalents, Florina (the cultivar used for the construction of the BAC library) being heterozygous for the *Vf*-region, only 2.5 haploid genome equivalents were present in the library for each of the two homologue chromosomes of this region. Nevertheless, it was possible to construct a contig without gaps on the chromosomal region containing the *Vf* gene.

The screening of a large number of plants allowed the identification of 29 recombinant plants that permitted the reduction of the genomic region containing *Vf* from 550 kb to 350 kb. This region can be covered by five clones (M18-2, M18-5, M18-6, M18-7 and M18-8, Fig. 6). A clear suppression of recombination has been detected between the markers QU97 and M7T. In fact, in the region of about 140 kb between these two markers no crossing-over was found, while in the flanking region between the markers M7T and AM19 an average of eight recombination events per 100 kb in more than 2000 plants analysed was found. The suppression of recombination near the *Vf* gene could be due to a high degree of polymorphism present between the two homologous chromosomes in the resistant parents used in the crosses (all heterozygous for the *Vf*-region) that could prevent the correct pairing of the chromosomes during meiosis, thus inhibiting recombination. Adjacent to this region with suppressed recombination, the crossing-over frequency goes back to “normal” and most probably this is why the suppression of recombination was not detected during the physical mapping experiments. The suppression of recombination could partially explain why it was not possible to reduce the region where *Vf* is located to less than 350 kb.

With the construction of the contig encompassing the whole *Vf*-region, the feasibility of the map-based cloning approach in fruit trees has been verified. The construction of a contig spanning the *Vf*-region is a fundamental step towards to the cloning of the *Vf* gene, allowing the further characterisation of the *Vf*-region.

4.4 IDENTIFICATION OF RESISTANCE GENE HOMOLOGUES IN THE VF-REGION

4.4.1 Identification of *HcrVf1*, *HcrVf2* and *HcrVf3*

The most interesting genes found in the contig spanning the *Vf*-region are the *HcrVfs* (Homologues of *Cladosporium fulvum* resistance genes of the *Vf*-region) that present homologies to the *Cf* resistance genes of tomato. The first indication of the presence of this gene class in the *Vf*-region was found by performing a screening of a cDNA library of *V. inaequalis* inoculated leaves of Florina (the same cultivar used for the construction of the BAC library). Using the inserts of the 5 BAC clones necessary to span the *Vf*-region (350 kb) as a probe, 50 different cDNAs were identified, among the 750,000 cDNAs screened. Three of those 50 presented similarities to the *Cf* resistance genes. The comparison of the partial sequences of the BAC subclones hybridizing to them indicated that the three cDNAs were not from the *Vf*-region, or at least not from the chromosome in cis with the resistance.

The high number of different cDNAs hybridising to the *Vf*-region, 50 in 350 kb, would correspond to an average gene density of one gene every 7 kb if all these different genes would really be encoded in the *Vf*-region. The high number of different cDNAs hybridising to the five BACs can also be explained by the presence of some repeated sequences in the *Vf*-region. These sequences can be similar to LRR, found in the *HcrVfs*, or to microsatellites also found in some BAC clones of the *Vf*-region (Vinatzer personal communication). The low number (three) of cDNAs presenting homologies to the *HcrVfs* in a total number of 750,000 cDNAs screened, the fact that all three were different from each other, and the absence of matching sequences to *HcrVf1* and *HcrVf2* (*HcrVf3* is not expressed and *HcrVf4* is not yet well characterised), are indications that this gene family has a low level of expression. The low level of expression of this gene family has also been confirmed by northern blot experiments using 10 µg of total RNA, where no hybridisation could be found (data not shown). The RT-PCR product obtained with *HcrVf3* specific primers (with RNA from Florina and Prima) presented 97% identity with *HcrVf3* and 94% identity with *HcrVf1* and *HcrVf2* (N.B. 100% match was found only with *HcrVf4*). If the *HcrVf3* specific PCR product had used as a probe in the northern blot experiment, it would

have hybridised to *HcrVf1*, *HcrVf2* and *HcrVf4* mRNAs and to all other unknown expressed genes of the family. Given the large size (about 3 kb) and the small differences in length between the *HcrVfs* (e.g. about 100 bp between *HcrVf1* and *HcrVf2*) it would probably have been impossible to distinguish the different mRNAs. The expression of an *HcrVf* could only be proven with RT-PCR using *HcrVfs* specific primers, sequencing and comparing the RT-PCR sequence with the genomic sequence.

The strategy of aligning and orienting the three cDNAs (8-1, 10-5 and 15-4) and the partial BAC subclones sequences to the *Cf-9* sequence, in order to develop a pair of primers (fv) able to amplify large parts of *HcrVfs*, was decisive for the identification of *HcrVf1*, *HcrVf2* and *HcrVf3*. In fact, due to the similarity of the genes and, in most cases, insufficient overlap of the *EcoRI* with the *Sau3AI* subclones, it was almost impossible to be sure of the correct assembling of the subclones. Only the comparison of the subclone sequences with that of a long PCR fragment sufficiently ensures the correctness of the sequence assembly. Moreover if the sequence obtained derives from RT-PCR experiments (as was the case for *HcrVf1*), the expression of the gene was automatically demonstrated.

Due to the sequence conservation up-stream and down-stream from the *HcrVf* ORF, it was possible to develop the primers *HcrVf1for/rev*. These primers permitted the amplification of the three sequences containing the three complete ORFs of the *HcrVf1* to 3. Comparing the restriction digestion patterns of these PCR products and knowing the relative overlap of the BAC clones, it was possible to estimate the position of *HcrVf1*, *HcrVf2* and *HcrVf3* on the contig.

Because M18-2 is not completely in the *Vf*-region (two recombinant plants present a crossing-over between the two markers at its extremities), it was necessary to map *HcrVf1* and *HcrVf2*. *HcrVf2* was mapped in the *Vf*-region, while up to now it has not been possible to map *HcrVf1*. It was unnecessary to map *HcrVf3* because the BAC clone from which it was obtained lies in the middle of the *Vf*-region (*HcrVf3* is not expressed and thus is not a *Vf* candidate).

To verify if the N-terminus of the genes was reached, RACE experiments were performed. A RACE product was found with a perfect match to the first 244 bp in the coding sequence and up to 36 bases up-stream from the presumed first methionine of *HcrVf2* ORF. Twenty-six bases up-stream from the end of the match a putative TATA box and twenty-two bases further on, a CAAT box was found on the genomic

sequence of *HcrVf2*. The analysis of the presumed ORF revealed that the first 29 amino acids may encode a signal peptide. All this evidence and the similarities with the *Cf-9* resistance gene of tomato (presence of all the domains) suggest that the 26 bases at the beginning of the RACE product with no similarity to the genomic sequence of *HcrVf2* could be a PCR artefact, and that the gene is most probably complete. Indirectly, *HcrVf1* and *HcrVf3* should also be complete, because a high level of similarity as been observed between the members of this gene family.

The hybridisation of *HcrVf2* to the *EcoRI* digestions of the “resistant” and “susceptible” BAC clones of the *Vf*-region allowed the identification of at least 14 *HcrVf* loci. Seven *HcrVf* loci are in cis (two of those are out of the *Vf*-region but on the 550 kb contig) and seven are in trans with the resistance. Three of the loci in cis were analysed (the fourth is only partially characterised). *HcrVf1* and *HcrVf2* are expressed genes, while the third, *HcrVf3*, is not. Of the two expressed genes, only *HcrVf2* maps definitely in the *Vf* region. The remaining loci, all of them deriving from the BAC clone M18-6, located in the *Vf*-region, have to be analysed in future. Two of them should contain the two sequences M18-6A and M18-6B obtained with the *fv* primers, which are similar to the *HcrVfs* but contain internal stop codons.

In Figure 11 it is possible to see that the band called *HcrVf1b* of *HcrVf1* is present in BAC clone M18-1, M18-2 and in M18-3. The second band, *HcrVf1a*, containing the rest of the sequence of *HcrVf1*, is missing in M18-3. These results suggest that *HcrVf1b* must be the extremity of M18-3. During the construction of the contig all the BAC clone ends were sequenced. The comparison of the corresponding end of M18-3 did not show any homology to *HcrVf1*. This result can only be explained assuming that M18-3 is a chimeric clone. The chimeric part of M18-3 should therefore be the T7-end. This end was not used for the construction of the contig, therefore not compromising the correct assembling of the contig. If this end had been used for the screening of the library, the mapping of each new probe would have revealed the problem.

4.4.2 Homologies of the *HcrVfs* to the *Cf* resistance genes (and other *R*-genes)

The amino acid sequence comparison of the three *HcrVfs* permitted the identification of the most conserved and the most variable regions. While the domains A and D to G are very conserved (95 % and 96 % identity respectively), the domains B and C are quite variable, each showing only 84 % and 86 % identity respectively. If the domain C is divided into two sub-domains C1 and C2 (Tab. 5), it is possible to see that most of the diversity is concentrated in C2. In fact, C1 shows 89 % identity, but C2 shows 82 % identity. This kind of distribution of the variability in the LRR has already been found in the *Cf* resistance genes (Parniske *et al.* 1997). For the *Cf* resistance genes it has been proposed that the region corresponding to the N-terminal part of the LRRs, our C2 domain, encodes for the specificity domain that is responsible for the recognition of the fungus avirulence gene product, and for this reason it is the part of the gene that is less conserved (Parniske *et al.* 1997). Another interesting feature observed in the *HcrVfs* and also in the *Cf* genes is an interruption of the LRR motif after the fourth repeat.

Other similarities between the *HcrVfs* and the *Cf* genes are evident. The two gene families belong to the same class of genes, showing the same seven domains (Tab. 4) in the same order, with the identical LRR consensus sequence. Furthermore, *HcrVfs* and *Cf* genes have both been predicted to encode extracytoplasmatic glycoproteins anchored to the cell membrane, with the majority of the extracytoplasmatic domain composed of LRRs. If it were possible to demonstrate that one or more *HcrVfs* found in the *Vf*-region are resistance genes, then they would belong to the class IV of resistance genes.

Some common features can be found comparing *Cf* and *Vf* resistance. *Cf* and *Vf* resistant plants carrying the genes in the homozygous state show an increased resistance compared to the heterozygous plants (Hammon-Kosak and Jones 1994, Gessler *et al.* 1997), thus indicating an incomplete dominance of these genes. It is possible that two copies of the gene result in higher amounts of active protein and this accounts for the discernible effect. An additional common feature is that the resistance level depends on the background (Hammon-Kosak and Jones 1994, Lamb and Hamilton 1970). The effect of the background should be considered in mapping experiments. This is especially true for genes such as *Vf*, which shows a high variation in the expression of the resistance. In these cases, parents known to show more

extreme resistance phenotypes should be preferred to simplify the scoring of the resistance phenotype.

A last characteristic of the *HcrVfs* that is common to many resistance genes is the clustering of many homologue genes in the same region. For example, in the *Vf*-region (in cis with the resistance) a number varying between 7 and 16 *HcrVf* homologues can be presumed.

Because *HcrVfs* are similar to the *Cf* resistance genes, and assuming that one of them is *Vf*, it can be speculated that a *Cf*-like mechanism of resistance is present in the apple. Joosten and de Wit (1999) presented three different possible mechanisms leading to resistance in tomato plants carrying *Cf-9*. In all three mechanisms the *Cf-9* protein, which is mostly extra-cytoplasmatic, interacts with a second membrane-anchored protein (transducer) with a cytoplasmatic domain carrying a kinase or that interacts with a kinase protein. Upon interaction of the *Avr9* protein with the *Cf-9* protein, the complex interacts with the transducer and defence response initiates. Alternatively, the *Avr9* protein interacts with another membrane-anchored protein (which could be the transducer itself or not) that afterwards interacts with the *Cf-9* protein, after which the defence response is initiated. It has been demonstrated that *Cf* proteins are not the primary receptors of the elicitors (*Avr9* protein) (Kooman-Gersmann *et al.* 1996), but it is possible that they interact with a lower affinity than the target. After recognition, the signal transduction involves phosphorylation cascades that can eventually activate effector proteins such as enzymes or transcription factors.

In the apple, the *Vf* mediated resistance may function with a similar mechanism. A protein called LRPKm1 has recently been cloned from Florina (Komjanc *et al.* 1999). LRPKm1 is predicted to encode a LRR-receptor like protein kinase that is present in multi copies in all apple cultivars (resistant and susceptible). Its expression is induced upon pathogen attack and salicylic acid treatment. The resistance mechanism could be described as follows: upon perception of the *AvrVf* protein by the *Vf* protein, the *Vf-AvrVf* complex interacts with LRPKm1 and the defence response is initiated by activating the kinase domain of LRPKm1. Alternatively, the LRPKm1 protein could interact directly with *AvrVf* to form a LRPKm1-*AvrVf* complex, which through the interaction with one of the *HcrVf* proteins trigger the kinase activity of LRPKm1, eventually starting the plant resistance response.

An increase of LRPK_{m1} expression after salicylic acid treatment (salicylic acid is known to improve expression of defence response proteins) could reinforce the signal transduction pathway. Being present in resistant and susceptible cultivars, but giving rise to an incompatible reaction only in the *Vf* cultivar Florina, LRPK_{m1} gene could be one of the proteins modifying *Vf* resistance. Different LRPK_{m1} allelic combinations (or of similar genes) could modulate the different resistance phenotypes observed. However, this mechanism is highly speculative.

5. OUTLOOK

The results presented in this thesis should form the foundation for further studies aiming at the cloning and understanding the mechanism of resistance provided by the *Vf* gene. First of all the remaining *HcrVfs* should be characterised. A good candidate, which is not yet completely characterised, is *HcrVf4*. *HcrVf4* is expressed and it is located in the *Vf*-region. However, *HcrVf4* is not the only *HcrVf* that has to be characterised. From the hybridisation of *HcrVf2* on the *EcoRI* digested BAC clones of the contig, it is known that other *HcrVfs* are present on the BAC clones M18-5 and M18-6, of which nothing is known except for the two PCR products obtained with the primer fv, containing truncated ORFs.

Furthermore, other recombinant plants for the *Vf*-region should be found to attempt to further reduce the region and therefore to be able to exclude some *HcrVfs*. However, as already demonstrated by the genetic mapping, recombination events in the *Vf*-region between the BAC clones M18-3 to M18-7 are extremely rare. In fact, in the 2071 plants analysed no recombinants could be found.

At the moment *HcrVf1* is still not mapped. The comparison of its sequence with the sequences of all other *HcrVfs* should permit the development of *HcrVf1* specific primers that will allow its mapping.

Whenever a complete sequence of an expressed *HcrVf* gene, which maps in the *Vf*-region is available, complementation experiments of a susceptible cultivar should be performed. The *HcrVf* that will transform a susceptible plant into a resistant one will be the *Vf* resistance gene. This kind of experiment will be the only one that will definitively demonstrate which *HcrVf* is the *Vf* gene. However, besides the homology between the *HcrVfs* and the *Cf* resistance genes, up to now there is no other evidence that *HcrVfs* could be involved in *Vf*-mediated resistance.

A possibility to have an indirect indication in a relatively short time that the *HcrVfs* are really involved in resistance, is given once more by the fv primers. Preliminary experiments have shown that PCR products of different sizes were amplified with these primers from 21 different apple cultivars, regardless of the fact that these plants were susceptible or resistant. A RT-PCR band of Golden Delicious (band L, Fig. 9) obtained with the fv primers has been cloned and sequenced and the same domains of

the *HcrVf* were found, thus indicating that *HcrVfs* are present and expressed also in a “susceptible” cultivar. This result can be explained in two ways. The first explanation is that the *fv* primers can amplify the “ephemeral” resistance genes present in apple cultivars (Sierotzki *et al.* 1994, Koch *et al.* 2000) as well as other resistance genes. If this is true, it will be interesting to use these primers in mapping experiments of several resistance genes to verify if a marker derived from these primers co-segregates with the resistance. If this is the case, an indirect demonstration that one of the *HcrVfs* is the *Vf* gene will be obtained. However, it will not be clear which *HcrVf* is the *Vf* gene. Although less likely, a second possible explanation could be that our candidate genes are present in all apple plants and they are not involved in the resistance.

To reduce the intrinsic risk of concentrating all the efforts on a single gene, the transformation of susceptible apple cultivars with the inserts of the five BAC clones spanning the *Vf*-region using the BIBAC technology (Hamilton 1997) should be continued (this part of the work has been initiated by the group of Professor Sansavini at the DCA Bologna). The transfer of the inserts of the BAC clones M18-5 and M18-6 and transformation experiments have already been performed and the results should be available some time in the near future.

6. CONCLUSIONS

The results obtained so far in the map-based cloning of the *Vf* resistance gene from *Malus floribunda* 821 are very promising. In fact, all the intermediate steps have been successfully completed and a cluster of putative resistance genes (*HcrVf*) has been identified in the *Vf*-region. In the first part of the work strong evidence of the accuracy of the position of the *Vf* gene on the genetic map was presented. In the second part, the feasibility of a chromosome walk employing the available apple BAC library and the two *Vf* flanking markers M18 and AL07 was demonstrated. Furthermore, the best starting points for the chromosome walk, M18 and AM19, were determined. In the third part, the chromosome walk was successfully completed, allowing the construction of a contig encompassing the whole *Vf*-region between M18 and AM19. The construction of the contig is a fundamental step towards the cloning of the *Vf* gene. Finally, the analysis of the *Vf*-region allowed the identification of a cluster of genes, *HcrVf*, homologues to the *Cf* resistance genes of tomato.

Preliminary results demonstrate that *HcrVf* homologues are present in all analysed cultivars (susceptible and resistant) and in other breeding selections carrying different resistance genes. If it is demonstrated that one of the *HcrVfs* is *Vf*, it could also be hypothesised that the *HcrVf* homologues, present in all apple cultivars, are ephemeral resistance genes overcome by different races of the fungus. The identification of ephemeral resistance genes would be a very important result since these genes could be used for a better deployment of the resistance. New strategies, such as resistance genes pyramiding and/or selection of “susceptible” cultivars with the best combination of “ephemeral” resistance genes, could be proposed and eventually used in mixed cultivar orchards to control apple scab in a more efficient and ecologically conscious way, reducing the 15 annual fungicide applications currently used.

7. REFERENCES

- Aderhold R (1902) Ein Beitrag zur Frage der Empfänglichkeit der Apfelsorten für *Fusicladium dendriticum* (Wallr.) Fuckel und deren Beziehungen zum Wetter. Arb. K. Biol. Anst. Landw. Forstw. 2:560-566
- Agrios, GN (1997) Plant Pathology (4th edition). Academic Press, San Diego (USA)
- Alberts B, Bray D, Lewis J, Raff M, Roberts K Watson JD (1990) Molekularbiologie der Zelle. VHC Verlagsgesellschaft mbH: 535
- Aldrich J and Cullis CA (1993) RAPD analysis in Flax: Optimization of yield and reproducibility using KlenTaq 1 DNA Polymerase, Chelex 100 and gel purification of genomic DNA. Plant Mol Biol Rep 11 (2):128-141
- Alston FH (1989) Breeding pome fruits with stable resistance to diseases. In Gessler, Butt and Koller (eds): Integrated control of pome fruit diseases II, OILB-WPRS Bulletin XII/6:90-99
- Arumuganathan K and Earle ED (1991) Nuclear DNA content of some important species. Plant Mol Biol Rep 9 (3): 208-218
- Bénaouf G and Parisi L (1997) Pathogenicity of *Venturia inaequalis* strains from *Malus floribunda* 821. In Berrie, Xu, Harris, Roberts, Evans, Barbara and Gessler, (eds): Integrated control of pome fruit diseases IV, OILB-WPRS Bulletin 20 (9): 8-11
- Bénaouf G and Parisi L (2000) Genetics of host-pathogen relationships between *Venturia inaequalis* races 6 and race 7 and malus species. Phytopathology 90 (3): 236-242
- Bent AF (1996) Plant Disease Resistance Genes: Function Meets Structure. The Plant Cell 8: 1757-1771
- Bent AF, Kunkel BN, Dahlbeck D, Brown KL, Schmidt R, Giraudat J, Leung J, Staskawicz BJ (1994) *RPS2* of *Arabidopsis thaliana*: A Leucine-Rich Repeat Class of Plant Disease Resistance Genes. Science 265: 1856-1860
- Boone DM (1971) Genetics of *Venturia inaequalis*. Ann Rev Phytopathology 9: 297-318
- Burke DT, Carle GF, Olson MV (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science 236:806-812
- Burr B, Evola SV, Burr FA, Backmann JS (1983). The application of restriction fragment length polymorphisms to plant breeding. In: Setlow J.K and Hollander A.(eds). Genetic Engineering Principles and Methods vol 5:45-49
- Büschges R, Hollricher K, Panstruga R, Simons G, Wolter M, Frijters A, van Daelen R, van der Lee T, Diergaarde P, Groenendijk J, Töpsch S, Vos P, Salamini F and Schultze-Lefert P (1997) The barley *Mlo* gene: a novel control element of plant pathogen resistance. Cell 88: 695-705
- Cai D, Kleine M, Kifle S, Harloff HJ, Sandal NN, Marcker KA, Klein-Lankhorst RM, Salentijn EMJ, Lange W, Stiekema WJ, Wyss U, Grundler FMW, Jung C (1997) Positional Cloning of a Gene for Nematode Resistance in Sugar Beet. Science 275: 832-834
- Cai L, Taylor JF, Wing RA, Gallagher DS, Woo SS, Davis SK (1995) Construction and characterization of a bovine bacterial artificial chromosome library. Genomics 29(2): 413-425

- Cheng FS, Weeden NF, Brown SK, Aldwinckle HS, Gardiner SE, Bus VG (1998) Development of a DNA marker for *Vm*, a gene conferring resistance to apple scab. *Genome* 41 (2): 208-214
- Chevalier M, Lespinasse Y and Renaudin S (1991) A microscopic study of different classes of symptoms coded by the *Vf* gene in apple for resistance to scab (*Venturia inaequalis*). *Plant Pathol.* 40: 249-256
- Chu G, Vollrath D, Davis RW (1986) Separation of large DNA molecules by contour-clamped homogeneous electric fields. *Science* 234: 1582-1585
- Conner PJ, Brown SK, Weeden NF (1997) Randomly amplified polymorphic DNA-based genetic linkage maps of three apple cultivars. *J Am Soc Hortic Sci* 122: 350-359
- Crandall CS (1926) Apple breeding at the University of Illinois. *Illinois Agr. Expt Sta. Bull.* 275: 341-600 0335 6231233
- Crosby JA, Janick J., Pecknold PC, Korban SS, O'Connor PA, Ries SM, Gofreda J, Voordeckers A (1992) Breeding Apples for Scab Resistance: 1945-1990. *Fruit Varieties Journal* 46(3): 145-166
- Crute IR and Pink DAC (1996) Genetics and utilization of pathogen resistance in plants. *Plant Cell* 8:1747-1755
- Dayton DF, Williams EB (1968) Independent genes in *Malus* for resistance to *Venturia inaequalis*. *Proc Am Soc Hort Sci* 93:89-94
- De Wit JGM (1997) Pathogen avirulence and plant resistance: a key role for recognition. *Trend in plant science* 2 (12): 452-458
- Delaney T, Uknes S, Vernooij B, Friedrich L, Weymann K, Negrotto D, Gaffney T, Gut-Rella M, Kessmann H, Ward E, Ryals J (1994) A central role of salicylic acid in plant disease resistance. *Science* 266: 1247-1250
- Dixon MS, Hatzixanthis K, Jones DA, Harrison K, Jones JDG (1998) The tomato *Cf-5* disease resistance gene and six homologs show pronounced allelic variation in leucine rich repeat copy number. *Plant Cell* 10: 1915-1925
- Dixon MS, Jones DA, Keddie JS, Thomas CM, Harrison K, Jones JDG (1996) The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. *Cell* 84: 451-459
- Flor HH (1971) Current status of the gene-for-gene concept. *Annu Rev Phytopathol* 9: 275-296
- Ganal MW, Young ND and Tanksley SD (1989) Pulsed field gel electrophoresis and physical mapping of large DNA fragments in the *Tm-2a* region of chromosome 9 in tomato. *Mol Gen Genet* 215: 395-400
- Gardiner SE, Basset HCM, Noiton DAM, Bus VG, Hofstee ME, White AG, Ball RD, Forster RLS, Rikkerink EHA (1996) A detailed linkage map around an apple scab resistance gene demonstrates that two disease classes both carry the *Vf* gene. *Theor Appl Genet* 93: 485-493
- Gessler C (1989) Genetics of the interaction *Venturia inaequalis-Malus*: the conflict between theory and reality. In Gessler, Butt and Koller (eds): *Integrated control of pome fruit diseases II*, OILB-WPRS Bulletin XII/6: 168-190
- Gessler C, Gianfranceschi L, Koller B, Seglias N, Sierotzki H, Koch T, Tenzer I, Blaise P (1995) Strategies of gene deployment in breeding apples with resistance to scab and mildew. *Agrobiotec atti del convegno Ferrara 1995*: 251-266
- Gessler C, Patocchi A, Kellerhals M and Gianfranceschi L (1997) Molecular markers applied to apple breeding and map-based cloning of resistance genes. In Berrie, Xu, Harris, Roberts, Evans, Barbara and Gessler, (eds): *Integrated control of pome fruit diseases IV*, OILB-WPRS Bulletin 20 (9): 105-109

- Gianfranceschi L, Koller B, Seglias N, Kellerhals M, Gessler C (1996) Molecular selection in apple for resistance to scab caused by *Venturia inaequalis*. *Theor Appl Genet* 93: 199-204
- Grant MR, Godiard L, Straube E, Ashfield T, Lewald J, Sattler A., Innes RW, Dangl JL (1995) Structure of the *Arabidopsis* RPM1 gene enabling dual specificity disease resistance. *Science* 269: 843-846
- Hamilton CM (1997) A binary-BAC system for plant transformation with high-molecular-weight DNA. *Gene* 200:107-116
- Hammond-Kosack KE and Jones JDG (1996) Resistance gene-dependent plant defense responses. *The plant cell* (8): 1773-1791
- Hemmat M, Weeden NF, Aldwinckle HS and Brown SK (1998) Molecular markers for the scab resistance (*Vf*) region in apple. *J Am Soc Hortic Sci* 123 (6):992-996
- Hough LF (1944) A survey of the scab resistance of the foliage on seedling in selected apple progenies. *Amer Soc Hortic Sci* 44: 260-272
- Hough LF, Shay JR and Dayton DF (1953) Apple scab resistance from *Malus floribunda* Sieb. *Am Soc Hortic Sci* 63: 341-347
- Iaonnou PI, Amemiya CT, Garnes J, Kroisel PM, Shizuya H, Chen C, Batzer MA and de Jong PJ (1994) A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nature Genetics* 6: 84-89
- Janick J, Cummings JN, Brown SK and Hemmat M (1996) Apples. In Janick J and Moore JN (eds), *Fruit breeding (vol. I): tree and tropical fruits*. John Wiley & Son, New York, NY, pp 1-77
- Johal GS, Briggs SP (1992) Reductase Activity Encoded by the HM1 Disease Resistance Gene in Maize. *Science* 258: 985-987
- Jones DA, Thomas CM, Hammond-Kosack KE, Balint-Kurti PJ, Jones JDG (1994) Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. *Science* 266: 789-793
- Joosten MHAJ and de Wit PJGM (1999) The tomato-*Cladosporium fulvum* interaction: a versatile experimental system to study plant-pathogen interactions. *Annu Rev Phytopathol* 37: 335-367
- King GJ, Alston FH, Brown LM, Chevreau E, Evans KM, Dunemann F, Janse J, Laurens F, Lynn JR, Maliepaard C, Manganaris AG, Roche P, Schmidt H, Tartarini S, Verhaegh J, Vrieling R (1998) Multiple field and glasshouse assessments increase the reliability of linkage mapping of the *Vf* source of scab resistance in apple. *Theor Appl Genet* 96: 699-708
- Kobe B, Deisenhofer J (1994) The leucine-rich repeat: a versatile binding motif. *Trends Biochem Sci* 19: 415-421
- Koch T (1998) Differential resistance and virulence in the interaction of *Malus x domestica* and *Venturia inaequalis*. ETHZ Diss.ETH Nr. 12547
- Koch T, Kellerhals M and C Gessler (2000) Virulence pattern of *Venturia inaequalis* field isolates and corresponding differential resistance in *Malus x domestica*. *J Phytopathology* 148: in press
- Koller B, Gianfranceschi L, Seglias N, McDermott J, Gessler C (1994) DNA-markers linked to the *Malus floribunda* 821 scab resistance. *Plant Mol Biol* 26: 597-602
- Komjanc M, Festi S, Rizzotti L, Cattivielli L, Cervone F, De Lorenzo (1999) Leucine Rich Repeat Receptor Like Protein Kinase (LRPKm1) gene is induced in *Malus x domestica* by *Venturia inaequalis* and Salicylic Acid. *Plant Mol Biol* 40: 945-957
- Konieczny A, Ausubel FM (1993) A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. *Plant J* 4: 403-410

- Kooman-Gersmann M, Vogelsang R, Vossen P, Van den Hooven H, Mahé E (1998) Correlation between binding affinity and necrosis-inducing activity of mutant AVR9 peptide elicitors. *Plant physiol* 117: 609-618
- Lamb RC and Hamilton JM (1970) Environmental and genetic factors influencing the expression of resistance to scab (*Venturia inaequalis* Cke. Wint.) in apple progenies. *J. Am. Soc. Hort. Sci.* 94: 554-557
- Lawrence GJ, Finnegan EJ, Ayliffe MA, Ellis JG (1995) The L6 Gene for Flax Rust Resistance Is Related to the Arabidopsis Bacterial Resistance Gene RPS2 and the Tobacco Viral Resistance Gene N. *Plant Cell* 7(8): 1195-1206
- Lisitsyn N, Lisitsyn N, Wigler M (1993) Cloning the differences between two complex genomes. *Science* vol. 259:946-951.
- MacHardy WE (1996) Inheritance of resistance to *Venturia inaequalis*. In: Apple scab, biology, Epidemiology and management. APS Press, St. Paul, Minnesota, pp 61-103
- Maliepaard C, Alston FH, van Arkel G, Brown LM, Chevreau E, Dunemann F, Evans KM, Gardiner S, Guilford P, van Heusden AW, Janse J, Laurens F, Lynn JR, Manganaris AG, den Nijs APM, Periam N, Rikkerink E, Roche P, Ryder C, Sansavini S, Schmidt H, Tartarini S, Verhaegh JJ, Vrieling-van Ginkel M, King GJ (1998) Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. *Theor Appl Genet* 97: 60-73
- Markussen T, Kruger J, Schmidt H, Dunemann F (1995) Identification of PCR-based markers linked to the powdery-mildew-resistance gene *Pl₁* from *Malus robusta* in cultivated apple. *Plant Breed* 114: 530-534
- Martin GB, Brommonschenkel SH, Chunwongse J, Frary A, Ganai MW, Spivey R, Wu T, Earle ED and Tanksley SD (1993) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* 262: 1432-1436
- Martin GB, Frary A, Wu T, Brommonschenkel SH, Chunwongse J, Earle ED and Tanksley SD (1994) A member of the tomato *Pto* gene family confers sensitivity to Fentiom resulting in rapid cell death. *Plant Cell* (6): 1543-1552
- Mehdy MC (1994) Active oxygen species in plant defense against pathogens. *Plant Physiol* 105: 467-472
- Meyerowitz E, Guild G, Prestidge L and Hogness D (1980) A new high-capacity cosmid vector and its use. *Gene* 11:271.
- 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-Rich Repeat Family of Plant Genes. *The Plant Cell* 10: 1307-1319
- Moragnte M and Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. *Plant J.* 3: 175-182
- Nakamura S, Asakawa S, Ohmido N, Fukui K, Shimizu N, Kawasaki S (1997) Construction of an 800-kb contig in the near-centromeric region of the rice blast resistance gene Pi-ta-2 using a highly representative rice BAC library. *Mol Gen Genet* 254(6): 611-620
- Parisi L, Lespinasse Y, Guillaumes J, Kruger J (1993) A new race of *Venturia inaequalis* virulent to apples with resistance due to the *Vf* gene. *Phytopathology* 83(5): 533-537
- Parker JE, Coleman MJ, Szabo V, Frost LN, Schmidt R, Van der Biezen R, Moores T, Dean C, Daniels MJ, Jones JDG (1997) The *Arabidopsis* downy mildew resistance gene *RPP5* shares similarity to the Toll and interleukin-1 receptors with *N* and *L6*. *Plant Cell* 9: 879-894

- Parniske M, Hammond-Kosack KE, Golstein C, Thomas CM, Jones DA, Harrison K, Wulff BBH, Jones JDG (1997) Novel Disease Resistance Specificities Result from Sequence Exchange between Tandemly Repeated Genes at the *Cf-4/9* Locus of Tomato. *Cell* 91: 821-832
- Parniske M, Wulff BBH, Bonnema G, Thomas CM Jones DA and Jones DGJ (1999) Homologues of the *Cf-9* disease resistance gene (*Hcr9s*) are present at multiple loci on the short arm of tomato chromosome 1. *Molecular Plant-Microbe Interaction* 12 (2): 93-102
- Paterson AH and RA Wing (1993) Genome mapping in plants. *Curr Opin in Biotech* 4:142-147
- Paterson AH, Brubaker CL and Wendel JF (1993) A rapid method for extraction of cotton (*Gossypium spp.*) genomic DNA suitable for RFLP or PCR analysis. *Plant Mol Biol Rep* 11(2): 122-127
- Pierce JC, Sternberg N and Sauer B (1992) A mouse genomic library in the bacteriophage P1 cloning system: organization and characterization. *Mammalian Genome* 3: 550-558
- Pryor T (1987) The origin and structure of fungal disease resistance genes in plants. *Trends in genetics* 3 (6): 157-161
- Roche P, Alston FH, Maliepaard C, Evans KM, Vrieling R, et al (1997) RFLP and RAPD markers linked to the rosy leaf curling aphid resistance gene (*Sd₁*) in apple. *Theor Appl Genet* 94: 528-533
- Rudloff CF, Schmidt M (1934) *Venturia inaequalis* (Cooke) Aderh. II. Zur Zuechtung schorf widerstandsfaehiger Apfelsorten. *Zuechter* 6: 288-294
- Rudloff CF, Schmidt M (1935) Der Erreger des Apfelschorfes *Venturia inaequalis* (Cooke) Aderh. Grundlagen und Moeglichkeiten fuer seine Bekaempfung auf zuechterischem Wege II. *Zuechter* 7: 65-74
- Salmeron JM, Oldroyd GED, Rommes CMT, Scofield SR, Kim H-S, Lavelle DT, Dahlbeck D, Staskawicz BJ (1996) Tomato *Prf* Is a Member of the Leucine-Rich Repeat Class of Plant Disease Resistance Genes and Lies Embedded within the *Pto* Kinase Gene Cluster. *Cell* 86: 123-133
- Seglias N (1997b) Genetische Kartierung quantitativer Merkmale beim Apfel. ETHZ Diss.ETH Nr. 12204
- Seglias N. and Gessler C (1997a). Genetics of apple powdery mildew resistance from *Malus zumi* (*PI2*). In Berrie, Xu, Harris, Roberts, Evans, Barbara and Gessler, (eds): Integrated control of pome fruit diseases IV, OILB-WPRS Bulletin 20 (9):195-208
- Shizuya H, Birren B, Kim U, Mancino V, Slepak T, Tachiiri Y, Simon M (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc Natl Acad Sci USA* 89: 8794-8797
- Sierotzki H, Eggenschwiler M, Boillat O, McDermott JM, Gessler C (1994) Detection of variation in virulence toward susceptible apple cultivars in natural populations of *Venturia inaequalis*. *Phytopathology* 84: 1005-1009
- Simons G, Groenendijk J, Wijbrandi J, Reijans M, Groenen J, Diergaarde P, Van der Lee T, Bleeker M, Onstenk J, de Both M, Haring M, Mes J, Cornelissen B, Zabeau M, and Vos P (1998) Dissection of the *Fusarium I2* Gene Cluster in Tomato Reveals Six Homologs and One Active Gene Copy. *Plant Cell* 10: 1055-1068.

- Song WY, Wang GL, Chen LL, Kim HS, Pi LY, Holsten T, Gardner J, Wang B, Zhai WX, Zhu LH, Fauquet C, Ronald P (1995) A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. *Science* 270: 1804-1806
- Stam P (1993) Construction of integrated genetic linkage maps by means of a computer package Joinmap. *Plant J* 3: 739-744
- Tanksley SD, Martin WG, Martin GB (1995) Chromosome landing: a paradigm for map-based gene cloning in plants with large genomes. *TIG* 11(2): 63-68
- Tartarini S (1996) RAPD markers linked to the *Vf* gene for scab resistance in apple. *Theor Appl Genet* 92 (7): 803-810
- Tartarini S, Gianfranceschi L, Sansavini S and Gessler C (1999) Development of reliable PCR markers for the selection of the *Vf* gene conferring scab resistance in apple. *Plant Breed* 118: 183-186
- Thomas CM, Dixon MS, Parniske M, Golstein C and Jones JDG (1998) Genetic and molecular analysis of tomato *Cf* genes for resistance to *Cladosporium fulvum*. *Phil. Trans. R.Soc. Lond. B* 353: 1413-1424
- Thomas CM, Jones DA, Parniske M, Harrison K, Balint Kurti PJ, Hatzixanthis K, Jones JDG (1997) Characterization of the tomato *Cf-4* gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in *Cf-4* and *Cf-9*. *The Plant Cell* 9 (12): 2209-2224
- Vinatzter BA, Zhang H-B and Sansavini S (1998) Construction and characterization of a bacterial artificial chromosome library of apple. *Theor Appl Genet* 97: 1183-1190
- von Heijne G (1986) A new methods for predicting signal sequence cleavage sites. *Nucl. Acids Res.* 14 (11): 4683-4691
- Wallace E (1913) Scab disease of apples. N.Y. (Cornell) *Abr. Expt. Sta. Bull.* 335:545-592
- Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsei L, Topaloglou T, Hubbell E, Robinson E, Mittmann M, Morris MS, Shen NP, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson TJ, Lipshutz R, Chee M, Lander ES (1998) Large-scale identification, mapping and genotyping of single-nucleotide polymorphisms in the human genome. *Science* 280 (5366): 1077-1082
- Wang ZX, Masahiro Y, Yamanouchi U, Iwamoto M, Monna L, Hayasaka H, Katayose Y and Sasaki T (1999) The *Pib* gene for rice blast resistance belongs to the nucleotide binding and leucine-rich repeat class of plant disease resistance genes. *Plant J* 19 (1): 55-64
- Whitham S, Dinesh-Kumar SP, Choi D, Hehl R, Corr C, Baker B (1994). The product of the tobacco mosaic virus resistance gene N: Similarity to toll and the interleukin-1 receptor. *Cell* 78(6): 1101-1115
- Williams EB and Kuc J (1969) Resistance in *Malus* to *Venturia inaequalis*. *Annu Rev of Phytopathol* 7: 223-246
- Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphism amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* 18: 6531-6535
- Wing RA, Zhang H-B, Tanksley S D (1994) Map-based cloning in crop plants. Tomato as a model system: I. Genetic and physical mapping of jointless. *Mol Gen Genet* 242: 681-688
- Woo SS, Jiang J, Gill BS, Paterson AH, Wing RA (1994) Construction and characterization of a bacterial artificial chromosome library of *Sorghum bicolor*. *Nucl Acids Res* 22: 4922-4931

- Yang D, Sanchez A, Khush GS, Zhu Y, Huang N (1998) Construction of a BAC contig containing the *xa5* locus in rice. *Theor Appl Genet* 97: 1120-1124
- Yang H, Krueger J (1994) Identification of an RAPD marker linked to the *Vf* gene for scab resistance in apples. *Plant Breeding* 112(4): 323-329
- Young ND, Tanksley SD (1989) RFLP analysis of the size of chromosomal segments retained around the *Tm-2* locus of tomato during backcross breeding. *Theor Appl Genet* 77: 353-359
- Zabeau M (1992) European Patent Application, publication no. EP 0534858.
- Zhang H-B, Choi S, Woo S-S, Li Z, Wing RA (1996) Construction and characterization of two rice bacterial artificial chromosome libraries from the parents of a permanent recombinant inbred mapping population. *Molecular Breeding* 2(1): 11-24
- Zhang H-B, Zhao X, Ding X, Paterson AH, Wing RA (1995) Preparation of megabase-size DNA from plant nuclei. *Plant J* 7: 175-184
- Zhou J, Loh Y-T, Bressan, RA, Martin GB (1995) The tomato gene *Pti1* encodes a serine/threonine kinase that is phosphorylated by Pto and is involved in the hypersensitive response. *Cell* 83: 925-935

CURRICULUM VITAE

ANDREA PATOCCHI

Born March 19th, 1971

Citizen of Switzerland

- 1977-1982 Scuola elementare, Bellinzona
- 1982-1986 Scuola Media, Bellinzona
- 1986-1990 Liceo, Bellinzona, Matura (Typus C)
- 1990-1991 Studies of Agronomy at the Swiss Federal Institute of Technology (ETHZ), Zürich
- 1991-1996 Studies of Natural Sciences at the Swiss Federal Institute of Technology (ETHZ), Zürich (Dipl. Natw. ETH, Xac)
- 1997-2000 Dissertation at the Institute of Plant Sciences, Plant Pathology Group, Swiss Federal Institute of Technology (ETHZ), Zürich

PUBLICATIONS

- Gessler C, Patocchi A, Kellerhals M and Gianfranceschi L (1997) Molecular markers applied to apple breeding and map-based cloning of resistance genes. In Berrie, Xu, Harris, Roberts, Evans, Barbara and Gessler, (eds): Integrated control of pome fruit diseases IV, OILB-WPRS Bulletin 20 (9):105-109
- Patocchi A, Gianfranceschi L and Gessler C (1999) Towards the map-based cloning of *Vf*: Fine and physical mapping of the *Vf* region. Theor Appl Genet 99: 1012-1017
- Patocchi A, Vinatzer BA, Gianfranceschi L, Taratarini S, Zhang HB, Sansavini S and 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
- Patocchi A, Vinatzer BA, Gianfranceschi L, Taratarini S, Sansavini S and Gessler C (1999) Construction of 550 kb BAC contig in apple spanning the genomic region containing the *Vf* scab resistance gene. Poster abstract from Plant and Animal Genome VII meeting. San Diego CA

RINGRAZIAMENTI

Il mio grazie di cuore va a tutte quelle persone che hanno permesso la realizzazione di questa tesi. Nella speranza di non dimenticare qualcuno e senza voler far classifiche un grazie particolare va a:

- CESARE GESSLER per avermi dato la possibilità di questa tesi e per la fiducia sempre accordatami
- LUCA GIANFRANCESCHI per tutto quanto mi ha insegnato e per le correzioni a tempo di record di questa tesi
- BORIS VINATZER per aver letto i miei 640 e-mails e avermene scritto 558 in tre anni di collaborazione!
- PROF. KLAUS APEL für die Übernahme des Doktorvaters
- PROF. SILVIERO SANSAVINI e tutto il resto del gruppo di Bologna per la sempre squisita accoglienza e l'ottima collaborazione
- ISABEL TENZER come sempre per tutto...
- BERNIE KOLLER e PHILIPPE BLAISE für die Unterstützung in Computerfragen
- Tutto il resto del gruppo di fitopatologia per la serena atmosfera di lavoro
- MARKUS KELLERHALS, MONIKA GÖRRE und KAREN WOLEWINSKI für Pflanzenmaterial
- KAREN RUDIN, CELESTE LINDE and EVE DILWORTH for the stylistic english corrections of this thesis
- MARCELLO ZALA per l'efficienza nelle ordinazioni
- ANGELA TENZER e GIOVANNI BROGGINI per avermi aiutato nei lavori di routine
- Tutti gli amici dello "Stall" per la compagnia nei momenti di relax
- E ovviamente un grazie particolare va a tutta la mia famiglia alla quale dedico questa tesi

Si ringrazia il Politecnico Federale di Zurigo, fondo 0-20-345-96 per avere finanziato questa ricerca.