Tree Genetics & Genomes (2013) 9:237–251
DOI 10.1007/s11295-012-0550-3

ORIGINAL PAPER

A candidate gene for fire blight resistance in *Malus* × *robusta* 5 is coding for a CC–NBS–LRR

Johannes Fahrentrapp · Giovanni A. L. Broggini · Markus Kellerhals · Andreas Peil · Klaus Richter · Elena Zini · Cesare Gessler

Received: 16 March 2012 / Revised: 11 May 2012 / Accepted: 11 July 2012 / Published online: 31 July 2012

Abstract Fire blight is the most important bacterial disease in apple (*Malus* × *domestica*) and pear (*Pyrus communis*) production. Today, the causal bacterium *Erwinia amylovora* is present in many apple- and pear-growing areas. We investigated the natural resistance of the wild apple *Malus* × *robusta* 5 against *E. amylovora*, previously mapped to linkage group 3. With a fine-mapping approach on a population of 2,133 individuals followed by phenotyping of the recombinants from the region of interest, we developed flanking markers useful for marker-assisted selection. Open reading frames were predicted on the sequence of a BAC spanning the resistance locus. One open reading frame coded for a protein belonging to the NBS–LRR family. The in silico investigation of the structure of the candidate resistance gene against fire blight of *M.* × *robusta* 5, *FB_MR5*, led us hypothesize the presence of a coiled-coil region followed by an NBS and an LRR-like structure with the consensus 'LxxLx[IL]xxCxxLxxL'. The function of FB_MR5 was predicted in agreement with the decoy/guard model, that FB_MR5 monitors the transcribed RIN4_MR5, a homolog of RIN4 of *Arabidopsis thaliana* that could interact with the previously described effector $AvrRpt2_{EA}$ of *E. amylovora*.

Keywords *Erwinia amylovora* · Positional cloning · Fine mapping · QTL analysis · R gene · Apple breeding

Communicated by E. Dirlewanger

Electronic supplementary material The online version of this article (doi:10.1007/s11295-012-0550-3) contains supplementary material, which is available to authorized users.

J. Fahrentrapp · G. A. L. Broggini · C. Gessler (✉)
Plant Pathology Integrative Biology Zurich (IBZ),
Swiss Federal Institute of Technology, ETH Zurich,
Zurich, Switzerland
e-mail: cesare.gessler@agrl.ethz.ch

M. Kellerhals
Agroscope Changins-Wädenswil (ACW),
Wädenswil, Switzerland

A. Peil
Julius Kühn-Institut, Federal Research Centre for Cultivated Plants, Institute for Breeding Research on Horticulture and Fruit Crops,
Dresden, Germany

K. Richter
Julius Kühn-Institut, Federal Research Center for Cultivated Plants, Institute for Resistance Research and Stress Tolerance,
Quedlinburg, Germany

E. Zini
Fondazione Edmund Mach–FEM,
Istituto Agrario San Michele (IASMA),
S. Michele all'Adige TN, Italy

Introduction

Annually, 71 Mio metric tons of apples are produced worldwide, and cultivars of *Malus* × *domestica* cover an area of 4.8 Mio hectares of land (FAO 2009). Fire blight (FB) is one of the most disastrous diseases in apple cultivation. The causative Gram-negative bacterium, *Erwinia amylovora*, affects apple (*M.* × *domestica*) and pear (*Pyrus communis*) orchards worldwide and causes important losses in production. Forty-six countries reported the presence of the bacterium (Van der Zwet 2006). With an unpredictable periodicity, FB causes severe losses of several million US dollars, and complete pome fruit orchards need to be eradicated. On the one hand, FB can be controlled by pruning and eradication, although pruning is often followed by recrudescence, whilst on the other, the application of different products reduces the risk of complete orchard loss. For

example, bio-control agents (e.g. *Bacillus subtilis* and *Pantoea agglomerans*), chemicals (e.g. prohexadione-Ca and acibenzolar-*S*-methyl) and antibiotics (e.g. streptomycin, oxytetracycline and kasugamycin) are used to protect the crops (Ngugi et al. 2011). The potential success of the control measure chosen depends upon the exact timing of the application. In Europe, the time of flowering is the period of highest infection risk due to the accessibility of *E. amylovora* to apple plants through the natural openings of the nectarthodes, leading to so-called blossom blight. The most effective control method on short term is the application of antibiotics. But due to the problem of selection of resistant strains and potentially remaining residues in soil, honey and the fruit itself, its use is strictly regulated in many countries. However, the selection of resistant strains of *E. amylovora* may result in loss of efficacy of the compound applied (McManus et al. 2002). Moreover, the potential risk increases that the resistance trait from *E. amylovora* will be transferred to human pathogenic bacteria. The risk for humans is even greater if the drug class applied is also used in human medicine, as in the case of streptomycin and tetracyclines. In order to avoid the use of antibiotics, the current research is attempting to achieve natural host resistance to this disease in orchards, either by conventional breeding or by new breeding strategies, including genetic engineering.

Natural resistance to *E. amylovora* has been described in different wild *Malus* species, and it has recently been used in conventional breeding programs (Kellerhals et al. 2009). In order to effectively introgress fire blight resistance into new selections, the strength of the resistance, its heritability and its agronomic usefulness must be evaluated. Breeding programs mostly involve labor-intensive and costly fire blight resistance phenotype assessments through inoculation assays. Knowing the position of the trait in the genome, the genetic marker that is tightly linked to a strong monogenetic resistance locus could be used to speed up the breeding process by reducing the number of genotypes that need to be phenotyped. If resistance genes from *Malus* would be available, they could be used to develop potentially more marketable cisgenic plants that only harbor species' own genes (Schouten et al. 2006).

Quantitative trait loci analysis can be used to investigate the position of monogenic traits. In plant–pathogen interactions, such resistance traits can be ascribed to resistance (*R*) genes. During host–pathogen interactions, the host's R gene products may undergo a so-called gene-for-gene interaction with pathogenic effectors. This interaction may be direct between an R gene product and a pathogen-derived avirulence (Avr) protein (Flor 1971) or indirect. The indirect interaction is explained with the 'guard model' which describes the interaction of the Avr protein with the 'guardee' of the host which is monitored by the R protein (Van der Biezen and Jones 1998). The 'decoy model' was recently introduced in which the guardee is replaced by the 'decoy' but in contrary to the guardee, the decoy is an alternative substrate and not a fitness-enhancing target of Avr (Block and Alfano 2011; Van der Hoorn and Kamoun 2008). The *R* genes can be classified into "seven major structural classes of *R* proteins" (Kruijt et al. 2005). Three of the *R* gene classes are known to be directed against bacteria: kinases such as Pto, CC/TIR-NBS-LRR such as RPS2 and receptor-like kinases such as Xa21 (Hammond-Kosack and Jones 1997). The CNLs (CC–NBS–LRR) are further characterized by the N-terminal conserved non-TIR motifs (Bai et al. 2002), containing the penta peptide 'EDVID' often separating two coiled coils (Rairdan et al. 2008). This region is followed by the P-loop region (synonyms: NBS and NB-ARC) containing 'Walker A' (kinase 1a) and 'Walker B' (kinase 2), RNBS-A, -B, -C and -D, GLPL and 'MHD' (DeYoung and Innes 2006; Van Ooijen et al. 2008).

Quantitative trait loci (QTLs) analyses were performed to identify the genetic regions of putative resistance genes. In apple, major QTLs are found in the *M.* × *domestica* cultivars 'Florina' and 'Fiesta', *Malus floribunda* 821, the ornamental cultivar 'Evereste' and the wild apple *Malus* × *robusta* 5 (MR5) (reviewed in Khan et al. 2011). The strongest QTL reported is located on the LG 3 of MR5 (Peil et al. 2007). It explains up to 80 % of the phenotypic variation between susceptibility [100 % percent lesion length (PLL)] and resistance (0 % PLL) to the strain Ea222_JKI (Peil et al. 2007).

Previous studies evaluated the resistance of MR5 with different results. MR5 has been classified as a "cultivar[s] showing diverse degree of susceptibility according to the inoculated strain" (Paulin et al. 1993). Similar results were found by Norelli and Aldwinckle (1986), who reported that the PLLs varied from 0 % PLL after Ea273 inoculation to 93 % using E4001A (Ea266). The latter strain isolated in Canada was obviously able to overcome the resistance of MR5. Another strain from Canada, E2002A (synonyms Ea3049 and Ea265), applied to segregating population of 'Idared' × MR5, also broke the resistance of MR5 (Peil et al. 2011). The facts that the resistance was broken twice by single strains and the high explanation of the phenotypic variation of the QTL suggest the presence of a resistance type matching the gene-for-gene model.

Recently, Parravicini et al. (2011) described eight putative *R* genes of apple against *E. amylovora* in a small 79-kb region that accounted for the FB resistance of 'Evereste'. They hypothesized that the two most likely functional genes, a kinase- and a CNL-encoding gene, could function similarly to Pto-Prf of tomato, which confer resistance to *Pseudomonas syringae* pv. *tomato* (Martin et al. 1993; Parravicini et al. 2011). However, no complementation experiments to demonstrate the resistance function of these genes have been reported so far.

In this paper, we report the enrichment of the top *M.* × *robusta* 5 LG 3 with new molecular markers followed by a phenotypic evaluation of the FB resistance of 33 recombinants from 2,133 seedlings of MR5 or of a resistant F1 of MR5 as the male parent crossed with susceptible mother plants. The locus responsible for resistance against FB was localized, spanned by MR5 DNA-harboring BACs and sequenced. The sequences were used for gene prediction. The transcribed mRNA was sequenced, and the candidate gene was further analyzed in silico in order to model the mode of function.

Materials and methods

Plant material

The susceptible *M.* × *domestica* cultivar 'Idared' was crossed with the fire blight resistant *M.* × *robusta* 5 (MR5) at the Julius-Kühn Institute (JKI, Dresden-Pillnitz, Germany). A total of 140 plants of this population, called 04208, were previously used for QTL analysis (Peil et al. 2007). A second and third cross of 'Idared' × MR5 was performed (called 09261 and 05211, respectively) to increase the total population size. Three resistant progenies (DA02_2,7, DA02_2,40 and DA02_1,27) of cross 04208 were crossed at Agroscope Changins-Wädenswil (ACW, Wädenswil Research Station, Switzerland) with the susceptible cultivar 'La Flamboyante' (sold as Mairac®) and the medium-level susceptible hybrid 'ACW11303' [(Arlet × Gloster) × Rewena], as well as with an uncharacterized mother plant. These crosses (0802-0804, 0805-0807, 0808; Table 1) were part of a breeding program towards resistance pyramiding at ACW (Switzerland). In total, 2,137 plants were raised in the greenhouse, transplanted to the field and pots, respectively, and used for mapping (Table 1).

Marker enrichment in the region of interest

The region of interest (roi) was defined as the region with a high certainty of containing the postulated resistance gene (s). The roi was narrowed down by deep data mining of the phenotypic and genotypic data of the original dataset generated by Peil et al. (2007). We re-mapped the fire blight resistance trait as a single gene using the method of Durel et al. (2009) to transform the phenotypic data into binary data. Individuals with fewer than three repetitions during each phenotyping experiment or which showed a deviation in the average PLL between experimental repetitions of greater than 30 % were rigorously excluded.

After DNA extraction (Frey et al. 2004), the simple sequence repeat (SSR) and single nucleotide polymorphism (SNP) markers reported to map on the top of LG 3 or which were developed on the top of the LG 3 of the whole genome sequence (WGS) of 'Golden Delicious' (GD) (Velasco et al. 2010) were applied to a subset of 92 individuals in order to verify their location on top of LG 3 (Table 2). The amplification of SSRs was either performed using fluorescently labeled primers in conventional PCR or by using a method described by Schuelke (2000). For this method, the forward primer is prolonged by a nucleotide sequence similar to a universal primer, which is fluorescently labeled. As a universal primer, we used either of the sequences FAM-gactgcgtaccaattcaaa or HEX-gactgcgtaccaattcacg. The SNP loci were amplified by conventional PCR and sequenced using the BigDye Terminator Kit 3.1 (Applied Biosystems, Carlsbad, CA, USA). Amplified fragments of the SNPs and SSRs were analyzed on an ABI3730 or an ABI3310xl capillary sequencer (Applied Biosystems) using the software programs 'Sequencher v4.8' (Gene Codes Corporation, Ann Arbor, MI, USA) and 'GeneMapper v4.0' (Applied Biosystems), respectively. Four SSR markers (FEM18, FEM57, Ch03e03 and MdMYB12) selected on

Table 1 Mapping population with a total number of 2,137 individuals

Population name	Crosses	*n*	Place of phenotyping experiment	Number of recombinants between markers FEM57 and FEM18
0802-0804	'ACW11303' × DA02_x	875	CH	18
0805-0807	'La Flamboyante' × DA02_x	511	CH	2
0808	XY × DA02_2,7	29	CH	0
04208	'Idared' × MR5	284	GER	6
09261	'Idared' × MR5	354	GER	5
AA01	'Idared' × MR5	84	GER	2

MR5 was the source of resistance in all crosses. 'ACW11303' was a medium-level susceptible accession from Agroscope Changins-Wädenswil that was used as a mother plant without a MR5 parentage. 'La Flamboyante' and 'Idared' were susceptible mother plants. DA02_x were resistant progenies of the population 04208 ('Idared' × MR5) phenotyped by Peil et al. (2007). The resistance donors are underlined

x one of three different pollen donor plants from the latter population with the numbers 2,7, 2,40 and 1,27, all progenies of MR5, *XY* an unknown mother plant with unknown susceptibility

Table 2 Molecular markers with primer sequences and allele sizes

Marker	Source	Type	Forward primer	Reverse primer	Quality	Allele(s) in MR5	Allele(s) in Idared
Ch03g07	Liebhard et al. (2002)	SSR	AATAAGCATTCAAAGCAATCCG	TTTTTCCAAATCGAGTTTCGTT	Perfect	125/145	125
Ch03e03	Liebhard et al. (2002)	SSR	GCACATTCTGCCTTATCTTGG	AAAACCCACAAATAGCGCC	Imperfect	184/(204)/206	204/206
EH034548	Norelli et al. (2009)	SNP	GACCAATTTGGATCTTGTAACTCC	CAGCTACCAATGTAGCAGTTAATCC	N/A	N/A	N/A
FEM07	This work[a]	SSR	TGGCTGGTTACTCCTTCACC	AGAGGAGCACAGGGAAATCA	Perfect	183/189	183/191
FEM09	This work[a]	SSR	TGGATGAATTTCAATTGGAGTAA	GGGGTAATAGAATCGCCCATA	Perfect	203	201/203
FEM11	This work[a]	SSR	GGAGGGGAAGTGAGGAGAGT	AGGAGGGGAGGAGGATAATG	Perfect	215	220
FEM12	This work[a]	SSR	CGGGTCGTGGACTAAGAAAA	GAAGCACGCATCATCTCCTT	Perfect	228/236	236/244
FEM14	This work[a]	SSR	GTAGCAATTGGGTTGCGACT	TTTTCCTCAGGTTTCTCAGCA	Imperfect	(179)/181/187	179/187
FEM18	This work[a]	SSR	AGAGCCACCAAAACCTGAGA	CGAAACGTCTCTTTCCTCCA	Perfect	224/234	234/257
FEM19	This work[a]	SSR	ATTCGCTTTCGTGAGGAAGA	GGGGATGCTGCAAGTTTAAG	Perfect	134/150	134/156
FEM47	This work[a]	SSR	CCAAATGTTGGGTTTCCACT	CTACACAGCTGGGGAGGAAG	Imperfect	(194)/209/217	194/217
FEM53	This work[a]	SSR	CTCAGCGGCTCTGTCTTCTT	GCCTTCAAAATTCGATGCTC	Perfect	166	165
FEM57	This work[a]	SSR	ACAGTCGGGTTTGAAGGAGA	CCACCCTGTTGAAGCAATC	Perfect	183/189	189
MR-ARGH31-like	Baldi et al. (2004)	SSR	GATACAGTCGGGTTTGAAGGAG	TTAACTAGCATAGCCATCATGC	Perfect	146/152	152
MdMYB12	Chagne et al. (2007)	SSR	CTCGGCAATCGGTAAAGCTA	TATGAACAGTGAAACCCTAACCCTA	Perfect	150/178	150
rp16k15	This work	SNP	CACACACAAATTTTGCCTTATTC	TCAAATGTGTCATTTCTGCAAC	N/A	N/A	N/A
t16k15	This work	SNP	CTCCTCTTTCAACTCTTTGTCC	TCAAATTGAGGAGAGAGAACTAGTTTC	N/A	N/A	N/A

The alleles given in parentheses were co-amplified regions without polymorphisms. The alleles in coupling with the resistance trait of FB_MR5 are underlined

[a] Developed on GD WGS by IASMA, Instituto Agrario San Michele all'Adige, Italy

GD WGS to spread equally in the roi were used to screen the whole population of 2,137 individuals in order to identify individuals showing recombination events in this interval. The other markers used for fine mapping were only applied to these recombinant individuals.

The genotypic data were processed and mapped using JoinMap3 (Van Ooijen and Voorrips 2001) with a logarithm of odds (LOD) threshold of 10 and manually checked for consistency. For stability and a realistic estimate of distances in cM, the map was calculated using the true genotypic and phenotypic data of each recombinant, and amended with the true genotypic data for the four selected SSR markers from the rest of the whole population. All missing genotypic and phenotypic data were replaced by deduced data under the assumption of a parental (e.g. non-recombinant) genotype/phenotype.

Fire blight phenotyping

Four to 12 replicates (on average 9.5 plants grafted onto rootstock M9) of each recombinant individual were trained to a single actively growing shoot with at least 15 cm in length. After being transferred to the quarantine greenhouse, each shoot was inoculated by cutting the two most juvenile and fully unfolded leaves with scissors dipped into a suspension of *E. amylovora* strain Ea222_JKI at 10^9 cfu/mL. The strain Ea222_JKI (previously named Ea222 Peil et al. 2007, 2008; Richter and Fischer 2000) used in this study was isolated on *Cotoneaster* in the Czech Republic (Richter and Fischer 2000) and is distinct to Ea222 isolated on *M.* × *domestica* cultivar '20 Ounce' (Norelli et al. 1988). Therefore, the strain was renamed from Ea222 to Ea222_JKI (Gardiner et al. 2012). The length of the shoot and degree of necrosis were recorded 21 days after inoculation and transformed into the percentage of lesion length per shoot (PLL). The binary data transformation method described by Durel et al. (2009) was used. The experiments with recombinants of the populations 04208 and 09216 (all 'Idared' × MR5 progenies) were carried out at JKI (Quedlinburg, Germany). The greenhouse conditions were the same as described by Peil et al. (2007). MR5 and 'Idared' plants were used as positive and negative control plants, respectively. Recombinants from crosses with 'La Flamboyante'/'ACW11303' (Table 1) were inoculated at ACW (Wädenswil Research Station, Switzerland). At ACW, the resistant parents (DA02_2,7, DA02_2,40 and DA02_1,27) and the susceptible *M.* × *domestica* cultivar 'Gala Galaxy' were used as controls. The greenhouse conditions during the whole experiment were 22 °C/18 °C (day/night) and 70 % relative humidity; at JKI, no additional light was provided, at ACW 10 h 400 W light per day was provided during growth of the shoots to the experimental length. The raw data obtained in Germany and Switzerland were handled separately until they were transformed into binary values.

Isolation, sequencing and analysis of the region of interest

A BAC library harboring MR5 DNA in pCC1BAC (Invitrogen, Carlsbad, USA) with HindIII cloning sites was constructed by Amplicon Express, Pullman, USA. A total of 36,864 BAC clones with an average insert size of 145 kb, representing approximately 3.5 times the diploid apple genome (Velasco et al. 2010), were spotted onto nylon filters. Filter hybridization was performed using a radioactively labeled probe designed on the sequence flanking the SSR marker Ch03e03 (forward primer: ggcatttcttgctcttctgc, reverse primer: ttggcagctgcaacatagac). The procedure was described in detail by Patocchi et al. (1999), although New England Biolab's 'NEBlot®Kit' (Ipswich, USA) was used to label the probe. The BAC DNA was extracted, and the RP and T7 extremities of the positive BACs were sequenced (Galli et al. 2010). After designing the primer using Primer3 (Rozen and Skaletsky 1999) on these sequences, the amplicons were amplified on the population's parents and sequenced to detect polymorphic sites such as SNPs in order to verify the source area of the BACs on LG 3. The BAC insert sizes were estimated by both the summation of the size of the bands obtained by HindIII digestions and pulsed-field gel electrophoresis after NotI digest (Broggini et al. 2007). The BAC clones covering the resistance locus were sequenced by 454-pyrosequencing (GS FLX (Titanium), Roche, Basel, Switzerland) and assembled using Newbler (v2.3, module: gsAssembler, Roche) using a minimum overlap of 40 bp with 90 % identity. Additional Sanger sequencing for gap closure was applied. The sequences and contigs of 454 sequencing were assembled with Sequencher (v4.8, Gene Codes Corporation). The ORFs were predicted using FGENESH with algorithms for tomato, *Arabidopsis thaliana* (dicot plants) and *Vitis vinifera* (Salamov and Solovyev 2000). The predicted proteins were further analyzed with the BLASTx (Altschul et al. 1990) and MotifScan to predict their function and specific motifs (Pagni et al. 2007).

Analysis and expression of candidate genes

Candidates belonging to one of the three classes of resistance genes against bacteria were analyzed using consensus sequences of known motifs, i.e. in the case of CNL for the presence of the kinases Walker A and B, the pentapeptide EDVID or the NB-ARC 'switch' MHD (Van Ooijen et al. 2008; Walker et al. 1982). Coiled-coil regions were predicted using Coils/pCoils (Lupas 1996; Lupas et al. 1991). Constitutionally transcribed total RNA was extracted from young, healthy (non-infected) apple leaves of MR5, 'Idared' and 'Golden Delicious' according to the instructions given by the manufacturer (Concert™ Plant RNA Reagent, Invitrogen). Potential DNA contamination was removed using the TURBO DNA-free™ Kit (Applied Biosystems). The

quality and yield were determined and visualized with the NanoDrop spectrophotometer and agarose gel electrophoresis. The absence of DNA was proven by PCR with RNA as template and PCR with primers for elongation factor with cDNA as template (Szankowski et al. 2009). The RNA was reverse transcribed to cDNA using 'Maxima™ Reverse Transcriptase' (Fermentas, Waltham, USA) with oligo dT. For gene expression, primer pairs at the start (ATG-forward: ATGGGGGGAGAGGCTTTTCTTGTGGCATTCCTCCAAG) and stop codons ($TGA_{(tomato)}$-reverse: TCAAATCATCTTCCAATCTATATCTATGTAAG and $TGA_{(dicot\&vitis)}$-reverse: TCACGGGAAATCGACCACCACACCTGGCC) were designed to amplify the whole coding sequence. Amplified DNA was cloned into pTZ57/RP (Fermentas) and transformed into ccdB survival cells (Invitrogen). Inserts of plasmids were sequenced by primer walking.

Results

Region of interest

By mining the phenotypic data, which were previously produced by Peil et al. (2007), 36 individuals of the total 140 were excluded from transformation into binary values due to low number of replicates, no experimental repetition or high average variability in PLL. The averages of phenotypic data from the remaining 104 individuals were transformed into binary data according to Durel et al. (2009), excluding a further 20 individuals around the median PLL. The fire blight resistance co-segregated with Ch03e03 in 83 individuals, and in one individual, a recombination was observed, indicating the position of the fire blight resistance locus as being 1.2 cM distal to Ch03e03. Thus, we determined the roi as being the interval from a locus between Ch03g07/Ch03e03 and the close end of LG 3.

Marker enrichment in the region of interest

Ten markers (EH034548, FEM09, FEM11, FEM12, FEM14, FEM18, FEM53, FEM57, MR-ARGH31-like and MdMYB12) were validated on a subset (n=92) to map to the top of LG 3 (Fig. 1a). It was assumed in accordance to roi definition that the fire blight resistance locus is located around Ch03e03. Four markers (FEM57, FEM18, MdMYB12 and Ch03e03) which span the roi in regular intervals were used to screen a total of 2,137 progenies. Four individuals showed different SSR alleles to the estimated parental ones and therefore were classified as outcrossers and excluded from further analysis. Thirty-three individuals showed a recombination in this interval. These 33 individuals were used for all subsequent studies. The validated markers plus three additional ones (FEM07, FEM19 and FEM47) were applied to the recombinants and mapped (Fig. 1b). The top 1.5 cM of LG 3 was now defined

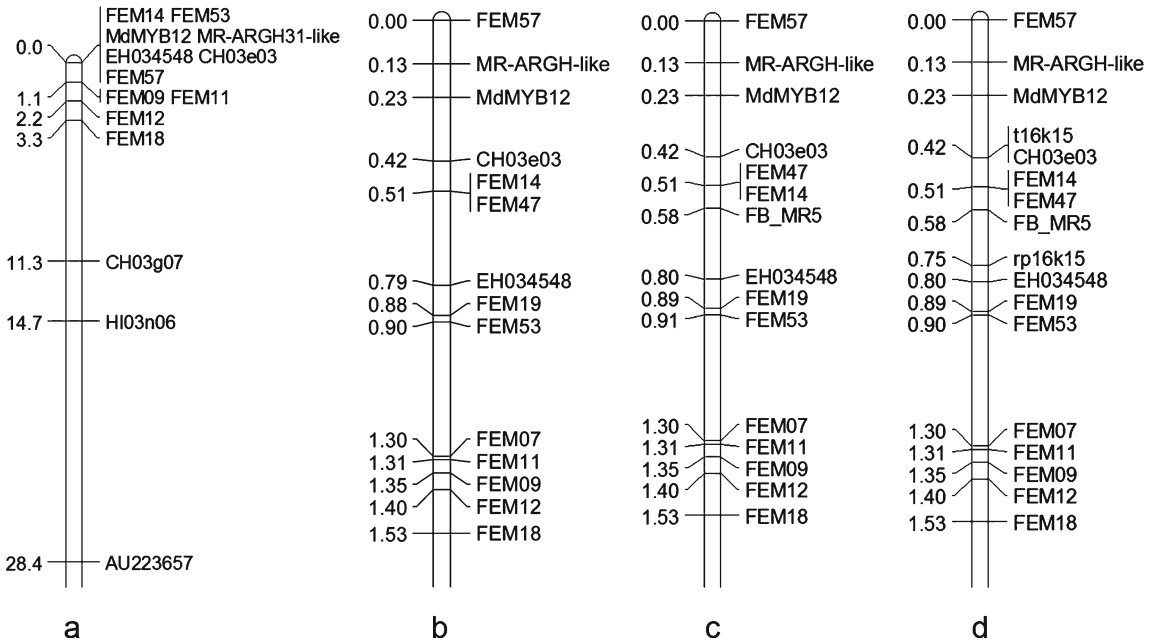

Fig. 1 Marker enriched tops of linkage group 3 of *Malus*× *robusta* 5. **a** Mapped markers on top of linkage group 3 of a subset of 'Idared'× MR5 population, n=92. **b** Fine mapping of the top 1.5 cM of linkage group 3 on n=2,133 individuals. **c** FB_MR5 (fire blight resistance locus) mapped as a single gene after the transformation of phenotypic data into binary data; five recombinants were found between FB_MR5 and EH034548 and one between FB_MR5 and FEM47/FEM14. **d** The genetic map as in **c**, but SNPs were added which were developed on BAC end sequences (rp16k15 and t16k15)

by 14 markers, with a median distance between each other of 0.05 cM and the biggest inter marker distance of 0.4 cM between FEM53 and FEM07. The most distal marker of LG 3 was FEM57, which was found to lengthen the linkage group by 0.4 cM compared to the previously published map of MR5 (Peil et al. 2007).

Phenotyping and mapping of the resistance locus

Following *E. amylovora* inoculation, the resistant parents MR5, DA02_1,27 and DA02_2,40 showed no lesions at all, whereas DA02_2,7, which was also carrying MR5 resistance, showed a PLL of 7 %. The susceptible cultivar 'Idared' showed an average PLL of 52 %, the parent 'ACW11303' a PLL of 14 % and 'La Flamboyante' a PLL of 29 %. The susceptible standard control 'Galaxy' showed a PLL of 77 %. The tested recombinant genotypes of populations 04208, 09261 and 05211 (all progenies of 'Idared' × MR5) showed lesions between 0 and 77 % of the shoots with a median of 1 % and an average of 20 % (Fig. 2a). The PLL of the recombinants of crosses 0802-0807 ranged from 0 to 25 % (Fig. 2b), with a median of 12 % and an average of 11 %.

Single gene analysis with transformed phenotypic data indicated the position of the resistance locus between markers FEM14/FEM47 and EH034548 flanking an interval of 0.29 cM (Fig. 1c). This interval was defined by six recombinants, five of which were between the resistance locus and EH034548 (Fig. 3). Linkage mapping placed the resistance locus at a map position of 0.58 cM away from

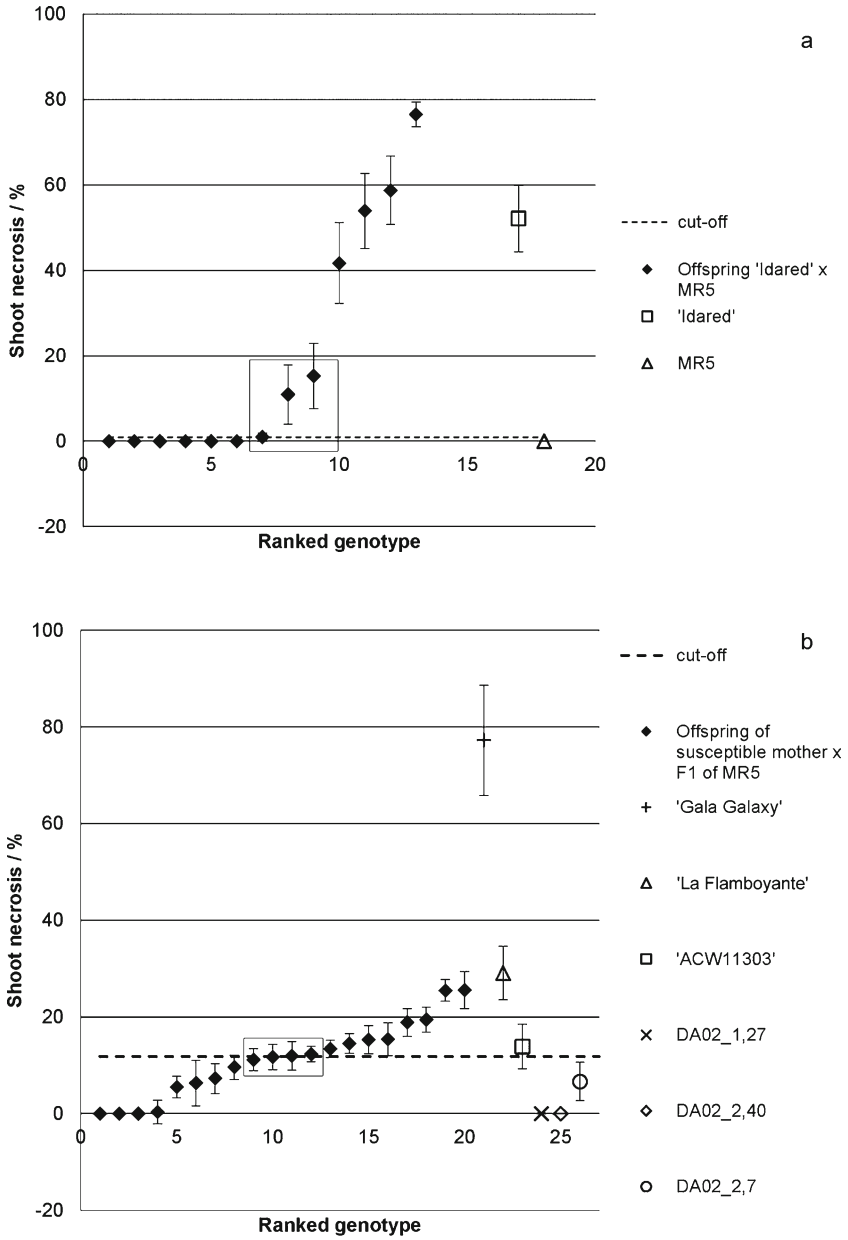

Fig. 2 Necrosis values in percent of shoot length of the phenotyped recombinants and controls. **a** The results of the experiments conducted at JKI (Germany). **b** The results obtained at ACW. All values are averages of 4–13 repetitions of each genotype. The *filled diamonds* show the recombinants, the *different unfilled symbols* show the controls. The data points, which were shown in *rectangles*, were excluded from the mapping process due to their position close to the cutoff value (median of average necrosis values of the genotypes tested). The cutoff values (**a** 1 % and **b** 12 %) between data transformation classes 'resistance' and 'susceptibility' are given as a *black line*. The *bars* display standard errors

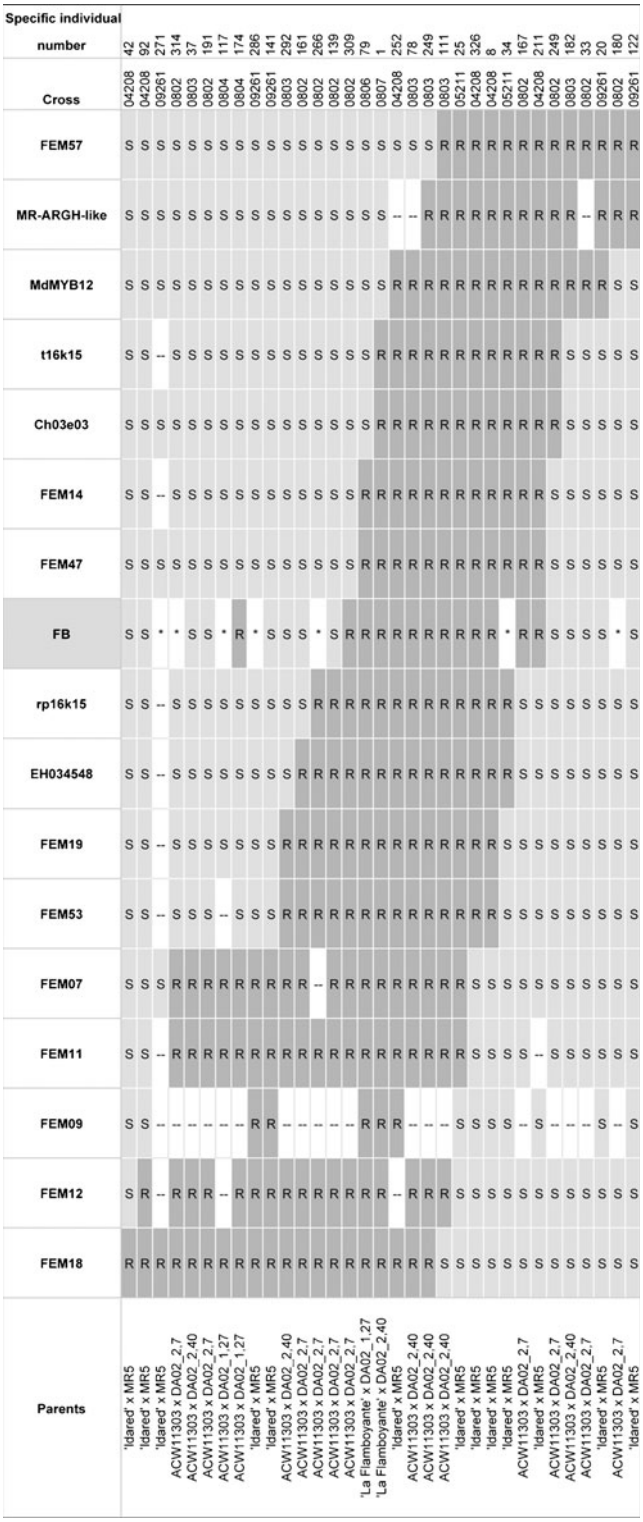

Fig. 3 Genotypic and transformed phenotypic data of 33 progenies of *Malus*×*robusta* 5. The progenies were recombinant between FEM57 and FEM18. The R alleles (*dark grey*) were in coupling with phenotypic fire blight resistance (FB), S alleles in repulsion with FB (*light gray*); *asterisk* indicates excluded phenotypic values (*white*) due to their position close to the cutoff value, *double-hypen* stands for missing data (*white*). For gene mapping using the software JoinMap4, the missing data, which were not located at the border of a crossing over, were substituted by the allele of the missing marker which does not produce a crossing over towards his neighboring markers. The missing data of FEM09 in all crosses with 'ACW11303' were due to the lack of a polymorphism between 'ACW11303' and MR5 for this marker

FEM57. One individual (0804-174, Fig. 3) was genetically susceptible, but its phenotype was resistant (0 % PLL).

Chromosome landing and sequencing of the resistance region

A total of six clones were detected by hybridization of the genomic library with the probe derived from the sequence flanking Ch03e03. One of which, the BAC clone 16k15 (insert size 162 kb), carried flanking markers FEM14/FEM47, as found by means of PCR (data not shown). The extremities of this BAC clone were sequenced, and SNP markers were developed on these sequences (t16k15 and rp16k15) and mapped to LG 3 (Fig. 1d). The marker rp16k15 was located one recombination closer to the FB resistance locus than EH034548, whereas the other extremity, t16k15, was located on the opposite side of the FB resistance locus. Therefore, the BAC 16k15 spanned the whole region of resistance locus. The marker rp16k15 was used as a new probe, which hybridized to another BAC named 72i24 (insert size, 250 kb). This BAC clone carried the alleles in repulsion to the resistance locus of the flanking markers FEM14/FEM47 and rp16k15. Consequently, 72i24 spanned the same region of the homolog chromosome, which does not contribute to the resistance of MR5 against *E. amylovora*. Both BAC clones 16k15 and 72i24 were sequenced, resulting in a total of 162 kb in 4 contigs and 251 kb in 22 contigs, respectively. The sequences of markers FEM14, FEM47, Ch03e03, t16k15 and rp16k15 were anchored to the four contigs of 16k15 (accession numbers HE805489–HE805492). The t16k15 and Ch03e03 sequences were anchored to one contig, whereas each of the remaining three contigs comprised one of the markers FEM14, FEM47 or rp16k15. The sequences of 72i24 were used as a database to detect the homolog sequences of the candidate genes.

Gene prediction

Three gene sets were predicted on the sequence of BAC 16k15 using different presets of FGENESH. The FGENESH algorithm for (1) *dicot* plants (based on *A. thaliana*) predicted 41 open reading frames (ORFs); (2) the algorithm for tomato predicted 47 ORFs and (3) the algorithm for *V. vinifera* predicted 25 ORFs. Genes with putative resistance function were identified with BLASTx (Tab. S1). A2.6 was found to be homolog to a putative serine/threonine kinase of

M. × *domestica* (AEJ72559) but only with poor query coverage of 25 % (68 of 272 amino acids). The sequences A4.11, A2.10 and V4.9 were found to share homologies with a TNL (TIR NBS-LRR) of *Arachis hypogaea*. Further A3.5 and T3.5 were homolog to a putative disease resistance RPP13-like protein 1-like of *V. vinifera*. All these BLASTx results could not be substantiated with MotifScan. Additional two predicted ORFs in each gene set (A2.2, A2.7, T2.2, T2.7, V2.1, V2.3) were similar to NB-ARC and LRR domains, which are common motifs of plant resistance genes. Only in A2.7, T2.7 and V2.3 the motifs (i.e.NB-ARC, Walker A and B, MHD) indicating functionality could be identified. The three predicted genes were located on the same sequence; all three shared the same start codon but showed two different splicing profiles, with the splicing and the stop codons predicted by the FGENESH*dicot* and *vitis* algorithms being the same. The mature transcript predicted by the FGENESH*dicot* and *vitis* algorithms contained one exon of 4,089 bp followed by two shorter exons (69 and 153 bp, respectively). The FGENESH*tomato* algorithm predicted an unspliced transcript of 4,167 bp with a different stop codon position. Therefore, the gene encoded in this transcript, which was predicted in two splicing variants, was the only putative candidate resistance gene.

Expression of the putative candidate gene

The transcript of the putative resistance gene was amplified with primers on the ATG codon as well as two different primers designed on the two differently predicted stop codons. Amplicons were obtained using both primer pairs (ATG-forward, $TGA_{(tomato)}$-reverse, $TGA_{(dicot\&vitis)}$-reverse) and were cloned into the pTZ57/RP vector for sequencing, showing that the introns predicted by FGENESH*dicot* and *vitis* parameters were not predicted correctly. The full ORF of the transcript T2.7 was sequenced and called *FB_MR5* (accession number: CCH50986.1), being the only candidate resistance gene of *M.* × *robusta* 5 against *E. amylovora*. A homolog nucleotide sequence for this gene could not be amplified from BAC 72i24 nor from genomic DNA of the susceptible cultivars 'Idared' and 'Golden Delicious'.

Structure analysis of FB_MR5

Analysis with MotifScan and Coils/pCoils of FB_MR5 predicted N-terminal two coiled-coil regions (amino acid (aa) positions of around 60–75 and 120–150) separated by an EDVID-like motif (aa position 80) followed by a predicted NB-ARC (start at aa 172). The latter domain comprised the following motifs: Walker A and Walker B, RNBS-B, GLPL and MHD (Fig. 4a). At the C-terminal region behind the NB-ARC, three LRRs were identified. The imperfect LRRs fulfilled the consensus LxxLxxLxLxxT and LxxLxxLxLxx(T/C)xxLxxIPxx, respectively, but lacked one 'Lxx' repeat at the start of the motif (Jones and Jones 1997). The minimal consensus of LRRs 'LxxLxL' (Kajava and Kobe 2002) increased the number of putative LRRs to eight. All eight LRRs were located behind MHD. In a de novo search for repetitive motifs using the software MEME (Bailey and Elkan 1994), we identified the consensus 'LxSL[EKR]ELxIx[GD]CxSL' (Fig. 4c) in the amino acids behind NB-ARC (aa 732-1387). This consensus was refined to an LRR-like motif 'LxxLx[IL]xxCxxLxxL', of which a perfect version was present three times and an imperfect version 20 times (Fig. 4b). A hydropathy plot (Fig. S2) resulted in values between −2 and +1.5; therefore, no indications were found for membrane-spanning elements (Kyte and Doolittle 1982). A protein BLAST analysis against NCBI's reference proteins found highest similarity (49 %) of FB_MR5 to the predicted CNL of *Populus trichocarpa* (Accession number XP_002328224). Protein BLAST against proteins predicted on GD genome sequence resulted in homology to three predicted peptide sequences on LG 3 of GD with highest similarity to MDP0000269188 of 62 % in ClustalW pairwise alignment. MDP0000269188 comprises a NB-ARC domain lacking the motifs Walker B and MHD, and Walker A is present only in imperfect version. BLASTn of the candidate against apple ESTs identified GO532686 isolated from bark of 'Royal Gala' covering 15 % of the query with 93 % identity.

Discussion

Firstly, the original data of Peil et al. (2007) were investigated and used to precisely map the FB resistance trait as a single gene using the method published by Durel et al. (2009), which accounts for a certain degree of biological variance in these fire blight inoculation experiments. Then, the mapping population was increased to 2,133 individuals, which allowed a mapping resolution of 0.05 cM. A total of 33 individuals were identified showing recombination in the region of interest on top of LG 3, where the resistance resides, and the resistance of these recombinants towards FB was assessed. This was done because only such individuals are informative in locating the exact map position of the resistance. Eventually, a genomic library of *M.* × *robusta* 5 was used to isolate the region of resistance, which was then sequenced and used to identify a putative resistance gene named *FB_MR5* belonging to the gene family encoding CNLs. Transcripts of this gene were reverse transcribed into cDNA, cloned into a vector and then sequenced to identify the correct splicing profile of this gene.

Fire blight remains the most feared disease in apple production as its appearance is highly erratic and the damage can potentially endanger every orchard. Classical

a

```
                                                                  CC1
MGGEAFLVAFLQVLVDKLAHREVFKYFGLVKGVDQKLKKWSATLSAIGAVLNDAEERQLTAKNNTLKLWLEDLRDLAFDV
EDVID                                        CC2
EDVLDKYATKMLKRQIQHAHSRTTSKLWNSIPDGVFNFNMNSEIQKISERLQEISEQKDQLNLKIDTGALTTRARRNISP
           NB-ARC                        Walker A
SSSQPDGPVIGRDEDKRKIVELLSKQEHRTVNFDVVAIVGMAGVGKTTLAGQVLNDMVATQTFQPAVWACVSDDFNLERV
                                    Walker B                    RNBS-B
TKQILESITSRQCTTEDYNKVQDYLHKELAGKKFLIVLDDVWKTCSYGEWMKLQSPFRDGAQGSKIIVTTRDTDVSKMMG
                                                GLPL
AATLVHNLEPMESSVCLQVFEQHAFLNSNDDKPPNYELLKEKIAAKCRGLPLAARTLGGVLLRKDTYEWEDILNNKLWSL

SNEHDILPVLRLTYFYLPSHLKRCFAYCSILPNDYEFEEKQMILLWMAEGFILPRPEDKKQIEDLGADYFRDLVSRSLFQ
          MHD
KSTKCISKYVMHDLIGDLARWAAGEICFRLEDKQNDDGEQLRCFPKARHSSYIRGLSDGVKRFEVFSELKYLRTFLPLRK
                                          LRR1                  LRR2
DSFWNYLSRQVAFDLLPKLQYLRVLSFNCYKITELPDSIGDLRYLRYLDLSYTDITSLPKSTSTLYNLQTLILEGCSKLK

ALPIDMSNLVNLRHLNNSNVSLLEDMPPQLGRLVNLQSLTKFVVSGGGGGDRSGIRELEFLMHLRGTLCISRLENVTDVE

DAQRANLNCKER
```

b

```
LDSLVLEWSHSSDTRETESAVLDMLQPHTK
LKELTIKSYAGKEFSSWVGVPLFSNMVLVRLEECNNCLSLPPLGKLPH
LKELYIRGMNAVESVGAEFYGECSLPFPLLETLEFVDMQHWKVWLPFQTDHRGSVFPC
LKTLLVRKCSKLEGKLPENLDS
LASLEIVKCEELLVSIANYKQ
LRQLNIDGCKGVVHTAAKVEFEL
LESLYLSNISELTSLQTGELCRNGLNM
VRDLKINGCEELTSSLKNEAILLQQLIS
LGRLEIEDNSLLVEELGKEADELLQLQILGCK
LEFLKLKKCKNLLKLPEGLNQLSS
LQELRIHECSSLVSFPDVGLPPS
LKDIEITECHSLIYFAKSQIPQN
LRRIQIRDCRSLRSLVDNEAVGSCSSSSHNC
LEYLNIERCQSLTLLSLSDQLVRA
LRELDIYDCEQLEFLAPDGLFCNNTNYF
LENFRIRRCQNLKSLPRLSGGIRGSN
LREIRITDCDRLEALPEDMHNFNS
LEKLIIDYREGLTCSFPANLTS
LMIWKVKSCKSLWELEWGLHRLTS
LRYLWIGGEDPDMVSFPPDMVRMETLLPKS
LTELSIGGFPNLKKLSSKGFQFLTS
LESLELWDCPKLASIPKEGLPLS
LTELCIYGCPVLKERCQPGKGRYWHKISHIPYIDIDWKMI
```

c

Fig. 4 The structure and motifs of FB_MR5. FB_MR5 was identified as CNL which was subdivided in **a** coiled-coil (CC) region and NB-ARC domain as well as **b** LRR-like motifs. **a** The CC regions and conserved motifs are *underlined*, NB-ARC is highlighted in *gray*. **b** Amino acids of the LRR-like motif are highlighted: L, *blue*; I, *yellow*; C, *green*; *gray*, residues that did not follow the consensus 'LxxLx[IL]xxCxxLxxL'. Classically known (Jones and Jones 1997) cytoplasmatic LRRs are *underlined*. **c** Motif logo created by de novo motif finder 'MEME' with the general consensus LxxLx[IL]xxCxxL. The *colors of the letters* indicate equal characteristics of amino acids; their height indicates the number of occurrences at the specific position

breeding is currently introgressing resistance from sources such as *M.* × *robusta* 5 or *M.* × *domestica* cv 'Evereste'. As the selection of fire blight-resistant offspring by phenotypic reactions upon inoculation is labor-intensive and costly, DNA markers linked to the resistance trait represent an alternative selection tool for marker-assisted selection.

Moreover, in order to evaluate the usefulness and durability of such resistance trait, the mechanisms of the plant–pathogen interaction need to be understood by studying the proteins involved in pathogen recognition, as well as the defense cascade induced. Finally, once a FB resistance gene has been identified and its functionality has been proven, it can be introduced into a commercial cultivar via genetic modification aiming to produce a cisgenic FB-resistant apple cultivar where selectable marker genes are absent and only species' own genes are present with their own regulatory sequences from a crossable donor (Gessler 2011). Cisgenesis has the additional benefits of maintaining all of the qualities that make a particular cultivar popular, whereas in classical apple breeding, a new cultivar is created with its own unique qualities (Schouten et al. 2006). Recently, the first cisgenic apple cultivar carrying the *HcrVf2* scab resistance gene was reported (Vanblaere et al. 2011), and such a GM cultivar could be further improved by adding a fire blight resistance trait. Thus, for these reasons, we attempted the positional cloning of the resistance against fire blight present in *M.* × *robusta* 5 (Peil et al. 2007). This resistance can be attributed to a single locus being a QTL explaining 80 % of the phenotypic variation, and putatively, it undergoes a gene-for-gene relationship as this resistance was broken twice by two different Canadian *E. amylovora* strains (Norelli and Aldwinckle 1986; Peil et al. 2011). The resistance may still be useful as single major *R* gene in Asia, Europe and other countries, where such strains have not yet been reported, or as pyramided FB resistance also in North America.

Mapping of the fire blight resistance locus

Re-analysis of the phenotypic and genotypic data produced by Peil et al. (2007) eliminated 7 out of 11 individuals, which showed a disagreement >30 % of the averaged PLLs between experimental repetitions. That confirmed that this method was useful and should be used in subsequent studies. The data of the remaining four individuals did not cause any double crossing overs. When the phenotypic data were mapped, only one recombinant individual indicated a position of the FB locus distal to Ch03e03 (equal to 1.2 cM), which is in contrast with the previously mapped distance of 9 cM (Peil et al. 2007). Due to the remaining disagreement of the re-mapped FB locus regarding the maximum of the QTL (about 5 cM), the reduced dataset was used for a re-QTL analysis. This resulted in a LOD curve (data not shown) that increased towards the LG end, similar to that demonstrated for 'Evereste' (Durel et al. 2009) missing a maximum peak. This was probably caused by the lack of markers distal to Ch03e03 (MR5) and Hi23d11y_E ('Evereste'), respectively. After adding two new markers to LG 12 in 'Evereste' (M45TA_403c_E and M35TA_256s_E), 2.7 and 3.7 cM distal to Hi23d11y_E, respectively, the maximum of the LOD curve corresponded to the position resulting of the single gene analysis of FB_E (Parravicini 2010). In order to solve the discrepancy in MR5 between the QTL curve maximum and the FB locus, the mapping population was increased to 2,133 individuals, and the roi was enriched with 15 new markers. As phenotyping for resistance/susceptibility against fire blight for a large number of progenies is cumbersome and costly, and since only individuals showing a recombination in the region of interest are informative in determining the map position of the resistance locus, we only considered these individuals with a crossing over in the roi that clearly encompassed the locus of the resistance trait inherited from MR5. New QTL mapping of this data was not feasible because only the recombinants were phenotyped. The population used to map the resistance of MR5 was obtained by combining different populations resulting from the crossing of resistant parents (either MR5 or one of its resistant F1) with highly susceptible and medium-level susceptible parents. The susceptibility levels of the mother plants (average PLL of 'Idared' 52 %, 'La Flamboyante' 29 % and 'ACW11303' 14 %) were clearly distinguishable from the resistance of MR5 and the resistant progenies of MR5 (average PLL 1.75 %) used as male plants. Separate single gene analysis of the recombinants produced by each mother plant showed the same position of the fire blight resistance locus and therefore acknowledged that the different populations could be pooled (data not shown).

The application of the method suggested by Durel et al. (2009) to transform the phenotypic data into binary data eliminated three and four individuals, respectively, from each inoculation experiment (Fig. 2a, b). The individual 0804.174 which showed a phenotype incongruent to its phenotype was not excluded by this data treatment. Taking into account their standard error of the mean (SEM) as an indicator of the putative switch of ranked classes after experimental repetition if the SEM crosses the cutoff value, one additional individual on both sides of the four excluded individuals of ACW subpopulation should be eliminated. Also, 'ACW11303' had a large SEM ranging across the threshold towards the resistant category, indicating the lower susceptibility level of this parent. The resistance contributed by 'ACW11303' might be a quantitative resistance what could be interpreted regarding the more linear distribution of the PLLs of the recombinants from the ACW population (Fig. 3). This quantitative resistance may also explain the resistant phenotype of 0804.174. Since qualitative resistance masks usually quantitative resistance (McDonald and Linde 2002), we used the data obtained from progenies of 'ACW11303' also for single gene mapping. But also, if all individuals would have been excluded being offspring of 'ACW11303' the genetic window of the FB resistance locus would be in the analyzed region ranging from Ch03e03/t16k15 to rp16k15 which is covered by the sequence of BAC 16k15. Furthermore, the two individuals defining the

genetic window where the resistance was mapped (0806_79 and 04208_211) were both clearly resistant with 0 % PLL.

Using the data from all 33 recombinants and the deduced phenotypic data from the non-recombinant individuals enabled single gene mapping of the position of the FB resistance locus between rp16k15 and FEM14/FEM47. This position confirmed the position of the previously published QTL and dislocated the position of the single gene analysis in the same study, which was located 9 cM distally of Ch03e03 (Peil et al. 2007). The nearest flanking markers of the resistance locus FB_MR5 (FEM14/FEM47 and rp16k15) were 0.23 cM distant from each other, which is comparable to other chromosome landing approaches for apple. Galli et al. (2010) described a distance of 0.5 cM between flanking markers (GmTNL1 and ARGH17) of the Rvi15 scab resistance locus, and supplemental two co-segregating markers (41A24T7 and 43M10RP) were developed. The flanking markers developed in the study of the QTL in 'Evereste', ChFbE01 and ChFbE08, were 0.18 cM apart, and marker ChFbE02-07 was co-segregating with the resistance locus (Parravicini et al. 2011). We did not identify any co-segregating marker in the present study useful for marker-assisted breeding; however, the use of two flanking markers located very close to the locus of interest is highly precise. If both flanking markers are evaluated to be in coupling with the trait locus, the trait locus will be present completely with the probability of 5.29×10^{-6} of a false positive due to double crossing over in 0.23 cM.

BAC sequencing

Hybridization with Ch03e03 identified a BAC (16k15) spanning the resistance region. The assembly of sequenced 454 pyrosequencing data of 16k15 plus an additional gap closure step via Sanger sequencing produced four contigs. Additional re-sequencing of the same BAC using Illumina technology (HiSeq 2000) could not close the remaining gaps either (data not shown), but provided additional evidence confirming the previously generated sequences. The resulting number of nucleotides (162 kbp) in the four contigs was almost equal compared to the number of base pairs of BAC 16k15, measured via PFGE and fingerprinting, which indicated that the gaps were relatively small. The region spanned by BAC 16k15 was also compared to the corresponding section of GD WGS (Velasco et al. 2010). This section ranged from marker t16k15 to EH034548 and contained approximately 23 % gaps (meta-contigs, without real sequenced data, Fig. S3), which indicated that the region was hardly sequenceable. Due to these results and the experience that non-sequenceable regions were mostly non-coding homo-polymeric stretches or repetitive elements such as SSRs, we ended the investigations to complete the BAC sequence.

Gene predictions

Few conclusions can be made on the gene density recorded in this work. Genes are not thought to spread equally within the genome (Barakat et al. 1999; Bernardi 2004) so it is difficult to compare densities between different genomic regions. For example, in 48.6 kb in a GMAL 2473-derived BAC clone, Galli et al. (2010) found 17 ORFs, 3 of which were predicted to be related to resistance. Regarding FB resistance, Parravicini et al. (2011) predicted a total of 23 genes in 78 kb of the ornamental apple cultivar 'Evereste' using the FGENESH*tomato* algorithm; 8 of these 23 genes were related to resistance. The number of genes predicted in our study (47 genes in 162 kb) was, on the one hand, in accordance with the latter two publications, although the number of *R* gene candidates with only one candidate, *FB_MR5*, was very low. On the other hand, the GD sequence predicted 0.78 genes per 10 kb (gene density), i.e. just one quarter of what we found (Velasco et al. 2010).

However, it is noteworthy that resistance genes are usually found in clusters of *R* genes and *R* gene analogs (*RGA*s) (Meyers et al. 2003). This is also true for many *R* genes and *RGA*s of apple (Baldi et al. 2004; Broggini et al. 2009; Calenge et al. 2005). In our case, we identified a sole *R* gene that was not embedded in a cluster of paralogs. This may be explained by the lack of long-term coevolution between *Malus* species and *E. amylovora* since *E. amylovora* came from North America and *Malus* species have their origin in Central Asia (Juniper and Mabberley 2006).

LRR-like structure

In total, 23 hydrophobic LRR-like motifs were found to fit the consensus 'LxxLx[IL]xxCxxLxxL'. Furthermore, L could be substituted by C and I due to their shared characteristics: aliphatic (I and L) as well as hydrophobic and non-polar (I, C, L) (Kyte and Doolittle 1982; Taylor 1986). The HHpred analysis (Söding 2005) detected a strong similarity between the secondary structure (*E* value<1.1E-36) of FB_MR5 and the proteins 3rgz_A (brassinosteroid-insensitive protein, *A. thaliana*) and 1ziw_A (TOLL-like receptor 3, *Homo sapiens*). Both are folded into a horseshoe-shaped tertiary structure like classic LRRs (Kobe and Deisenhofer 1994), suggesting that the same structure is formed by the LRR-like protein part. In addition, the protein homology modeler 'SWISS-MODEL' (Arnold et al. 2006) indicated a horseshoe-shaped protein for the LRR-like part of FB_MR5 (Fig. S4). The number of LRR-like motifs is in accordance with the number of LRRs in other described CNLs: Galli et al. (2010) found 15, 15 and 29 imperfect LRRs in three Rvi15 (Vr2) resistance protein candidates, respectively. Parravicini et al. (2011) found 11–12 imperfect LRRs in MdE-EaN, a CNL of *Malus* cultivar 'Evereste'.

The functionality of the proteins in these latter two studies has not yet been validated. Examples of functional resistance proteins of less related species possess between 14 (RPS2, *A. thaliana*) and 27 (L6, *Linum usitatissimum*) LRRs (Jones and Jones 1997), respectively. These models and facts support the putative functionality of the LRR-like motif.

Assumed mode of function of FB_MR5

FB_MR5 was classified as belonging to the family of CNL proteins. The members of this protein family are often resistance genes. In the plant resistance gene database 55 CNLs are listed; 26 of these act against fungi, 10 against oomycetes, 7 against viruses, 6 against bacteria, 5 against nematodes and 1 against aphids (Sanseverino et al. 2010). The six CNLs that act against bacteria are Bs2, Prf, RPM1, RPS2, RPS5 and Xa1. Bs2, RPM1 and Xa1 were not reported as CNLs but as NBS-LRRs without a coiled-coil region (Grant et al. 1995; Tai et al. 1999; Yoshimura et al. 1998). Therefore, only Prf, RPS2 and RPS5 can be directly compared to FB_MR5. These three proteins confer the resistance following the decoy/guard model being in the position of the guard. The Prf of tomato is known as the guard of Pto, a serine–threonine kinase that recognizes AvrPto and AvrPtoB of *P. syringae* pv. *tomato* (Pedley and Martin 2003). The RPS2 from *A. thaliana* is activated by the effector AvrRpt2 via the cleavage of RIN4, the decoy/guardee of RPS2 (Caplan et al. 2008). Furthermore, RIN4 is also a decoy/guardee of RPM1, which is activated after phosphorylation of RIN4 by AvrRpm1 or AvrB. The RPS5 guards the protein kinase PBS1, which is AvrPphB susceptible (Caplan et al. 2008). The corresponding Avr proteins are all secreted to host cells via type three secretion systems (T3SS). This implies also for the pathosystem *E. amylovora*/*Malus* an interaction between host *R* gene and pathogen T3SS or T3SS-delivered effector as suggested also for the fire blight candidate resistance proteins of the cultivar 'Everest' the kinase MdE-EaK7 and the CNL MdE-EaN (Parravicini et al. 2011). The authors hypothesized that they function in a similar system to that of Pto and Prf of tomato. Inferentially, the two possible hypotheses for the function of FB_MR5 are that (1) the effectors or the T3SS itself could be targets of the candidate resistance gene product FB_MR5 or (2) that FB_MR5 has an undiscovered decoy/guardee that recognizes homologs of the aforementioned effectors of *E. amylovora*.

No conclusion about the mode of function of FB_MR5 could be made from a ClustalW pairwise alignment with Prf, RPS2 and RPS5, where 20–22 % identity was calculated between the aa sequences. However, the search for effectors in *E. amylovora* encourages the second hypothesis: No Avr products corresponding to Pto and PBS1 homologs of MR5 were discovered in *E. amylovora* by BLAST search of WGS of *E. amylovora* strains ATCC 49946 and CFBP1430. Instead, homolog of AvrRpt2 of *P. syringae*, $AvrRpt2_{EA}$, was identified in the WGS of *E. amylovora*. Previous studies showed that $AvrRpt2_{EA}$ elicits a hypersensitive response in *A. thaliana* after infiltration with *P. syringae* pv. *tomato* DC3000 expressing $avrRpt2_{EA}$ (Zhao et al. 2006). Also, $avrRpt2_{EA}$- strains had a significantly lower effect on pear fruits (Zhao et al. 2006). In the constitutional transcriptome of unchallenged MR5 (Fahrentrapp et al., unpublished), we also identified a homolog of RIN4, the RPS2 decoy/guardee. RIN4 homologs were also found on chromosome 5 and 10 of GD, and we amplified putative homolog sequences with similar size from cultivars 'Idared' and 'Galaxy'. For these reasons, we suggest that the three proteins FB_MR5, RIN4_MR5 and $AvrRpt2_{EA}$ could act in a similar 'decoy'/'guard' system to RPS2, RIN4 and AvrRpt2.

Conclusion

The FB_MR5 CNL is the second CNL found in apple related to fire blight resistance and the third candidate resistance gene of *Rosaceae* against *E. amylovora* to be expressed and cloned. The FB_MR5 CNL comprises a new LLR-like motif with the consensus LxxLx[IL]xxCxxLxxL putatively involved in protein–protein interaction, like classic LRRs. The results of the comparison to known CNLs and of the search for interacting proteins indicate that the mode of function of FB_MR5 could be in congruence with the decoy/guard model together with RIN4_MR5 and $AvrRpt2_{EA}$. Future tasks are (1) to provide evidence of the functionality of the designated *R* gene in complementation assays and (2) to describe the mode of interaction between *E. amylovora* and *M.* × *robusta* 5, including (3) clarifying the function of the LRR-like motif within FB_MR5 as well as (4) how the resistance was overcome by the two Canadian *E. amylovora* strains. Furthermore, if the functionality of the *R* gene can be confirmed, the flanking markers of *FB_MR5* could be used for highly precise MAS in classical breeding. The markers would also enable the pyramiding of different FB resistance genes and other important apple diseases such as scab and powdery mildew. The isolated *R* gene could be used to transfer the trait in well-established apple cultivars and for engineering cisgene apple plants.

Acknowledgments We acknowledge the Genetic Diversity Center of ETH Zurich, Switzerland, for sequencing, fragment analysis and bioinformatics support, as well as LeRoux P-M and Baumgartner I from Agroscope Changins-Wädenswil (Switzerland) and Malnoy M from FEM/IASMA, Italy, for technical support. For financial funding, we thank the Federal Office for Agriculture FOAG of Switzerland (project: ZUEFOS) as well as the D-A-CH (German–Austrian–Swiss project: 310030L_130811).

References

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

Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22:195–201

Bai J, Pennill LA, Ning J, Lee SW, Ramalingam J, Webb CA, Zhao B, Sun Q, Nelson JC, Leach JE, Hulbert SH (2002) Diversity in nucleotide binding site-leucine-rich repeat genes in cereals. Genome Res 12:1871–1884

Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol 2:28–36

Baldi P, Patocchi A, Zini E, Toller C, Velasco R, Komjanc M (2004) Cloning and linkage mapping of resistance gene homologues in apple. Theor Appl Genet 109:231–239

Barakat A, Han DT, Benslimane A-A, Rode A, Bernardi G (1999) The gene distribution in the genomes of pea, tomato and date palm. FEBS Lett 463:139–142

Bernardi G (2004) The organization of plant genomes. In: Bernardi G (ed) New comprehensive biochemistry. Elsevier, Amsterdam, pp 227–240

Block A, Alfano JR (2011) Plant targets for *Pseudomonas syringae* type III effectors: virulence targets or guarded decoys? Curr Opin Microbiol 14:39–46

Broggini GAL, Le Cam B, Parisi L, Wu C, Zhang HB, Gessler C, Patocchi A (2007) Construction of a contig of BAC clones spanning the region of the apple scab avirulence gene *AvrVg*. Fungal Genet Biol 44:44–51

Broggini GAL, Galli P, Parravicini G, Gianfranceschi L, Gessler C, Patocchi A (2009) *HcrVf* paralogs are present on linkage groups 1 and 6 of *Malus*. Genome 52:129–138

Calenge F, Van der Linden CG, Van de Weg E, Schouten HJ, Van Arkel G, Denance C, Durel CE (2005) Resistance gene analogues identified through the NBS-profiling method map close to major genes and QTL for disease resistance in apple. Theor Appl Genet 110:660–668

Caplan J, Padmanabhan M, Dinesh-Kumar SP (2008) Plant NB-LRR immune receptors: from recognition to transcriptional reprogramming. Cell Host Microbe 3:126–135

Chagne D, Carlisle C, Blond C, Volz R, Whitworth C, Oraguzie N, Crowhurst R, Allan A, Espley R, Hellens R, Gardiner S (2007) Mapping a candidate gene (MdMYB10) for red flesh and foliage colour in apple. BMC Genomics 8:212

DeYoung BJ, Innes RW (2006) Plant NBS-LRR proteins in pathogen sensing and host defense. Nat Immunol 7:1243–1249

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

FAO (2009) Food and agricultural organization of the United Nations Statistical database. FAO, Rome

Flor H (1971) Current status of gene-for-gene concept. Annu Rev Phytopathol 9:275

Frey JE, Frey B, Sauer C, Kellerhals M (2004) Efficient low-cost DNA extraction and multiplex fluorescent PCR method for marker-assisted selection in breeding. Plant Breed 123:554–557

Galli P, Patocchi A, Broggini GAL, Gessler C (2010) The *Rvi15* (*Vr2*) apple scab resistance locus contains three TIR-NBS-LRR genes. Mol Plant Microbe Interact 23:608–617

Gardiner S, Norelli J, de Silva N, Fazio G, Peil A, Malnoy M, Horner M, Bowatte D, Carlisle C, Wiedow C, Wan Y, Bassett C, Baldo A, Celton J-M, Richter K, Aldwinckle H, Bus V (2012) Putative resistance gene markers associated with quantitative trait loci for fire blight resistance in Malus 'Robusta 5' accessions. BMC Genet 13:25

Gessler C (2011) Cisgenic disease resistant apples: a product with benefits for the environment, producer and consumer. Outlooks Pest Manag 22:216–219

Grant M, Godiard L, Straube E, Ashfield T, Lewald J, Sattler A, Innes R, Dangl J (1995) Structure of the *Arabidopsis* RPM1 gene enabling dual specificity disease resistance. Science 269:843–846

Hammond-Kosack KE, Jones JD (1997) Plant disease resistance genes. Annu Rev Plant Physiol Plant Mol Biol 48:575–607

Jones DA, Jones JDG (1997) The role of leucine-rich repeat proteins in plant defences. Adv Bot Res 24:89–167

Juniper BE, Mabberley DJ (2006) The story of the apple. Timber Press, Portland

Kajava AV, Kobe B (2002) Assessment of the ability to model proteins with leucine-rich repeats in light of the latest structural information. Protein Sci 11:1082–1090

Kellerhals M, Székely T, Sauer C, Frey J, Patocchi A (2009) Pyramiding scab resistances in apple breeding. Erwerbs-Obstbau 51:21–28

Khan M, Zhao YF, Korban S (2011) Molecular mechanisms of pathogenesis and resistance to the bacterial pathogen *Erwinia amylovora*, causal agent of fire blight disease in Rosaceae. Plant Mol Biol Rep 30(2):247–260

Kobe B, Deisenhofer J (1994) The leucine-rich repeat: a versatile binding motif. Trends Biochem Sci 19:415–421

Kruijt M, De Kock MJD, De Wit PJGM (2005) Receptor-like proteins involved in plant disease resistance. Mol Plant Pathol 6:85–97

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157:105–132

Liebhard R, Gianfranceschi L, Koller B, Ryder CD, Tarchini R, Van De Weg E, Gessler C (2002) Development and characterisation of 140 new microsatellites in apple (*Malus*×*domestica* Borkh.). Mol Breed 10:217–241

Lupas A (1996) Prediction and analysis of coiled-coil structures. In: Russell FD (ed) Methods Enzymol. Academic, Waltham, pp 513–525

Lupas A, Van Dyke M, Stock J (1991) Predicting coiled coils from protein sequences. Science 252:1162–1164

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

McDonald BA, Linde C (2002) Pathogen population genetics, evolutionary potential, and durable resistance. Annu Rev Phytopathol 40:349–379

McManus PS, Stockwell VO, Sundin GW, Jones AL (2002) Antibiotic use in plant agriculture. Annu Rev Phytopathol 40:443–465

Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW (2003) Genome-wide analysis of NBS-LRR–encoding genes in *Arabidopsis*. Plant Cell 15:809–834

Ngugi HK, Lehman B, Madden LV (2011) Multiple treatment meta-analysis of products evaluated for control of fire blight in the eastern United States. Phytopathology 101:512–522

Norelli JL, Aldwinckle HS (1986) Differential susceptibility of *Malus* spp cultivars Robusta 5, Novole, and Ottawa 523 to *Erwinia amylovora*. Plant Dis 70:1017–1019

Norelli JL, Aldwinckle HS, Beer SV (1988) Virulence of *Erwinia amylovora* strains to *Malus* sp. Novole plants grown in vitro and in the greenhouse. Phytopathology 78:1292–1297

Norelli JL, Farrell RE, Bassett CL, Baldo AM, Lalli DA, Aldwinckle HS, Wisniewski ME (2009) Rapid transcriptional response of apple to fire blight disease revealed by cDNA suppression subtractive hybridization analysis. Tree Genet Genomes 5:27–40

Pagni M, Ioannidis V, Cerutti L, Zahn-Zabal M, Jongeneel CV, Hau J, Martin O, Kuznetsov D, Falquet L (2007) MyHits: improvements to an interactive resource for analyzing protein sequences. Nucleic Acids Res 35:W433–W437

Parravicini G (2010) Candidate genes for fire blight resistance in apple cultivar 'Evereste'. IBZ Plant Pathology, ETH, Zurich, p 149

Parravicini G, Gessler C, Denance C, Lasserre-Zuber P, Vergne E, Brisset MN, Patocchi A, Durel CE, Broggini GAL (2011) Identification of serine/threonine kinase and nucleotide-binding site-leucine-rich repeat (NBS-LRR) genes in the fire blight resistance quantitative trait locus of apple cultivar 'Evereste'. Mol Plant Pathol 12:493–505

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

Paulin JP, Lachaud G, Lespinasse Y (1993) Role of the aggressiveness of strains of *Erwinia amylovora* in the experimental evaluation of susceptibility of apple cultivars to fire blight. Acta Hortic (ISHS) 338:375–376

Pedley KF, Martin GB (2003) Molecular basis of *Pto*-mediated resistance to bacterial speck disease in tomato. Annu Rev Phytopathol 41:215–243

Peil A, Garcia-Libreros T, Richter K, Trognitz FC, Trognitz B, Hanke MV, Flachowsky H (2007) Strong evidence for a fire blight resistance gene of *Malus robusta* located on linkage group 3. Plant Breed 126:470–475

Peil A, Hanke MV, Flachowsky H, Richter K, Garcia-Libreros T, Celton JM, Gardiner S, Horner M, Bus V (2008) Confirmation of the fire blight QTL of *Malus*×*robusta* 5 on linkage group 3. Acta Hortic (ISHS) 793:297–303

Peil A, Flachowsky H, Hanke M-V, Richter K, Rode J (2011) Inoculation of *Malus*×*robusta* 5 progeny with a strain breaking resistance to fire blight reveals a minor QTL on LG5. Acta Hortic (ISHS) 896:357–362

Rairdan GJ, Collier SM, Sacco MA, Baldwin TT, Boettrich T, Moffett P (2008) The coiled-coil and nucleotide binding domains of the potato Rx disease resistance protein function in pathogen recognition and signaling. Plant Cell 20:739–751

Richter K, Fischer C (2000) Stability of fire blight resistance in apple. Acta Hortic (ISHS) 538:267–270

Rozen S, Skaletsky H (1999) Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132:365–386

Salamov AA, Solovyev VV (2000) *Ab initio* gene finding in *Drosophila* genomic DNA. Genome Res 10:516–522

Sanseverino W, Roma G, De Simone M, Faino L, Melito S, Stupka E, Frusciante L, Ercolano MR (2010) PRGdb: a bioinformatics platform for plant resistance gene analysis. Nucleic Acids Res 38:D814–D821

Schouten HJ, Krens FA, Jacobsen E (2006) Cisgenic plants are similar to traditionally bred plants. EMBO Rep 7:750–753

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol 18:233–234

Söding J (2005) Protein homology detection by HMM–HMM comparison. Bioinformatics 21:951–960

Szankowski I, Waidmann S, Degenhardt J, Patocchi A, Paris R, Silfverberg-Dilworth E, Broggini GAL, Gessler C (2009) Highly scab-resistant transgenic apple lines achieved by introgression of *HcrVf2* controlled by different native promoter lengths. Tree Genet Genomes 5:349–358

Tai TH, Dahlbeck D, Clark ET, Gajiwala P, Pasion R, Whalen MC, Stall RE, Staskawicz BJ (1999) Expression of the *Bs2* pepper gene confers resistance to bacterial spot disease in tomato. Proc Natl Acad Sci U S A 96:14153–14158

Taylor WR (1986) The classification of amino acid conservation. J Theor Biol 119:205–218

Van der Biezen EA, Jones JD (1998) Plant disease-resistance proteins and the gene-for-gene concept. Trends Biochem Sci 23:454–456

Van der Hoorn RAL, Kamoun S (2008) From guard to decoy: a new model for perception of plant pathogen effectors. Plant Cell 20:2009–2017

Van der Zwet T (2006) Present worldwide distribution of fire blight and closely related diseases. Acta Hortic (ISHS) 704:35–36

Van Ooijen JW, Voorrips RE (2001) JoinMap 3.0, software for the calculation of genetic linkage maps. Plant Research International, Wageningen

Van Ooijen G, Mayr G, Kasiem MMA, Albrecht M, Cornelissen BJC, Takken FLW (2008) Structure–function analysis of the NB-ARC domain of plant disease resistance proteins. J Exp Bot 59:1383–1397

Vanblaere T, Szankowski I, Schaart J, Schouten H, Flachowsky H, Broggini GAL, Gessler C (2011) The development of a cisgenic apple plant. J Biotechnol 154:304–311

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D, Salvi S, Pindo M, Baldi P, Castelletti S, Cavaiuolo M, Coppola G, Costa F, Cova V, Dal Ri A, Goremykin V, Komjanc M, Longhi S, Magnago P, Malacarne G, Malnoy M, Micheletti D, Moretto M, Perazzolli M, Si-Ammour A, Vezzulli S, Zini E, Eldredge G, Fitzgerald LM, Gutin N, Lanchbury J, Macalma T, Mitchell JT, Reid J, Wardell B, Kodira C, Chen Z, Desany B, Niazi F, Palmer M, Koepke T, Jiwan D, Schaeffer S, Krishnan V, Wu C, Chu VT, King ST, Vick J, Tao Q, Mraz A, Stormo A, Stormo K, Bogden R, Ederle D, Stella A, Vecchietti A, Kater MM, Masiero S, Lasserre P, Lespinasse Y, Allan AC, Bus V, Chagne D, Crowhurst RN, Gleave AP, Lavezzo E, Fawcett JA, Proost S, Rouze P, Sterck L, Toppo S, Lazzari B, Hellens RP, Durel C-E, Gutin A, Bumgarner RE, Gardiner SE, Skolnick M, Egholm M, Van de Peer Y, Salamini F, Viola R (2010) The genome of the domesticated apple (*Malus*×*domestica* Borkh.). Nat Genet 42:833–839

Walker JE, Saraste M, Runswick MJ, Gay NJ (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J 1:945–951

Yoshimura S, Yamanouchi U, Katayose Y, Toki S, Wang Z-X, Kono I, Kurata N, Yano M, Iwata N, Sasaki T (1998) Expression of Xa1, a bacterial blight-resistance gene in rice, is induced by bacterial inoculation. Proc Natl Acad Sci U S A 95:1663–1668

Zhao YF, He SY, Sundin GW (2006) The *Erwinia amylovora avrRpt2 (EA)* gene contributes to virulence on pear and *AvrRpt2(EA)* is recognized by *Arabidopsis RPS2* when expressed in *Pseudomonas syringae*. Mol Plant Microbe Interact 19:644–654