Theor Appl Genet (2013) 126:359–367
DOI 10.1007/s00122-012-1984-6

ORIGINAL PAPER

Positional cloning of a candidate gene for resistance to the sunflower downy mildew, *Plasmopara halstedii* race 300

Jérôme Franchel · Mohamed Fouad Bouzidi · Gisèle Bronner · Felicity Vear · Paul Nicolas · Said Mouzeyar

Received: 12 March 2012 / Accepted: 15 September 2012 / Published online: 29 September 2012
© Springer-Verlag Berlin Heidelberg 2012

Abstract The resistance of sunflower to *Plasmopara halstedii* is conferred by major resistance genes denoted *Pl*. Previous genetic studies indicated that the majority of these genes are clustered on linkage groups 8 and 13. The *Pl6* locus is one of the main clusters to have been identified, and confers resistance to several *P. halstedii* races. In this study, a map-based cloning strategy was implemented using a large segregating F2 population to establish a fine physical map of this cluster. A marker derived from a bacterial artificial chromosome (BAC) clone was found to be very tightly linked to the gene conferring resistance to race 300, and the corresponding BAC clone was sequenced and annotated. It contains several putative genes including three toll-interleukin receptor-nucleotide binding site-leucine rich repeats (TIR-NBS-LRR) genes. However, only one TIR-NBS-LRR appeared to be expressed, and thus constitutes a candidate gene for resistance to *P. halstedii* race 300.

Communicated by A. Bervillé.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-012-1984-6) contains supplementary material, which is available to authorized users.

J. Franchel · M. F. Bouzidi · F. Vear · P. Nicolas · S. Mouzeyar (✉)
Clermont Université, Université Blaise Pascal, UMR INRA-UBP 1095 GDEC, BP 10448, 63000 Clermont-Ferrand, France
e-mail: said.mouzeyar@univ-bpclermont.fr

J. Franchel
Clermont Université, Université Blaise Pascal, UMR 547 PIAF, BP 10448, 63000 Clermont-Ferrand, France

J. Franchel
INRA, UMR 547 PIAF, 63100 Clermont-Ferrand, France

G. Bronner
Clermont Université, Université Blaise Pascal, Laboratoire "Microorganismes: Génome et Environnement", BP 10448, 63000 Clermont-Ferrand, France

G. Bronner
CNRS, UMR 6023, LMGE, 63171 Aubiere, France

Introduction

Sunflower downy mildew is caused by the obligate parasitic oomycete, *Plasmopara halstedii*. The resistance of sunflower to this parasite is controlled by single dominant genes denoted *Pl* (Vranceanu and Stoenescu 1971), each being effective against one of the many races that have been described worldwide. Genetic studies have shown that the interaction between sunflower and downy mildew races probably complies with the gene for gene concept as initially described by Flor (1947). *Pl1* was the first *Pl* gene to be localised on linkage group 8 of the sunflower map (Mouzeyar et al. 1995). Other *Pl* genes were later mapped and proved to be clustered (Vear et al. 1997; Bachlava et al. 2011).

Sequence comparisons of the plant resistance genes that have been cloned show that the majority of these genes belong to a highly represented gene family containing two domains; a nucleotide binding site (NBS) and a stretch of leucine rich repeats (LRR) (Dangl and Jones 2001). Some of the NBS-LRR proteins contain either a coiled-coil (CC) domain or a toll-interleukin receptor (TIR) domain at their N-termini (McHale et al. 2006). Degenerate primers targeting conserved regions within these sequences have been used to clone resistance gene candidates (RGC), which mapped regions containing clusters containing resistance genes to downy mildew (Gentzbittel et al. 1998; Gedil et al.

2001a; Radwan et al. 2003). Subsequent molecular analysis showed at least two major loci conferring resistance against *P. halstedii*. The *Pl6* locus seems to contain mainly TIR-NBS-LRR sequences (Bouzidi et al. 2002), whereas the *Pl5/Pl8* locus seems to contain mainly CC-NBS-LRR sequences (Radwan et al. 2004). In a comprehensive work, Radwan et al. (2008) mapped unique NBS-LRR sequences to 167 loci throughout the 17 sunflower linkage groups, thus giving a more complete picture of the genomic distribution of NBS-LRR encoding genes in this plant. Interestingly, most of these loci mapped to linkage groups 8 and 13, thus confirming previous findings (Gentzbittel et al. 1998; Bert et al. 2001; Gedil et al. 2001b; Bouzidi et al. 2002; Radwan et al. 2003, 2004; Slabaugh et al. 2003). Recently, Bachlava et al. (2011) identified several bacterial artificial chromosome (BAC) clones proximal to *Pl8* and *Pl14* in linkage group 13 and showed that they contain additional RGC with non-TIR-like motifs, although they are phylogenetically distant from the two non-TIR-NBS-LRR classes of resistance genes.

The cloning of functional *Pl* genes would help us gain insights into the molecular basis of downy mildew resistance in sunflower. This study thus focused on the *Pl6* locus conferring resistance to several downy mildew races, using 3,072 F2 individuals and a physical map. In addition, using a map-based strategy, we exploited the molecular markers described in a previous work (Bouzidi et al. 2002) to clone a TIR-NBS-LRR candidate gene for resistance to race 300 of *P. halstedii*. A 111 Kb BAC clone was completely sequenced and annotated. It contains three TIR-NBS-LRR sequences of which only one is expressed; it is thus a putative functional RGC, and therefore, a candidate gene for resistance to *P. halstedii* race 300.

Materials and methods

Plant genotypes and disease assessment

The YDQ sunflower line contains the *Pl6* locus which confers resistance against races 100, 300, 700, 703, 710 of *P. halstedii* (Bouzidi et al. 2002). The H52 sunflower line (from ARS, South Africa) is susceptible to all known races of *P. halstedii*. A total of 3,072 F2 individuals were produced from a cross between these two lines. F2 seedlings with recombinations in the *Pl6* locus were identified with flanking codominant markers and were selfed to obtain the F3 families.

The F3 families were tested for resistance to *P. halstedii* races 300, 703 and 710. For each F3 progeny, at least 20 seedlings were infected and grown, as described by Mouzeyar et al. (1993). For the few F3 progenies which were difficult to define, F4 progenies were tested alike, thus making it possible to classify the F2 plants as homozygous resistant, homozygous susceptible or heterozygous (Vear et al. 1997).

Mapping of the *Pl6* locus

Sequence tagged site (STS) markers within the *Pl6* locus, described by Bouzidi et al. (2002), were used to develop codominant PCR markers flanking the *P. halstedii* resistance genes. Two primer pairs (NBS8 and NBS1–3) were used to screen 3072 F2 plants, to identify recombinant individuals within the *Pl6* locus. The NBS8 pair is made up of the NBS8For (GAACTTTGTGGATCATCTCTACGA) and NBS8Rev (CTATCAAATTTGCTCATTTGACTT) primers and the NBS1–3 pair contains the NBS8For and NBS1–3Rev (TCTATCTCACTCCCTACCACA) primers.

A sunflower BAC library (Bouzidi et al. 2006) made up from isolated nuclei of the resistant parent (YDQ) was screened using two TIR-NBS-LRR probes (Ha-NBS2 and Ha-NBS8) that are linked to the *Pl6* locus (Bouzidi et al. 2002). The BAC library contained 147,456 clones with an estimated average insert size of 118 Kb and was arrayed onto 16 nylon filters. The probes were labelled radioactively using the random priming method (Amersham Megaprime kit), and hybridized to the filters under stringent conditions using hybridization methods (Sambrook et al. 1989).

DNA from the 245 BAC clones was digested with *Hin*dIII that was end-labelled with [32P]dATP. Digested DNA was then subjected to 3.5 % polyacrylamide DNA sequencing gel electrophoresis. The gels were dried and autoradiographed, and the fingerprints were scanned into images and edited using Image 4.0 software (Soderlund et al. 1997). Version 4.7 of the FPC program (Soderlund et al. 2000) was then used to assemble the BAC contigs using the default parameters.

The BAC ends were then sequenced (GENOME Express, Grenoble). About 240 primer pairs (Eurogentec) were designed using the BAC clone ends and tested on parental (YDQ, H52) and bulk DNA (according to their phenotype in downy mildew tests with race 300). PCR-based markers segregating between the susceptible and resistant parents were then mapped using MAPMAKER 3.0 with a minimum LOD score of 3.0 and a maximum recombination fraction of 0.45 (Lander et al. 1987).

DNA and RNA extraction

Young leaf tissue from F2 plants was collected and freeze dried. DNA was isolated using the Metabisulfite method (Peterson et al. 1997) adapted to a large number of plants. Total cellular DNA was then isolated from individual seedlings. Briefly, freeze-dried leaves were ground with a mechanical mill, dispersed in 600 µL of lysis buffer

(100 mM Tris/0.5 M NaCl/50 mM EDTA/20 mM sodium metabisulfite), mixed gently by inversion and then incubated at 95 °C for 45 min. The crude extract was centrifuged at 4,000×*g* at room temperature for 20 min. Two hundred microliters of Isopropanol containing 1 M ammonium acetate was added to 200 µL of supernatant, and the solution was mixed by inversion and centrifuged at 4,000×*g* at room temperature for 20 min. The precipitated DNA was washed with 600 µL ethanol and centrifuged at 3,000×*g* at room temperature for 20 min. The supernatant was removed and the DNA was dried and suspended in distilled water. DNA from the leaves of F2 recombinants was isolated using the CTAB method, as described by Saghai Maroof et al. (1984).

The DNA inserts from BAC clones were isolated using the alkaline lysis method (Sambrook et al. 1989).

For RNA extraction, healthy or infected plants were used by the whole seed immersion method. Young leaf tissue from the control or infected resistant and susceptible parents was collected and freeze-dried. About 250 mg (dry weight) of freeze-dried tissue were ground and dispersed in 900 µL CTAB buffer (0.1 M Tris/1.4 M NaCl/20 mM EDTA pH 8/2 % cetylmethylammonium bromide/3 % β-mercaptoethanol), mixed by vortexing and incubated at 65 °C for 5 min before 900 µL of phenol acid was added and vortexed, and then centrifuged at 19,000×*g* at 4 °C for 10 min. Isopropanol (v/v) was then added to the supernatant, and the solution was mixed by vortexing, incubated at −80 °C for 2 h and centrifuged at 19,000×*g* at 4 °C for 20 min. The precipitated nucleic acids were washed with 600 µL ethanol (100 %) and centrifuged at 10,000×*g* at 4 °C for 5 min. The supernatant was removed and nucleic acids were dried and re-suspended in distilled RNAse free water. Lithium chloride (3 M final) was added, then the solution was vortexed, incubated overnight at 4 °C and centrifuged at 19,000×*g* at 4 °C for 45 min. RNA were washed with 600 µL ethanol (70 %) and centrifuged at 15,000×*g* at 4 °C for 5 min. The supernatant was removed and the RNA dried and re-suspended in distilled RNAse free water. The RNA solution was incubated at 37 °C for 30 min in the presence of 1 U DNAse buffer, 10 U DNAse (Roche) and 20 units RNAsine (Promega). The solution was washed with phenol–chloroform and then with chloroform and centrifuged at 10,000×*g* at 4 °C for 10 min. The RNA were then precipitated, dried and re-suspended in distilled RNAse free water.

Analysis of expression

Two oligonucleotide primer pairs (TIR3F1R2: 5′TGTCTTGAAAGATTGATTTAGTGT3′ and 5′CCCTCTATCCTTTTCAGTTTTCTG3′) and (TIR3F2R2: 5′TTGATCTCAGTTGGTGCAAGGA3′ and 5′CCCTCTATCCTTTTCAGTTTTCTG3′) were used. They were designed in the LRR domain using the 3 TIR-NBS-LRR sequences in the BAC clone sequenced (this study). Ha-EF-1 (5′AGGCGAGGTATGATGAAATTGTCA3′ and 5′GTCTCTTGGGCTCATTGATTTGGT3′) was used as the reference genes.

Reverse transcription was performed with the Invitrogen SuperScript™ (11904-018) kit. Quantitative RT-PCR amplification was carried out with 2 µL of 1/10 diluted RT product. In the presence of 0.13 mM of each dNTP, 0.3 µL of *Taq* DNA Polymerase (Advantage 2, Clontech, France), 1× *Taq* Polymerase Buffer, 0.17 nM of each primer, 0.5 µL of SYBR Green I (Sigma) diluted 1/1,000 in 20 % DMSO and 1 µL of Fluorescein (BioRad) diluted 1/1,000 in water. qRT-PCR was carried out in a Biorad i-cycler thermocycler under the following conditions: initial denaturation at 94 °C for 2 min, 50 cycles of 94 °C for 15 s, 64 °C for 15 s and 72 °C for 20 s. Amplicons were then checked for specificity by sequencing and by subjecting them to melt curve analysis, using a 0.5 °C incremental increase in temperature for 10 s, starting at 55 °C. The relative expression of the TIR-NBS-LRR sequence was quantified using the $2^{-\Delta Ct}$ method, according to the method described by Schmittgen and Livak (2008) with the sunflower elongation factor Ha-EF-1α as a reference gene (Radwan et al. 2005).

DNA sequencing and analysis

A shotgun library of the BAC clone was obtained by mechanical cleavage (Hydroshear, GeneMachines). Fragments ranging in size from 2.5 to 5 Kb were ligated to PUC18 and transformed into *E. coli* using Electro Ten Blue (Stratagene). DNA sequencing was carried out using the ET Terminator sequencing kit (Amersham Bioscience). The quality of sequences was controlled using Phred (Ewing and Green 1998; Ewing et al. 1998), Cross-match (Ewing and Green 1998; Ewing et al. 1998) and BLASTN (Altschul et al. 1997). Assembly of the sequence was achieved using Phrap (Ewing and Green 1998; Ewing et al. 1998). The complete sequence is available in Genbank under accession number HM543461.

Gene structures were predicted on the basis of Genscan (Burge and Karlin 1998) using *A. thaliana* as a matrix. The predicted coding DNA sequence (CDS) and mRNA were then compared to the NR protein database (NCBI), using BLASTX with an *E*-value threshold of $E < 10^{-5}$. If predictions did not align with complete subject proteins in the NR database, the whole genomic segments corresponding to the predicted transcript were blasted against the NR protein database using the same parameters. Intron/exon boundaries and exon reading frames were verified and fine-tuned manually based on a comparison of Genscan predictions and BLASTs hits using the Artemis viewer (Mural 2000). The

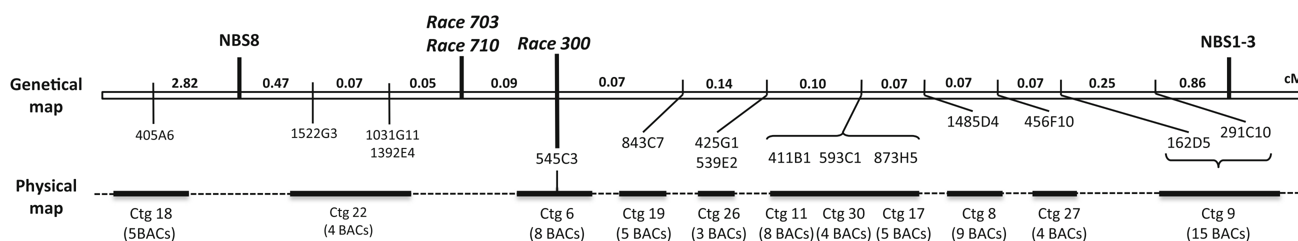

Fig. 1 Detailed map of the sunflower *Pl6* locus containing resistance genes to *P. halstedii* races 300, 703 and 710. The marker names correspond to the BAC clones from which they were derived. CTG stands for contig and the number of BAC clones in each contig is indicated in parentheses. Each contig is symbolised by a *thick black line* and *dots* indicate gaps between the contigs. Genetic distances are given in cM and are based on recombination data from 3,072 F2 individuals

clone sequence was also analysed with RepeatMasker (Chen 2004) to detect repeated elements as well as pseudogenes.

The clone sequence was compared (tblastn) against itself and a dot plot in order to detect regions of high similarity. Finally, the three putative TIR-NBS-LRR genes were translated into proteins and aligned with muscle (Edgar 2004).

Results

Genetic mapping and progeny testing of resistance to downy mildew

Thirteen RGA markers linked to the *Pl6* locus that had previously been identified and sequenced (Bouzidi et al. 2002) were used to develop two codominant STS markers (NBS8 and NBS1-3) flanking the *Pl6* locus, in order to screen 3,072 F2 individuals and identify recombinants between these markers (Fig. 1). A total of 160 F2 individuals were found to be recombinant in the 6.3 cM interval between NBS8 and NBS1-3 that contained the *Pl6* locus.

In order to locate the *Pl6* locus precisely on the molecular map, the 160 F2 recombinant individuals were selfed. One hundred and fifty-one F3 families corresponding to 151 F2 individuals were obtained and tested for their resistance to *P. hasltedii*. In cases where progeny tests were unable to clearly distinguish between susceptible and resistant, the corresponding F4 families were tested. Each family was tested with races 300, 703 and 710 of *P. halstedii.* Segregation in the *Pl6* locus between the gene conferring resistance to race 300 and those conferring resistance to races 703 and 710 was found (Fig. 1), thus confirming the findings of Bouzidi et al. (2002).

Generation of additional markers for fine mapping of the *Pl6* locus

Screening at high stringency of the library of 147,456 BAC clones from the resistant parent, YDQ identified 245 positive clones potentially linked to the *Pl6* locus. These 245 clones were assembled in 36 contigs and 17 singletons by *Hin*dIII digestion and FPC analysis (Soderlund et al. 2000) (Fig. 1).

Both ends of all the BAC clones were sequenced and analysed using BLASTX and placed in four classes, as a function of the hits they matched: non-coding sequence, putative gene, transposable element, resistance-related gene. This BAC end analysis revealed that the DNA in this resistance locus consists predominantly of non-coding sequences or putative genes (64 %). Most of the sequences corresponded to transposable elements which are supposed to act in reshaping the genome, especially at resistance loci (Meyers et al. 2003). All the resistance-related genes analysed were of the TIR-NBS-LRR class, which is one of the three NBS/LRR subdivisions described in plants alongside the CC-NBS-LRR and the NBS-LRR gene classes.

Using the sequences of the BAC ends, 240 specific primers pairs were designed and STS were developed in order to increase the density of markers within the *Pl6* locus. Of the 239 amplicons tested (Supplementary Table S1), 25 new STS markers could be mapped on the YDQ × H52 cross (Supplementary Table S2). One STS marker specific to BAC Contig #6 was found to be absolutely linked (i.e., no recombination) to resistance to race 300 using the 3,072 F2 individuals. This contig could therefore be considered as a good candidate for harbouring the resistance gene to *P. halstedii* race 300.

Analysis of the BAC contig #6 tightly linked to resistance 300

As southern analyses of the BAC contig #6 displayed at least two putative resistance genes, the largest and more representative BAC clone 242G9 was sequenced using the shotgun strategy resulting in a 111,896 bp sequence (Fig. 2). Annotation of the sequence showed that it contained eight genes or pseudogenes corresponding roughly to one-third of the sequence, five putative transposable elements (TE) corresponding to another third of the

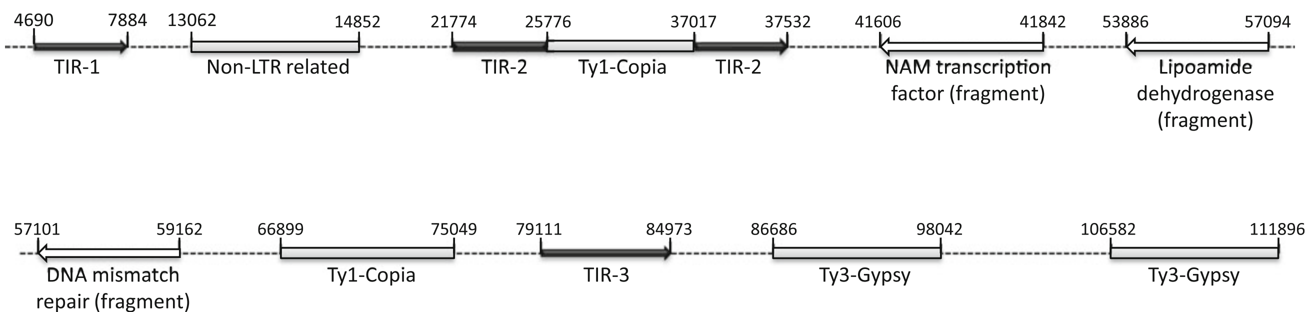

Fig. 2 Annotation of BAC clone 242G9 (accession number HM543461). *Arrows* indicate the relative orientation of the predicted CDS and *numbers* indicate their position

sequence and the remaining third corresponded to non-coding intergene sequences (Fig. 2).

Isolated functional domains were identified, although no similarity with any coding sequence was found around them. These functional domains, well supported through the BLAST analysis, were scattered along the sequence and corresponded to various functions: a NAM transcription factor (PFAM02365), a transket_pyr domain from pyruvate dehydrogenase E1 beta subunit (mitochondrial), a PRK06416 domain from a lipoamide dehydrogenase 1, a ABC_M-SH2_euk domain fragment from protein DNA mismatch repair MSH2 (CD03285), a ribosomal protein-like and a RNA-directed polymerase RT non-LTR related (PTHR 19446) (Fig. 2).

RepeatMasker analysis enabled the identification of ten segments similar to putative transposable elements from *A. thaliana*. These predictions were congruent with the BLAST annotations. Two Ty1-copia (TE_1 and TE_2) and two Ty3-gypsy (TE_3 and TE_4) retroelements were predicted. TE_1 possessed two domains of copia-type retroelement rve and RVT_2 and two potential LTR regions of 662 bp with 99 % identity, suggesting the recent insertion of the element. This TE could span up to 11 Kb, which is within the range of a copia element length. Ty1-copia sequences range from 3,925 bp for PRD1 in pea to 12,088 bp in the BAR1 of barley (Peterson-Burch and Voytas 2002) and more than 16 kb in *Helianthus annuus* (Buti et al. 2011). TE_3 is a 11,465 bp complete Ty3-gypsy like protein, since two 470 bp segments with 100 % identity and potentially corresponding to LTR domains were detected both upstream and downstream of the gypsy gag, RVP_2, RT and rve domains. A 98 % identity duplicated segment of 2.8 Kb was identified at both 3,077 bp downstream of the 5′ putative LTR of the Ty3-gypsy and 189 bp after the 3′ end of this TE. TE_4 was truncated in its 3′ end towards the end of the BAC clone sequence (Fig. 2 and supplementary file Figure S1).

Three predicted TIR-NBS-LRR genes (TIR_1, TIR_2 and TIR_3) were similar to RS6-8 and PU3 resistant-like proteins from *Helianthus annuus* (AAN73007.1, AAL07535.1). However, TIR_1 lacked the typical TIR domain, as well as a partially deleted LRR region and is interrupted by a 70-mer TA microsatellite. So it was assumed that TIR_1 was non-functional and corresponded to a pseudogene. TIR_2 corresponded to an almost complete sequence of a TIR-NBS-LRR gene, although it was disrupted by the insertion of TE_1. Indeed, the TBLASTX of the BAC clone against itself exhibited a DNA fragment downstream of the 3′ end of TE_1, which was very similar to the 3′ end of both TIR_1 and TIR_3. The gene referred to as TIR_3 seemed to be complete as it contained a complete TIR domain, an NB-ARC domain and an LRR region with at least nine LRR domains, as identified by InterProscan. TIR_3 thus corresponded to a potentially functional TIR-NBS-LRR resistant gene (Fig. 2).

The multiple alignments of TIR_1, TIR_2 and TIR_3 allowed us to identify conserved segments and motifs. Most of the mutations seemed to be deleted (Fig. 3). The TIR_2 gene seemed to be entire when compared with TIR_3; in other words, TIR_2 exhibited the three conserved domains (TIR, NB-ARC and 9 LRR) except for a deletion and the presence of a stop codon in the reading frame of the gene beyond the TE_1 insertion. Nevertheless, a 378 bp sequence identified as 3′ UTR, which was very similar in TIR_1 and TIR_3, was not recovered in TIR_2. TIR_1 was characterised by the presence of an internal fragment of the TIR domain, a conserved segment from the 3′ end of the TIR domain to the first quarter of the NB-ARC domain, and an LRR domain exhibiting two deletions when compared with the TIR_2 and TIR_3 LRR regions. The 3′ end of TIR_1 was conserved with close similarity to the two other R genes including its 3′ UTR.

TIR_3 expression

Expression of the CDS identified as TIR_3 was monitored by qRT-PCR using either healthy or infected sunflower plants. The kinetic showed that TIR_3 was expressed constitutively in YDQ plants. TIR_3 mRNAs were accumulated at a low level and were detectable in both the plant infected with *P. halstedii* and in healthy plants. No

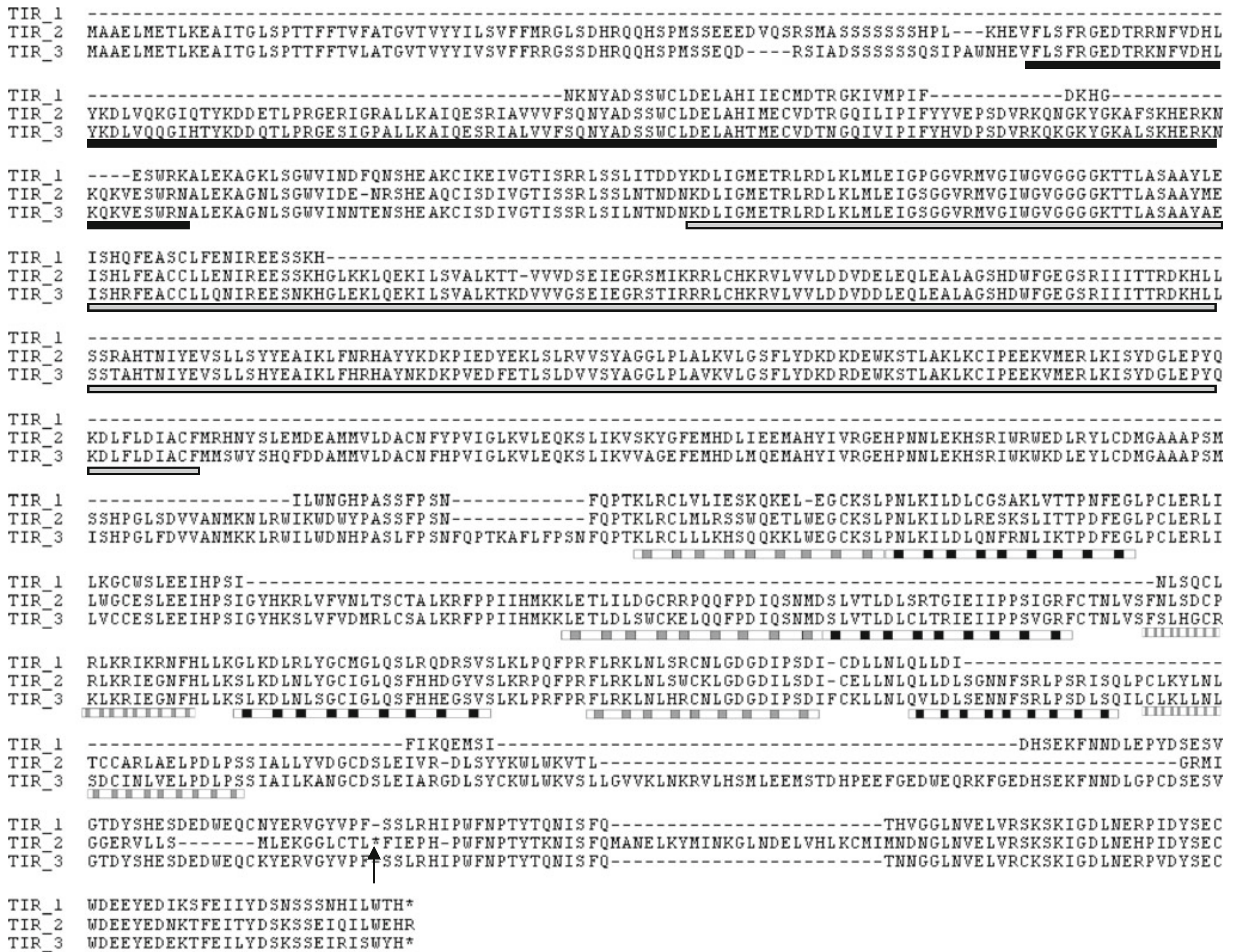

Fig. 3 Alignment of the amino acid sequences of the three TIR-NBS-LRR genes predicted in BAC clone 242G9 (accession number HM543461). *Dots* indicate deletions. A *black rectangle* indicates the toll/interleukin-1 receptor homology (TIR) domain. A *grey rectangle* indicates the NB-ARC domain and *black white* or *grey white* rectanglesindicate the LRR motifs. *Arrow* indicates the premature stop codon in the predicted TIR_2 protein

induction of TIR_3 expression was observed within the 15 days, following infection by *P. halstedii* race 300. As a control to verify the qRT-PCR findings at different points of the kinetic, an equal signal was observed for the Elongation factor 1α gene in all RNA samples (Fig. 4). However, all attempts to detect mRNA corresponding to TIR_1 and TIR_2 failed (data not shown).

Discussion

We present in this study the cloning of a sunflower candidate gene for resistance to *P. halstedii* race 300, the causal agent of downy mildew. For this purpose, we exploited the results obtained by Bouzidi et al. (2002), who described several dominant STS markers within the *Pl6* locus conferring resistance to several races of *P. halstedii*. Some of these markers were used to develop codominant markers flanking the disease resistance gene cluster. These codominant markers were then used to screen 3072 F2 individuals segregating for resistance to downy mildew, and 151 recombinants were found. In order to determine additional markers more closely linked to the *Pl* genes, we used PCR markers derived from the BAC end sequences of clones within the *Pl6* locus. One out of 25 polymorphic markers was found to be completely linked to resistance to *P. halstedii* race 300 in the segregating population. One BAC clone was then selected, shotgun sequenced and annotated, and one TIR-NBS-LRR candidate gene was identified.

Physical map of the *Pl6* locus

Screening of a sunflower BAC library constructed using DNA from the resistant line YDQ with two STS markers

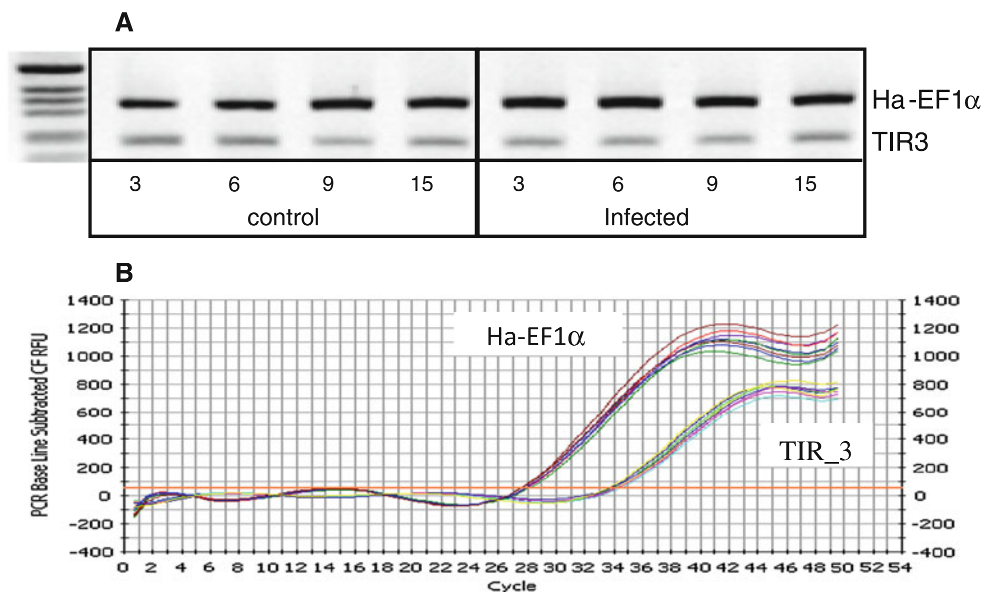

Fig. 4 **a** Semi-quantitative RT-PCR *diagram* showing the amplification of cDNA from the constitutive control Ha-EF1α and the TIR-NBS-LRR gene TIR_3. *Numbers* indicate days after infection. **b** Real-time RT-PCR showing the kinetic of cDNA amplification from the constitutive control Ha-EF1α and the TIR-NBS-LRR gene

was able to identify 245 strongly positive clones. The probes used to screen the library were the most distinct, according to their sequence homology (Bouzidi et al. 2002). It can, therefore, be assumed that the majority of potentially positive BAC clones were recovered. Furthermore, given the large representation of this class of genes in both sunflower and other plants (Bouzidi et al. 2002; Gedil et al. 2001b; Meyers et al. 2003), the hybridization and washing conditions were extremely stringent in order to prevent any cross-hybridization with BAC clones containing NBS-LRR sequences originating from chromosomal regions other than the *Pl6* locus. When the 245 BAC clones were subjected to *Hind*III digestion and arranged in contigs, 36 contigs were identified as containing 2–15 clones. Given that the average length of BAC clones is approximately 120 Kb, the total length of the *Pl6* locus is probably larger than 4 Mb (120 × 36), which is similar to the *Dm3* locus in lettuce (Meyers et al. 1998a). However, the gaps that remain between the contigs still need to be filled so as to enable a more accurate estimation of the size of this locus. Plant resistance genes belonging to the NBS-LRR class are often clustered within large chromosomal regions (Young 2000; Radwan et al. 2008; Bachlava et al. 2011). This genomic organisation is thought to favour rearrangements between sequences within these clusters, giving rise to new specificities or leading to the extinction of some others (Meyers et al. 1998b). This large physical breadth of the *Pl6* locus was also expected, because several sunflower resistance genes have been mapped to this locus (Bouzidi et al. 2002; Gedil et al. 2001a); and these genes confer resistance not only to downy mildew, but also to the rust caused by *Puccinia helianthi* (Lawson et al. 1998).

The *Pl6* locus contains several NBS-LRR sequences

Screening of the *Pl6*-derived sunflower BAC library with a highly stringent TIR-NBS-LRR probe identified 245 positive clones. These clones were then subjected to *Hind*III digestion and probed with different NBS-LRR probes. Each BAC clone contained between two and fives copies of TIR-NBS-LRR-related sequences (data not shown). This result indicates that this locus may contain several copies belonging to this class of resistance genes, as reported previously by Bouzidi et al. (2002). In addition, among the BAC end sequences of these 245 clones (i.e., 490 sequences), 23 end sequences displayed strong homology with TIR-NBS-LRR sequences in the databases. The clustering of resistance genes has been described extensively in different plant species, and the sunflower *Pl6* locus is probably as complex as *Dm3* in lettuce, which contains 24 NBS-LRR genes (Meyers et al. 1998a) even though it does not contain similar sequences.

Identification of genes

Even though gaps still remain between the BAC contigs, and the physical map described in this paper is of the first generation, there is a strong evidence that the BAC clone thus identified contains a good candidate for resistance to *P. halstedii* race 300. Complete annotation of contiguous 111,896 bp revealed 13 potential coding sequences. Four CDS shared homology with LTR-containing retrotransposons, two corresponded to Ty1-copia (Pseudoviridae) and two to Ty3-gypsy (Metaviridae). One Ty1-copia (TE1) is postulated to be a recent insertion because the gene interrupted does not seem to have evolved in a degenerative

manner. The LTR of this element shared 99 % identity and may have recently become active.

Of the two Ty3-gypsy ORFs detected, one is likely to be entire (11.5 kb) and presents a 470 bp LTR on both sides. It is within the range of previously described retrotransposons in sunflower which are between 7 and 14 kb (Buti et al. 2011). Similarly, a sunflower BAC clone found in databanks (FJ269356) possesses 3 Ty3-gypsy ranging from 9 to 15 kb. Interestingly, Bachlava et al. (2011) reported that 80, 60 and 30 % of three BAC clones in the *Pl8* and *Pl14* resistance genes clustered on linkage group 13 encoded retroelements.

Mutations due to the movement of transposable elements are common in different plant genera (Johns et al. 1985; Grandbastien et al. 1989; Vignols et al. 1995), and Ty3-gypsy is known to induce mutations in sunflower (Tang et al. 2006). However, this is the first time that a Ty1-copia has been shown to interrupt a resistance gene in sunflower, hence creating a non-functional gene. Ty3-gypsy and Ty1-copia had previously been found in sunflower (Santini et al. 2002; Buti et al. 2011), and it is believed that they played a role in *Helianthus* genome evolution by enlarging the genome of *Helianthus* homoploid hybrid species (Baack et al. 2005) and in interrupting genes. The same pattern of gene and retroelement diversity, and of coding versus non-coding sequence ratios, was found in another sunflower BAC clone (accession number FJ269356).

Of the 13 CDSs identified in the BAC clone, three share homology with TIR-NBS-LRR plant resistance genes, and all of them may be candidate genes for the resistance of cultivated *H. annuus* to *P. halstedii* race 300. However, TIR_1 is not only truncated but also interrupted by a microsatellite, which means that this sequence is not likely to be functional. This region of the BAC clone contains no functional CDSs with TIR_1, while it contains a putative non-LTR retroelement for which only a reverse transcriptase domain could be detected.

TIR_2 is interrupted due to the insertion of a 7.3 kb-length sequence corresponding to the TE1 Ty1-copia element. This insertion alters the predicted CDS by creating a premature stop codon that would lead to a potentially truncated protein. In addition, RT-PCR with specific primers targeting these sequences failed to amplify any detectable product in healthy or infected plants, suggesting that these sequences are probably not functional.

The remaining sequence TIR_3 corresponds to a CDS potentially coding for a 1,099 aminoacids protein. This sequence presents all the characteristics of previously described resistance genes (Meyers et al. 1999) that are the TIR, NBS-ARC and 9 LRR domains. This sequence is constitutively expressed in leaves, and thus constitutes the best candidate gene for resistance to *P. halstedii* race 300. Nevertheless, transgenic susceptible sunflower lines containing this candidate gene need to be produced to verify this hypothesis.

Several sunflower loci containing NBS-LRR-related sequences have been described recently, and the disease resistance locus located at the end of one arm of linkage group LG8 is the largest (Radwan et al. 2008; Bachlava et al. 2011). This locus contains 54 different NBS-LRR sequences out of the 167 that were mapped during this study (Radwan et al. 2008). Interestingly, some of these sequences share a high degree of similarity with TIR_3 sequences, but none of them was found to be identical to this candidate gene. Thus, the work by Radwan et al. (2008) and the results described in the present paper indicate that the *Pl6* locus on linkage group 8 is very large and may contain dozens of functional resistance genes against various pests. Similarly, our work has supplied additional genomic resources that will be useful for the selection of specific resistance genes during sunflower breeding programmes.

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped blast and psi-blast: a new generation of protein database search programs. Nucleic Acids Res 25(17):3389–3402

Baack EJ, Whitney KD, Rieseberg LH (2005) Hybridization and genome size evolution: timing and magnitude of nuclear DNA content increases in *Helianthus* homoploid hybrid species. New Phytol 167(2):623–630

Bachlava E, Radwan OE, Abratti G, Tang S, Gao W, Heesacker AF, Bazzalo ME, Zambelli A, Leon AJ, Knapp SJ (2011) Downy mildew (*Pl8* and *Pl14*) and rust (RAdv) resistance genes reside in close proximity to tandemly duplicated clusters of non-TIR-like NBS-LRR-encoding genes on sunflower chromosomes 1 and 13. Theor Appl Genet 122:1211–1221

Bert PF, Tourvieille De Labrouhe D, Philippon J, Mouzeyar S, Jouan I, Nicolas P, Vear F (2001) Identification of a second linkage group carrying genes controlling resistance to downy mildew (*Plasmopara halstedii*) in sunflower (*Helianthus annuus* L.). Theor Appl Genet 103(6):992–997

Bouzidi MF, Badaoui S, Cambon F, Vear F, De Labrouhe DT, Nicolas P, Mouzeyar S (2002) Molecular analysis of a major locus for resistance to downy mildew in sunflower with specific PCR-based markers. Theor Appl Genet 104(4):592–600

Bouzidi MF, Franchel J, Tao Q, Stormo K, Mraz A, Nicolas P, Mouzeyar S (2006) A sunflower bac library suitable for pcr screening and physical mapping of targeted genomic regions. Theor Appl Genet 113(1):81–89

Burge CB, Karlin S (1998) Finding the genes in genomic DNA. Curr Opin Struct Biol 8(3):346–354

Buti M, Giordani T, Cattonaro F, Cossu R, Pistelli L, Vukich M, Morgante M, Cavallini A, Natali L (2011) Temporal dynamics in the evolution of the sunflower genome as revealed by sequencing and annotation of three large genomic regions. Theor Appl Genet 123(5):779–791

Chen N (2004) Using RepeatMasker to identify repetitive elements in genomic sequences. Curr Protoc Bioinformatics Chapter 4, Unit 4 10

Dangl JL, Jones JD (2001) Plant pathogens and integrated defence responses to infection. Nature 411(6839):826–833

Edgar RC (2004) Muscle: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32(5): 1792–1797

Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res 8(3):186–194

Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred I. Accuracy assessment. Genome Res 8(3):175–185

Flor HH (1947) Inheritance of reaction to rust in flax. J Agr Res 74:241–262

Gedil MA, Slabaugh MB, Berry S, Johnson R, Michelmore R, Miller J, Gulya T, Knapp SJ (2001a) Candidate disease resistance genes in sunflower cloned using conserved nucleotide-binding site motifs: genetic mapping and linkage to the downy mildew resistance gene *Pl1*. Genome 44(2):205–212

Gedil MA, Wye C, Berry S, Segers B, Peleman J, Jones R, Leon A, Slabaugh MB, Knapp SJ (2001b) An integrated restriction fragment length polymorphism—amplified fragment length polymorphism linkage map for cultivated sunflower. Genome 44(2):213–221

Gentzbittel L, Mouzeyar S, Badaoui S, Mestries E, Vear F, Tourvieille De Labrouhe D, Nicolas P (1998) Cloning of molecular markers for disease resistance in sunflower *Helianthus annuus* L. Theor Appl Genet 96(3):519–525

Grandbastien MA, Spielmann A, Caboche M (1989) Tnt1, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. Nature 337(6205):376–380

Johns MA, Mottinger J, Freeling M (1985) A low copy number, copia-like transposon in maize. EMBO J 4(5):1093–1101

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1(2):174–181

Lawson WR, Goulter KC, Henry RJ, Kong GA, Kochman JK (1998) Marker-assisted selection for two rust resistance genes in sunflower. Mol Breed 4:227–234

McHale L, Tan X, Koehl P, Michelmore RW (2006) Plant NBS-LRR proteins: adaptable guards. Genome Biol 7(4):212

Meyers BC, Chin DB, Shen KA, Sivaramakrishnan S, Lavelle DO, Zhang Z, Michelmore RW (1998a) The major resistance gene cluster in lettuce is highly duplicated and spans several megabases. Plant Cell 10(11):1817–1832

Meyers BC, Shen KA, Rohani P, Gaut BS, Michelmore RW (1998b) Receptor-like genes in the major resistance locus of lettuce are subject to divergent selection. Plant Cell 10(11):1833–1846

\Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnan S, Sobral BW, Young ND (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. Plant J 20(3): 317–332

Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW (2003) Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. Plant Cell 15(4):809–834

Mouzeyar S, Tourvielle De Labrouhe D, Vear F (1993) Histopathological studies of resistance of sunflower (*Helianthus annuus* L.) to downy mildew (*Plasmopara halstedii*). J Phytopathol 139: 289–297

Mouzeyar S, Roeckel-Drevet P, Gentzbittel L, Philippon J, Tourvieille De Labrouhe D, Vear F, Nicolas P (1995) RFLP and RAPD mapping of the sunflower *Pl1* locus for resistance to *Plasmopara halstedii* race 1. Theor Appl Genet 91(5):733–737

Mural RJ (2000) Artemis: a tool for displaying and annotating DNA sequence. Brief Bioinform 1(2):199–200

Peterson DG, Boehm KS, Stack SM (1997) Isolation of milligram quantities of nuclear DNA from tomato (*Lycopersicon esculentum*), a plant containing high levels of polyphenolic compounds. Plant Mol Biol Report 15(2):148–153

Peterson-Burch BD, Voytas DF (2002) Genes of the pseudoviridae (ty1/copia retrotransposons). Mol Biol Evol 19(11):1832–1845

Radwan O, Bouzidi MF, Vear F, Philippon J, De Labrouhe DT, Nicolas P, Mouzeyar S (2003) Identification of non-TIR-NBS-LRR markers linked to the *Pl5*/*Pl8* locus for resistance to downy mildew in sunflower. Theor Appl Genet 106(8):1438–1446

Radwan O, Bouzidi MF, Nicolas P, Mouzeyar S (2004) Development of PCR markers for the *Pl5*/*Pl8* locus for resistance to *Plasmopara halstedii* in sunflower, *Helianthus annuus* L. from complete CC-NBS-LRR sequences. Theor Appl Genet 109(1):176–185

Radwan O, Mouzeyar S, Venisse JS, Nicolas P, Bouzidi MF (2005) Induction of a sunflower CC-NBS-LRR resistance gene analogue during incompatible interaction with *Plasmopara halstedii*. J Exp Bot 56:2683–2693

Radwan O, Gandhi S, Heesacker A, Whitaker B, Taylor C, Plocik A, Kesseli R, Kozik A, Michelmore RW, Knapp SJ (2008) Genetic diversity and genomic distribution of homologs encoding NBS-LRR disease resistance proteins in sunflower. Mol Genet Genomics 280(2):111–125

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA 81(24):8014–8018

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory, New York

Santini S, Cavallini A, Natali L, Minelli S, Maggini F, Cionini PG (2002) Ty1/copia- and ty3/gypsy-like DNA sequences in *Helianthus* species. Chromosoma 111(3):192–200

Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc 3(6):1101–1108

Slabaugh MB, Yu JK, Tang S, Heesacker A, Hu X, Lu G, Bidney D, Han F, Knapp SJ (2003) Haplotyping and mapping a large cluster of downy mildew resistance gene candidates in sunflower using multilocus intron fragment length polymorphisms. Plant Biotechnol J 1(3):167–185

Soderlund C, Longden I, Mott R (1997) Fpc: a system for building contigs from restriction fingerprinted clones. Comput Appl Biosci 13(5):523–535

Soderlund C, Humphray S, Dunham A, French L (2000) Contigs built with fingerprints, markers, and FPC v4.7. Genome Res 10(11): 1772–1787

Tang S, Hass CG, Knapp SJ (2006) Ty3/gypsy-like retrotransposon knockout of a 2-methyl-6-phytyl-1,4-benzoquinone methyltransferase is non-lethal, uncovers a cryptic paralogous mutation, and produces novel tocopherol (vitamin e) profiles in sunflower. Theor Appl Genet 113(5):783–799

Vear F, Gentzbittel L, Philippon J, Mouzeyar S, Mestries E, Roeckel-Drevet P, Tourvieille De Labrouhe D, Nicolas P (1997) The genetics of resistance to five races of downy mildew (*Plasmopara halstedii*) in sunflower (*Helianthus annuus* L.). Theor Appl Genet 95(4):584–589

Vignols F, Rigau J, Torres MA, Capellades M, Puigdomenech P (1995) The brown midrib3 (bm3) mutation in maize occurs in the gene encoding caffeic acid *O*-methyltransferase. Plant Cell 7(4):407–416

Vranceanu VA, Stoenescu FM (1971) Pollen fertility restorer gene from cultivated sunflower (*Helianthus annuus* L.). Euphytica 20(4):536–541

Young ND (2000) The genetic architecture of resistance. Curr Opin Plant Biol 3(4):285–290