ORIGINAL PAPER

Comparative sequence analysis of the potato cyst nematode resistance locus *H1* reveals a major lack of co-linearity between three haplotypes in potato (*Solanum tuberosum* ssp.)

Anna Finkers-Tomczak · Erin Bakker · Jan de Boer · Edwin van der Vossen · Ute Achenbach · Tomasz Golas · Suwardi Suryaningrat · Geert Smant · Jaap Bakker · Aska Goverse

Abstract The *H1* locus confers resistance to the potato cyst nematode *Globodera rostochiensis* pathotypes 1 and 4. It is positioned at the distal end of chromosome V of the diploid *Solanum tuberosum* genotype SH83-92-488 (SH) on an introgression segment derived from *S. tuberosum* ssp. *andigena*. Markers from a high-resolution genetic map of the *H1* locus (Bakker et al. in Theor Appl Genet 109:146–152, 2004) were used to screen a BAC library to construct a physical map covering a 341-kb region of the resistant haplotype coming from SH. For comparison, physical maps were also generated of the two haplotypes from the diploid susceptible genotype RH89-039-16 (*S. tuberosum* ssp. *tuberosum/S. phureja*), spanning syntenic regions of 700 and 319 kb. Gene predictions on the genomic segments resulted in the identification of a large cluster consisting of variable numbers of the CC-NB-LRR type of *R* genes for each haplotype. Furthermore, the regions were interspersed with numerous transposable elements and genes coding for an extensin-like protein and an amino acid transporter. Comparative analysis revealed a major lack of gene order conservation in the sequences of the three closely related haplotypes. Our data provide insight in the evolutionary mechanisms shaping the *H1* locus and will facilitate the map-based cloning of the *H1* resistance gene.

Communicated by C. Gebhardt.

A. Finkers-Tomczak and E. Bakker contributed equally.

A. Finkers-Tomczak (✉) · E. Bakker · U. Achenbach ·
T. Golas · S. Suryaningrat · G. Smant · J. Bakker · A. Goverse
Laboratory of Nematology, Wageningen University,
P.O. Box 8123, 6700 ES Wageningen, The Netherlands
e-mail: Anna.Tomczak@wur.nl

J. de Boer
Wageningen UR Plant Breeding, P.O. Box 386,
6700 AJ Wageningen, The Netherlands

E. van der Vossen
Plant Research International BV, P.O. Box 16,
6700 AA Wageningen, The Netherlands

A. Finkers-Tomczak · E. Bakker · G. Smant · J. Bakker ·
A. Goverse
Centre for BioSystems Genomics, P.O. Box 98,
6700 AB Wageningen, The Netherlands

Introduction

Potato (*Solanum tuberosum*) is the fourth most important food crop in the world. Worldwide more than 19 million ha of potatoes are grown with a total economic value higher than 31 billion US$ (http://www.potato2008.org/en/world/index.html). Potato was brought to Europe by Spanish explorers in the sixteenth century from the Andean region of South America (Bradshaw and Ramsay 2005). Since then, it has been cultivated as a food crop by vegetative propagation of tubers. Together with the potatoes, the potato cyst nematodes (PCN) species *Globodera pallida* and *G. rostochiensis* became introduced into the old world. With the rapidly increasing acreage of cultivated potato in years following its introduction, the nematodes spread along over the European continent, and later on into other potato growing regions all over the world. Most current estimates suggest that the annual potato production worldwide is diminished by at least 10% due to infestations with potato cyst nematodes (Oerke et al. 1994).
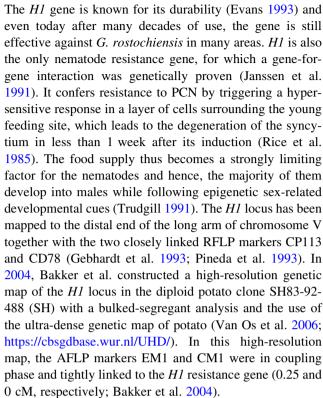
Potato cyst nematodes have a narrow host range parasitizing exclusively on members of the Solanaceae family,

including potato, eggplant and tomato (Sullivan et al. 2007). Potato cyst nematodes are sedentary endoparasites, whose survival and reproduction fully depends on a sophisticated feeding site induced by the nematodes inside a plant root (Gheysen and Mitchum 2009; Sobczak and Golinowski 2009). At the end of their life cycle, the fertilized adult females die and their dead hardened bodies form a protective cyst containing the eggs with the next generation of juveniles (Williamson and Hussey 1996). The cyst allows dormant PCN to survive for many years in the soil in temperate climate regions. Moreover, the juveniles of PCN will only hatch from the eggs when a suitable host plant is present. This long survival in the absence of host plants makes crop rotation not a very cost-effective method to control this pathogen. On the other hand, control of PCN by soil disinfection with unspecific pesticides presents its own disadvantages because of the environmental impact of the chemicals used. To prevent this damage the EU has recently introduced legislation to remove from use almost all active compounds effective against plant-parasitic nematodes (Rosso et al. 2009). Therefore, breeding nematode-resistant cultivars is gaining currently weight as a durable and environmentally friendly alternative for the other methods to control PCN.

So far, 14 PCN resistance gene loci have been mapped on eight linkage groups in potato. These resistances originate from wild potato species like *S. tuberosum* ssp. *andigena*, *S. vernei*, *S. spegazzinni*, *S. tuberosum* ssp. *tuberosum*, *S. tarijense* and *S. sparsipilum* (reviewed by Tomczak et al. 2009). They confer partial (*Gro1.2, Gro1.3, Gro1.4* (Kreike et al. 1996); $GpaIV_{adg}^{s}$ (Moloney et al. 2010); *Gpa* (Kreike et al. 1994); *Gpa5, Gpa6* (Rouppe van der Voort et al. 2000); *Grp1* (Rouppe van der Voort et al. 1998); $GpaV_{spb}^{S}$, $GpaXI_{spl}^{S}$ (Caromel et al. 2005); *GpaM1, GpaM2, GpaM3* (Caromel et al. 2003a, 2003b); $GpaXI_{tar}^{l}$ (Tan et al. 2009)) or nearly absolute (*H1* (Bakker et al. 2004; Kreike et al. 1993; Pineda et al. 1993); *GroVI* (Jacobs et al. 1996); *Gro1* (Ballvora et al. 1995; Barone et al. 1990; Paal et al. 2004) and *Gpa2* (Rouppe van der Voort et al. 1997; Van der Vossen et al. 2000)) resistance to one or more PCN pathotypes. The loci associated with nematode resistance often map to the regions known for carrying resistances to other plant pathogens, so called hot spots for resistance (Gebhardt and Valkonen 2001). To date, two PCN resistance genes have been characterized at the molecular level, viz. *Gpa2* from *S. tuberosum* ssp. *andigena* (Van der Vossen et al. 2000) and *Gro1* from *S. spegazzinni* (Paal et al. 2004). They both encode NB-LRR proteins, representing the largest class of resistance genes in plants.

The *H1* resistance gene was discovered in 1952 in *S. tuberosum* ssp. *andigena,* a genotype present in the Commonwealth Potato Collection. Since then, it has been introgressed into many commercially available cultivars.

The *H1* gene is known for its durability (Evans 1993) and even today after many decades of use, the gene is still effective against *G. rostochiensis* in many areas. *H1* is also the only nematode resistance gene, for which a gene-for-gene interaction was genetically proven (Janssen et al. 1991). It confers resistance to PCN by triggering a hypersensitive response in a layer of cells surrounding the young feeding site, which leads to the degeneration of the syncytium in less than 1 week after its induction (Rice et al. 1985). The food supply thus becomes a strongly limiting factor for the nematodes and hence, the majority of them develop into males while following epigenetic sex-related developmental cues (Trudgill 1991). The *H1* locus has been mapped to the distal end of the long arm of chromosome V together with the two closely linked RFLP markers CP113 and CD78 (Gebhardt et al. 1993; Pineda et al. 1993). In 2004, Bakker et al. constructed a high-resolution genetic map of the *H1* locus in the diploid potato clone SH83-92-488 (SH) with a bulked-segregant analysis and the use of the ultra-dense genetic map of potato (Van Os et al. 2006; https://cbsgdbase.wur.nl/UHD/). In this high-resolution map, the AFLP markers EM1 and CM1 were in coupling phase and tightly linked to the *H1* resistance gene (0.25 and 0 cM, respectively; Bakker et al. 2004).

Because of their tight linkage with *H1* resistance, we used the markers CM1 and EM1 to screen a BAC library of SH to construct a physical map spanning the *H1* locus. In addition, a physical map was also constructed for the matching regions of the two haplotypes of the susceptible diploid potato genotype RH89-039-16 (RH). Analysis of the three homologous genomic sequences revealed the presence of a large cluster of CC-NB-LRR genes on all three haplotypes. However, the sequences further revealed a substantial variation in the genomic organization at the *H1* locus among the three closely related haplotypes. We found only two co-linear regions, one of which is located outside the *R* gene cluster. Our data demonstrate that the *H1* resistance gene is located in a region that is highly divergent between haplotypes, occupied by large *R* gene clusters, rich in repeats and showing repressed recombination rates. All these findings implicate that a map-based cloning strategy may be not sufficient for cloning the functional *H1*. Hence, the extensive comparative sequence analysis presented in this paper may facilitate the identification of this durable nematode resistance gene.

Materials and methods

Plant material and DNA extraction

A diploid mapping population of a cross between the two potato clones SH83-92-488 (further referred to as SH) and

RH89-039-16 (further referred to as RH) has been previously developed for dissecting nematode resistances (Rouppe van der Voort et al. 1997). The female parent SH contains in a background of *S. tuberosum* an introgression segment originating from the wild accession *S. tuberosum* ssp. *andigena* CPC1673, which carries resistance to pathotype Ro1 and Ro4 of *Globodera rostochiensis*. The male parent RH has been selected for its fertility and the production of vigorous offspring and lacks resistance to any known PCN pathotype. It descends from the cross between SUH2293 and BC1034 and contains *S. tuberosum* ssp. *tuberosum* and *S. phureja* in genetic background. Plant genomic DNA was extracted from frozen leaf tissue of in vitro plants using the DNeasy Plant Mini Kit (QIAGEN Benelux B.V., Venlo, The Netherlands).

BAC library screening

The BAC library from the diploid clone SH83-92-488 consists of 97,920 clones and represents between 9 and 10 haploid genome equivalents. This BAC library is an extension of the BAC library containing 60,000 clones described by (Rouppe van der Voort et al. 1999). Pooling

and preparation of the BAC clones into 255 pools for screening was done as previously described in Kanyuka et al. (1999) and Bakker et al. (2003). Plate pools were screened using primer pairs based on genetic markers and *R* gene sequences (Table 1). The BAC coordinates of positive clones were identified by columns/rows PCR on bacterial cultures. The single BAC DNA was extracted from 500 ml culture on LB medium supplemented with 12.5 μg/ml of chloramphenicol for selection. Plasmid DNA was isolated using "Very Low-Copy Plasmid/Cosmid Purification" protocol included in the Plasmid Midi Kit (Qiagen, Hilden, Germany).

The RH BAC library consists of 78,336 clones (5× coverage of the diploid genome) and originates from the diploid potato clone RH89-039-16 (Borm 2008). To screen the RH library, a primer pair (H1probe3F: aca ttg gat gag cta aca ag; H1probe3R: atg act cca ccg att aga tc) was designed on the nucleotide binding domains of the *R* gene homologues in the region with 93–98% nucleotide identity. A 272-bp DNA probe was synthesized by PCR using an incubation at 94°C for 2 min, followed by 30 cycles of 94°C for 30 s, 58°C for 30 s and 72°C for 90 s, while finishing with an elongation step at 72°C for 5 min.

Table 1 Overview of markers and primer combinations used for screening of the BAC library and mapping of the BACs

Marker/primers pair	Type	F, R	Primer sequence	Anneal. temp. (°C)	Restriction enzymes	Selective nucleotides	Band size (bp)	References
H1LRR		F	CCTTATGTTAAACACCTCCTC	55			888	
		R	CGCAATATCTCCAAAACTGTC	55				
H1NBS		F	TTTCAAAGTTGGAAGAAGTTGG	55			689	
		R	GATCTTTTCTACGGTATCAGG	55				
CMI	AFLP				*Sac*I/*Mse*I	aga/cac	233	Bakker et al. (2004)
EMI	AFLP				*Eco*RI/*Mse*I	atg/gca	152	Bakker et al. (2004)
C43M51 318bp	AFLP				*Sac*I/*Mse*I	ata/cca	322	Van Os et al. (2006)
P22M39 152bp	AFLP				*Pst*I/*Mse*I	gt/aga	154.8	Van Os et al. (2006)
C39M50_53bp	AFLP				*Sac*I/*Mse*I	aga/cat	54.2	Van Os et al. (2006)
57R	SCAR	F	TGCCTGCCTCTCCGATTTCT	63			450	
		R	GGTTCAGCAAAAGCAAGGACGTG	63				
110L	SCAR	F	GGCCCTCCCCGATGATAATTAGTTTC	65			560	
		R	GGCTGTTATGGGTTATTTGGTGGGC	65				
202Sp6	SNP (LNA)	F	GGCAATGCTATGATGGAECTAG	64			216	
		R	TACGTTTAAAAGEACCTZCCAC	64				
micro_ttg_20	SSR	F	TGACTCCCGCTTTGGTTATC	60			231	
		R	AAGTGGGGTTTGGAGGGTAG	60				
micro_taa_129	SSR	F	TAGAATACGTGGGGGACTCG	59			277	
		R	AATCCTTTGCCAAATCATGG	59				

Hybridization was performed on filters onto which the RH BAC library was spotted in duplo. Preparation of the filters and hybridization were performed by Greenomics, Wageningen, The Netherlands.
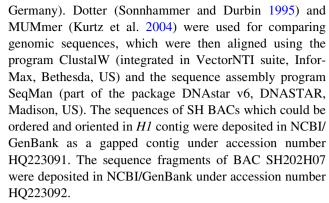
Physical mapping

To build scaffolds and contigs of positively selected SH BACs, the BAC inserts were fingerprinted using AFLP with *Eco*RI/*Mse*I restriction enzymes and adapters without selective nucleotides as described by (Vos et al. 1995). Similarly, the RH BACs were fingerprinted with a capillary sequencer as described by (Borm 2008). Band files were processed by the software package Finger Printed Contigs (FPC, Soderlund et al. 1997) to construct the minimal tiling paths.

Genetic mapping

AFLP analysis was performed using primer combinations from the UHD map of potato (http://www.plantbreeding.wur. nl/potatomap/) and the protocol described previously (Vos et al. 1995). The DNA sequences of the PCR primers, the corresponding thermal cycling conditions and details about the AFLP markers are listed in Table 1. Simple sequence repeat (SSR) primer pairs (Table 1) were used on BAC DNA and genomic potato DNA using the following thermal cycling conditions: 94°C for 5 s followed by 25 cycles of 94°C for 30 s, 56°C for 30 s and 72°C for 30 s, followed by 7 min incubation at 72°C. Visualisation of SSR markers was carried out using a Li-cor sequencer (Li-cor, Lincoln, NB, USA) according to the manufacturer's description. To map the SSR markers, a mapping population of 136 F1 genotypes from the cross between the diploid potato clones RH × SH was available (Rouppe van der Voort et al. 1997). The 45 most informative genotypes from a mapping population were selected using the software package MapPop (http://www. bio.unc.edu/faculty/vision/lab/mappop/) based on the maximum number of recombination events distributed over the genome. The UHD map of potato (Van Os et al. 2006) was used as input for MapPop. Segregating bands were mapped with the software package BINmap+ (Borm 2008). Mapping of SH BACs was performed using SCAR, SNP and AFLP markers (Table 1) on the set of 30 genotypes used previously for the construction of an high-resolution map (Bakker et al. 2004), and on 34 new recombinants based on recombination events between two PCR markers flanking the *H1* locus.

BAC sequencing and sequence analysis

Whole BAC clones were sequenced by shotgun sequencing with 6×, 10× or a full coverage (Greenomics, Wageningen, The Netherlands and GATC Biotech AG, Konstanz, Germany). Dotter (Sonnhammer and Durbin 1995) and MUMmer (Kurtz et al. 2004) were used for comparing genomic sequences, which were then aligned using the program ClustalW (integrated in VectorNTI suite, Infor-Max, Bethesda, US) and the sequence assembly program SeqMan (part of the package DNAstar v6, DNASTAR, Madison, US). The sequences of SH BACs which could be ordered and oriented in *H1* contig were deposited in NCBI/GenBank as a gapped contig under accession number HQ223091. The sequence fragments of BAC SH202H07 were deposited in NCBI/GenBank under accession number HQ223092.

Genes were annotated by combining predicted open reading frames from the gene finder program (FGENESH (Salamov and Solovyev 2000), using tomato as a model) with alignments of homologous sequences in public databases using Blastn and Blastp algorithms (Altschul et al. 1990). Annotation was also supported by finding a similarity with domains from curated domain databases at Interpro (Zdobnov and Apweiler 2001). The above-mentioned three types of data were used for manual curation resulting in the identification of genes at each haplotype.

Long terminal repeats (LTR) retrotransposons were identified using the LTR-finder tool (Xu and Wang 2007). Transposable elements were identified using CENSOR (Kohany et al. 2006) and transposon-related genes were identified with BLASTX search on NCBI (Altschul et al. 1990). Tandem repeats were predicted using Tandem Repeats Finder (Benson 1999). Simple repeats were identified with help of a DNA microsatellite repeat search utility called SPUTNIK (http://www.cbib.u-bordeaux2.fr/pise/sputnik.html).

Results

Physical map construction of the H1 locus for the resistant haplotype SH0

According to the previously constructed high-resolution map of the *H1* locus (Bakker et al. 2004), the AFLP markers CM1 and EM1 are in coupling phase with the *H1* resistance gene in the diploid potato genotype SH83-92-488. Marker CM1 could not be separated from resistance, while EM1 is separated by 0.25 cM. These findings were confirmed by screening an additional 1,116 progeny of the same SH × RH F1 mapping population (data not shown). Testing the additional progeny for CAPS marker 239E4left added 17 recombinants, which diminished its genetic distance from resistance to 2.1 cM. To identify the genome segment corresponding to markers CM1 and EM1, the SH BAC library comprising 97,920 clones with an estimated five times coverage of the diploid genome (Rouppe van der

Voort et al. 1999) was screened for the presence of these markers. This resulted in the identification of eight BAC clones, five of them selected with marker CM1 (i.e., SH196L20, SH057A05, SH110I21, SH245E19, and SH185K02) and three with marker EM1 (SH192L10, SH210E14, SH224A08), which were used for constructing a physical map of the resistant haplotype, which will be referred to as SH0.

For each of the eight selected BAC clones, an AFLP fingerprint was made with the restriction enzymes *Eco*RI and *Mse*I (Vos et al. 1995). Based on common bands in the fingerprints, a minimal tiling path was obtained consisting of four BACs (SH192L10; SH210E14; SH110I21, and SH057A05), which were all sequenced by shotgun sequencing. Analysis of the sequences confirmed the presence of the AFLP markers CM1 and EM1 and the presence of eight NB-LRR resistance gene homologues. Two other BACs reacting to either the CM1 or EM1 probes (SH245E19 and SH185K02) were fully overlapping with the minimal tiling path and as such, they were excluded from further study. The two remaining BACs (SH196L20 and SH224A08) that did not fit in the minimal tiling path were also sequenced. Surprisingly, their sequences contain neither the markers EM1 nor CM1 and do not overlap with any other BACs in the contig. They do, however, harbour three NB-LRR sequences that are highly homologous (88–90% similarity) to those on the four BACs that form one contig.

To extend the physical map of SH in this region, specific primer pairs were then developed on the sequence of: (1) conserved stretches of the LRR domains from the 11 resistance gene homologues present on the BACs, (2) the BAC ends delineating the minimal tiling path and (3) the BAC ends of the two BACs that could not be placed in the minimal tiling path. Screening the SH BAC library resulted in the identification of a novel BAC clone (SH202H07) that was positive for primers designed on one end of the minimal tiling path (SH057A05), for the LRR, and for a BAC end of BAC SH224A08. After sequencing this BAC, the overlapping regions could be confirmed and the physical map was extended with BAC SH224A08 and BAC SH202H07, resulting in an estimated total length of 341 kb.

In order to resolve the genetic position and haplotype of BACs SH057A05 and SH202H07, we designed a SCAR-marker for the end of BAC SH057A05 (57R) and an SNP marker for the end of BAC SH202H07 (202Sp6) and tested their behaviour in an SH × RH mapping population. Both markers were mapped to the *H1* locus and always co-segregated with nematode resistance. Using the sequences of the AFLP markers EM1 and CM1, which are closely linked to *H1* resista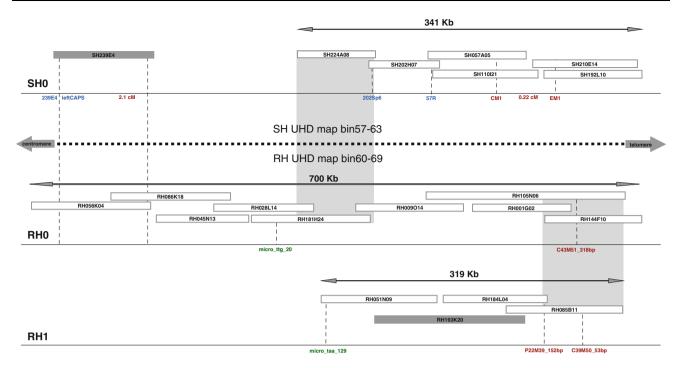nce, we further orientated the physical map and delimited the *H1* region for the resistant haplotype SH0 as presented in Fig. 1.

Physical map construction of the H1 locus for the two susceptible haplotypes of RH

To develop a similar physical map of the *H1* locus in the susceptible diploid potato clone RH89-039-16, a probe was designed based on the NBS region of the resistance gene homologues (RGHs) identified on the SH0 genome sequence. A BAC library of RH consisting of 78,336 clones and an estimated five times coverage of the diploid genome was spotted in duplo on hybridization filters. Hybridizing the probe under high-stringency conditions on the filters resulted in 157 positive BACs. Like for the SH BACs, the 157 RH BACs were fingerprinted using an AFLP reaction with the restriction enzymes *Eco*RI and *Mse*I, without selective nucleotides. Overlap between BACs was determined with FPC, resulting in 26 contigs. For each contig, one representative BAC was selected for sequencing. Sequence analysis revealed that 11 BACs harbour NB-LRR genes with high similarity to those present on the *H1* locus in SH0 (i.e. RH001G02, RH009O14, RH181H24, RH045N13, RH051N09, RH085-B11, RH186C17, RH154J09, RH125E08, RH053N17, and RH140O20; 69–100% aa identity). Two of them (RH181H24 and RH009O14) showed a minimal overlap with each other. Querying the HTGS nucleotide database of GenBank with the sequences of the 11 identified RH BACs using blastn algorithm (megablast) resulted in selecting 6 additional BACs (i.e. RH105N06, RH028L14, RH086K18, RH184L04, RH144F10 and RH056K04) which harbour *H1* homologues and/or showed an overlap with any of the 11 initially sequenced BACs from the RH library .

By using software for rapid aligment of genomes (MUMmer, Kurtz et al. 2004), followed by a BLAST search against the BAC-end-sequences database (BAC-end-tool, Borm 2008) and blastn (Altschul et al. 1990), it was possible to build two contigs of two (RH085B11 and RH184L04) and nine BACs (RH001G02, RH009O14, RH181H24, RH045N13, RH105N06, RH028L14, RH086-K18, RH144F10 and RH056K04), respectively. Application of the BAC-end-tool a second time resulted in the extension of the smallest contig with two additional BACs, i.e. RH051N09 and RH193K20. For RH051N09 the BAC sequence was available, whereas for RH193K20 only BAC-end sequences were available. Only 12 out of in total 17 identified and sequenced RH BACs could be physically mapped in *H1* region of RH, while remaining 5 anchored further from this region or on other chromosomes of potato (data not shown). Hence, the estimated physical lengths of the two contigs are approximately 700 and 319 kb.

Fig. 1 An integrated physical and genetic map of the *H1* locus on chromosome V of the resistant haplotype of the diploid potato clone SH83-92-488 (SH0), and two susceptible haplotypes of the diploid potato clone RH89-039-16 (RH0 and RH1). *Light grey* areas are corresponding to the co-linear regions between haplotypes. The *dotted line* shows orientation of the map on the chromosome and the *number placed below and above this line* refers to its relative position in the ultra-high-density map of SH/RH (Van Os et al. 2006). The *grey bars* represents a non-sequenced BAC clones. Physical distance is indicated in kilobase

Based on the presence of the three AFLP markers C34M51_(318bp), P22M39_(152bp) and C39M50_(53bp), which were previously mapped in the *H1* locus of RH in the ultra-dense genetic map of potato (Van Os et al. 2006), we could place the two contigs on two different haplotypes (from now on referred as RH0 and RH1; Fig. 1). Additionally, two BACs (RH051N09 and RH181H24) belonging to contigs mapped on two different haplotypes could be mapped in the same region (RH UHD map bin65–67 and bin60–62, respectively) by using SSR markers (Table 1). This further supported the anchoring of both contigs to the *H1* region on chromosome V of potato. Furthermore, identification of BAC RH056K04 sequence stretches of 96 and 99.1% identity with, respectively, the left and right end of the SH239E4 BAC, from which CAPS marker 239E4left was derived (Bakker et al. 2004), provides additional support for anchoring the RH0 contig to the *H1* genetic map.

Co-linearity between the three haplotypes SH0, RH0 and RH1 at the H1 locus

To study the co-linearity between the three physical maps, we have made a pairwise sequence comparison of all possible combinations of contiguous BAC sequences of SH0, RH0, and RH1 in a dot plot followed by a direct alignment of sequences. This allowed us to delineate a 92 kb region in SH0 and RH0 showing 100% identity (Fig. 2). The beginning of the co-linear stretch seems to be inverted, but we cannot completely exclude single BAC sequence assembly errors. In RH0 and RH1 a stretch of 102 kb was found with 95% overall sequence identity (Fig. 2). The sequences of RH1 and SH0, however, are completely non-linear and seem to share only a number of tandem repeats.

All the RH BACs and SH202H07 were sequenced with 6× coverage, resulting in 3–16 BAC sequence fragments. These fragments were assembled into scaffolds based on sequence identity/overlap between single BACs and co-linearity between the three haplotypes. Five BAC sequence scaffolds of SH0, 16 of RH0 and 12 of RH1 could be ordered and oriented (Fig. 3). For the remaining 31 scaffolds, this was not possible due to lack of sequence similarity between the 3 haplotypes, and insufficient overlap with other BACs from the same haplotype.

The co-linear region between RH0 and SH0 harbours 3 RGHs that are 100% identical. Between RH0 and RH1, a 102-kb stretch of 95% identity does not harbour any RGHs. The remaining part of the SH0, RH0 and RH1 sequences do not share any identical RGHs. Pairwise identity between full-length RGHs from all three haplotypes varies from 69 to 100% at the protein level (data not shown). Lack of synteny conservation observed between the major parts of SH0 and RH0 maps could be explained by the presence of the segment derived from *S. tuberosum* ssp. *andigena*
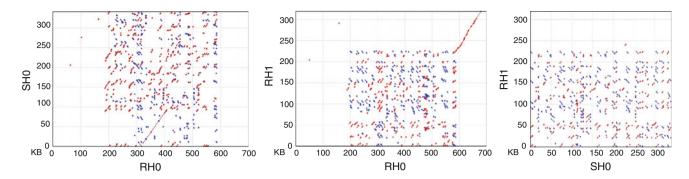
Fig. 2 Dot-plot (MUMmer) graphs comparing pairs of haplotypes: SH0 versus RH0, RH1 versus RH0 and RH1 versus SH0. *Grey* or *red lines* show forward matches and *black* or *blue lines* show the reverse matches between two sequences. The units for the labeled axes are bases
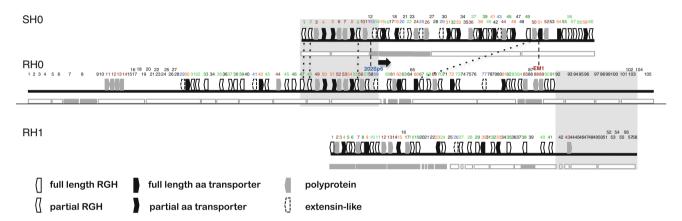

Fig. 3 Schematic overview of the genomic organization of the *H1* locus in SH and RH. Position and orientation of the ORFs were determined based on the genomic sequence of the resistant haplotype SH0 and the two susceptible haplotypes RH0 and RH1 derived from the diploid potato clones SH and RH, respectively. *Empty bars* represent sequence contigs with known orientation and order, and *grey bars* represent contigs for which orientation and order could not be determined. *Light grey areas* are corresponding to the co-linear regions between haplotypes. The putative start of the introgression fragment from *S. tuberosum* spp. *andigena* is indicated by a *black arrow*. Positions of all predicted ORFs are indicated by *numbers* that correspond to the numbers in Table S2. All ORFs annotated as RGHs, transposons, amino acid transporters and extensin-like genes are shown as *rectangles with arrowheads* indicating the direction of transcription. *Dotted line* connects RGHs that are highly similar between SH0 and RH0

that harbours the *H1* gene and has been introgressed into the *S. tuberosum* ssp. *tuberosum* background of SH. In that case we would expect that the region from marker 202Sp6 to the centromere is derived from *S. tuberosum* ssp. *tuberosum*. The two RGHs from the centromeric side of 202Sp6 are identical to RGHs from RH0 and are therefore no candidates for the functional nematode resistance gene *H1*. However, it is possible that a region between marker 202Sp6 and BAC SH239E4, which ends show very high identity with RH0 BACs, contains more divergent RGH sequences co-segregating with resistance. This should be investigated in the future. From the side of the telomere, the region harbouring functional candidates for the *H1* gene is delimited by the EM1 marker, which is not linked to resistance in a total of five per 2,189 offsprings (0.22 cM). This narrows down the genomic fragment carrying *H1* candidate resistance genes to 160 kb, between the markers 202Sp6 and EM1.

Genomic organization of the H1 locus in SH and RH: transposable elements and other repeats

While combining the gene predictions and sequence similarity data, we found 60, 105, and 58 open reading frames in three sequenced haplotypes SH0, RH0 and RH1 of the *H1* locus, respectively (Fig. 3; Table S2). This is an average of one gene per 5.2–6.6 kb. Twelve, 15 and 9 open reading frames in the SH0, RH0 and RH1 haplotypes code for proteins typically associated with transposable elements. In total, transposable elements occupy 17, 14 and 12% of the total sequence at SH0, RH0 and RH1, respectively.

The largest transposons found at all three haplotypes belong to the class of LTR-retrotransposons. SH0 includes eight predicted LTR-retrotransposons ranging in size from 2.6 to 11 kb. Six of them are in between resistance gene homologs, while two of them contain RGH coding

sequences. Sixteen LTR-retrotransposons were predicted (1.7–15 kb in size) in RH0. They are mostly interspersed between RGHs, but one LTR-retrotransposon carries an RGH-like sequence and three are located outside the *R* gene cluster. In RH1, four LTR-retrotransposons (4.8–13.8 kb in size) are present within a 39-kb fragment and one of them also carries an RGH-like sequence. The retrotransposons in SH0 and RH0 belong to the class I transposable element superfamilies Copia (4 and 6) and Gypsy (4 and 10) (Wicker et al. 2007). The RH1 haplotype, however, mostly harbours transposable elements from the Gypsy superfamily, clustered within a 40-kb region. Outside the RGH region in RH0, we found two additional clusters of transposable elements both consisting of at least three predicted retrotransposons. These latter retrotransposons stem from the Copia, Gypsy and LINE superfamilies (class I transposable elements), but are situated next to a class II transposable element named Mutator.

As the RH0 haplotype consists of a 400 kb large *R* gene cluster, flanked by regions that contain no RGHs, we could compare these two regions in terms of repetitive sequence composition. The region outside the *R* gene cluster has a lower percentage of repetitive sequences than the cluster itself (11 compared to 14%). The number and total length of the LTR-retrotransposons is lower outside the RGH cluster than inside, wherein LTR-retrotransposons account for more that 40% of the repeats. Furthermore, the resistance cluster in SH0 consists of 17% repeat sequences with an average length of 204 bp, while the RH0 and RH1 haplotypes have lower repeats content with smaller average sizes (14% and 189 bp and 12% with 163 bp, respectively). We have compared the transposable element content of the *H1* cluster and two other resistance gene clusters, viz. *Gpa2/Rx1* (AF265664) from potato and *Bs2* (AY702979) from pepper. At the *H1* locus (SH0) transposable elements occupy 3–4% more of the sequence than at the *Rx1/Gpa2* locus, but 8% less than at the *Bs2* locus. On average the LTR-retrotransposons are more abundant in the *Rx1/Gpa2* and *Bs2* clusters (60% of total TE length compared to 50% for *H1*). Furthermore, in the *H1* cluster the number of predicted LTRs larger than 1 kb belonging to the Copia-like family equals the number of Gypsy-like elements (1:1 ratio), while in both *Rx1/Gpa2* and *Bs2* we found a more typical ratio for potato of 2:1 Gypsy to Copia LTR-retrotransposons.

In addition to the large LTR-retrotransposons, which account for more than 40% of the total sequence encoding TEs, a substantial amount of smaller TEs and other repetitive sequences were discovered at the *H1* locus (Table S1). Sequence fragments with similarity to TEs from other orders of class I elements (LINE and SINE) and to the TEs from class II such as Mutator, Helitron, hAT, EnSpm and few others were identified in all three

haplotypes based on the sequence identity. Ten simple repeats were found in SH0, while RH0 and RH1 harbour 45 and 29 simple repeats, respectively. More than 5% of the sequence of each haplotype was predicted to comprise tandem repeats. SH0 contains 147 tandem repeats with a unit length varying from 7 to 335 bp and a copy number ranging from 1.9 to 20.4. RH0 contains 277 tandem repeats with a unit length from 7 to 335 bp and a copy number ranging from 1.8 to 23.9, while RH1 contains 116 tandem repeats of 9 to 232 bp long that are repeated 1.8 to 119.5 times.

Comparing the open reading frames in the haplotype sequences with expressed sequence tag data from the SGN database (Mueller et al. 2005) suggests that 40–51% of them code either for truncated proteins and short protein fragments. In the physical map of RH0, which covers the largest distance, two large regions (192 kb to the north and 102 kb to the south) without predicted RGHs flank the region harbouring 17 RGH-like sequences, which delimits the complete RGH cluster to approximately 400 kb. Two regions situated outside the predicted *R* gene clusters in RH0 and RH1 (Fig. 3, ORFs: RH0 1-28; 92–105 and RH1 42-58) harbour genes with homology to known functional genes, including DNA repair proteins, several transferases, sugar transporters, integrases, aspartic proteinases precursor, lipases, putative C2H2-type zinc finger proteins, glutamate decarboxylases, UDP-glucose pyrophosphorylases and ubiquitin protein ligases. The most significant Genbank hits and corresponding e-values are listed in Table S2.

Genomic organization of the *H1* locus in SH and RH: resistance gene homologs

The resistance gene homologues in all three haplotypes are located in single clusters, interspersed with several other genes, mostly occurring as single copies and coding for putative proteins without predicted homology. Two gene classes, however, occur as multiple copies and code for proteins with known function, namely, amino acid transporters and extensin-like proteins (Table S2; Fig. 3). Thirteen, nine and six copies of genes with high similarity to a transmembrane amino acid transporter from *Populus trichocarpa* (50% identity, XP_002316138) are present in SH0, RH0 and RH1, respectively. A transcript encoding a similar transmembrane amino acid transporter was not found in the potato expressed sequence tag database (at SGN), indicating that the gene is not abundantly expressed in potato. In comparison to the poplar genes most of the amino acid transporter-like genes we found in the *H1* cluster are truncated or contain indels, suggesting that they are non-functional. The genes with homology to an extensin-like gene show the highest similarity with an extensin-like protein from *Solanum tuberosum* (CAA06000)

and five, four and one copies are present at SH0, RH0 and RH1, respectively. An expressed sequence tag with 98% identity to the extensin-like protein was found in the potato ESTs collection (SGN-E558357).

A total of 55 RGHs were predicted in the haplotypes SH0, RH0, and RH1. They showed the highest amino acid identity (52%) to *Rpi-blb2*, a gene that confers resistance to *Phythophthora infestans* in *S. bulbocastanum* and that belongs to the CC-NB-LRR class of plant resistance genes (Van der Vossen et al. 2005). In SH0 17 RGHs were identified, of which only 5 were predicted to encode complete CC-NB-LRR type of resistance proteins. Two other open reading frames indicated as truncated (Fig. 3, SH0, ORFs 1 and 60) are located at the ends of BAC inserts. For these ORFs a part of the coding sequence information is missing and we are unable to predict whether these genes are full length. The rest of the open reading frames are most likely pseudogenes due to large deletions in the N-terminal, CC domains and LRR domains, or due to the occurrence of premature stop codons. In RH0 we found 25 RGHs, while 17 RGHs seem to be present in the RH1 haplotypes. Eight genes in RH0 and five genes in RH1 likely encode complete resistance proteins, while eight RGHs contain either deletions or premature stop codons resulting in truncated *R* genes. For one RGH some sequence information is missing because of its position at a BAC end. Position and orientation of all the RGHs located in this genomic region in SH0 and RH were determined and presented in Fig. 3. A detailed overview of the annotation of the *H1* locus is shown in Table S2.

Discussion

Here we describe the construction of physical maps of the *H1* locus on chromosome V in the diploid *S. tuberosum* clone SH83-92-488 and the corresponding genomic region in the two haplotypes of the diploid susceptible potato clone RH89-039-16. The *H1* locus was introgressed from the subspecies *S. tuberosum* ssp. *andigena* because it confers durable resistance to specific pathotypes of the potato cyst nematode *Globodera rostochiensis*. Although the *H1* resistance gene has not been identified yet, comparing the resistant haplotype sequence with the two susceptible haplotypes at the *H1* locus provides crucial insights that may help us to home in on the gene. The sequence information obtained in this study can also be used to develop specific markers and a candidate gene approach for the identification of the *GroVI* gene, which is another single dominant resistance gene against *G. rostochiensis* that has been mapped in a syntenic region on chromosome V at an introgression segment from *Solanum vernei* (Jacobs et al. 1996).

The sequences of the three haplotypes indeed revealed a number of remarkable features of this locus in *S. tuberosum*. First of all, each haplotype harbours a large cluster of resistance gene homologues from the CC-NB-LRR resistance gene class with significant sequence similarity to the late blight resistance gene *Rpi-blb2* identified at chromosome VI of potato (Van der Vossen et al. 2005). Surprisingly, only three RGHs on the resistant haplotype (SH0) have identical genes on one of the susceptible haplotypes (RH0), while no matching sequences were found on the second susceptible haplotype (RH1). One could argue that positive selection in the RGHs at the *H1* locus has been so extensive at this locus that accelerated evolution yielded highly diverse RGHs. However, also outside the coding regions no similarity was observed causing an overall lack of synteny between the different haplotypes in a major part of the *R* gene cluster. In fact, only 92 of 341 kb of the resistant haplotype is co-linear with the susceptible haplotype (with 100% identity in RH0), while another segment of only 102 out of 700 kb in RH0 is co-linear with RH1 (with 95% identity).

The physical maps were constructed by using the two AFLP markers CM1 and EM1. On the genetic map CM1 was at 0 cM distance from the *H1* resistance, while EM1 was positioned at 0.25 cM distance. The region between these two markers in the physical maps spans more than 65 kb. A remarkable difference in the recombination rate was observed between two sub-regions of the physical map of the resistant haplotype SH0. In the centromeric part of the SH map 65 kb equals 0.25 cM genetic distance, whereas no recombination was observed in the northern region of the map resulting in a complete linkage of a more than 170 kb genomic fragment with the resistance trait. Such a considerable disparity between the genetic and physical distances pointing at a strong repression of recombination was found previously in the maps of the *I3* resistance locus in tomato (Lim et al. 2008), the *R1* locus in potato (Kuang et al. 2005) and the *Mla* locus in barley (Wei et al. 1999). For the *Mi-1* locus, which was introgressed into *L. esculentum* from its wild relative *L. peruvianum,* an even more severe suppression of recombination could be observed resulting in the complete linkage of a 550 kb region to resistance. This repression of recombination is thought to be the consequence of the exogenous origin of the DNA segment, resulting in hemizygosity between haplotypes. A similar phenomenon was also observed for other *R* genes introgressed from wild species in various crosses (Ganal and Tanksley 1996). In case of *Mi-1*, the proximity to the centromere may also contribute to the repression of recombination in tomato (Kaloshian et al. 1998). Hence, the suppression of recombination that we observed north of the *H1* locus in resistant plants is likely caused by the sequence divergence between the

S. tuberosum ssp. *andigena* introgression segment and *S. tuberosum* ssp. *tuberosum* background. It is therefore hypothesized that the structural differences between the homologous chromosomes could interfere with chromosome pairing and crossing-over during meiosis (Ballvora et al. 2007).

After aligning the sequences of the three physical maps, complete loss of co-linearity was found in a region of 250 kb between the SH0 haplotype and the two RH haplotypes. A significant homology in this part of the map is observed only between two fragments carrying RGHs sequences, but the position of these fragments is not syntenic when both maps are aligned and their flanking sequence differ substantially. As this region is flanked by a region almost identical between the SH0 and RH0 haplotype, and harbours four RGH-like sequences, it would be interesting to see if sequence co-linearity continues further on the centromeric side of the *H1* locus. The finding of two other co-linear BACs from SH and RH that have been preliminary mapped centromeric to the *H1* cluster (data not shown) as well as the identification within the RH0 sequence both ends of the BAC SH239E4, from which marker 239E4left was generated and mapped at a distance of 2.1 cM from *H1* resistance, suggest that this indeed may be the case. A lack of co-linearity was observed also for a major part between the two susceptible haplotypes, reflecting possible differences in the genetic background of RH89-039-16 in this region likely related to presence of genetic material derived from *S. phureja* (Ramanna 1983). Beyond this part, in the region on the telomeric side, it was not possible to further distinguish between the RH0 and RH1 haplotypes. Apparently, the two haplotypes are very homozygous in this region resulting in a single BAC contig when constructing the physical map. Such large differences in synteny between haplotypes are comparable to that observed for the *R1* cluster in *S. demissum* (Kuang et al. 2005) and *S. tuberosum* (Ballvora et al. 2002). High sequence divergence was also found between the resistant and susceptible haplotype of SH83-92-488 at the *Rx1/Gpa2* locus in potato, as these two haplotypes could only be aligned using SSR markers (Butterbach 2007). In heterozygous, outbreeding species like potato such a high divergence between haplotypes potentially generates a high diversity of immune receptors, which is advantageous in responding to quickly changing pathogen populations (Hughes and Yeager 1998). In general, the high level of natural intraspecific polymorphisms between haplotypes can often be correlated with complex *R* gene clusters and provides evidence for strong evolutionary forces shaping these parts of the plant genome.

Transposable elements (TEs) in general are thought to have a major impact on the differentiation of plant species at the level of genome structure (Bennetzen 2005). In various plant species, including potato, tomato, wheat, *Poncirus*, *Arabidopsis*, wheat, barley and soybean, *R* gene clusters like the *H1* locus were found to co-localize with repetitive sequences and TEs (Ballvora et al. 2007; Gao and Bhattacharyya 2008; Kuang et al. 2005; Noel et al. 1999; Panstruga et al. 1998; Seah et al. 2007; Wei et al. 2002; Wicker et al. 2001; Yang et al. 2003). An extreme example is the *Mla* region in barley, where all major classes of TEs are represented forming two large nested complexes, flanking two RGHs sequences, likely contributing to repression of recombination in this region (Wei et al. 2002). Presence of similar copies of TEs in one *R* gene cluster enable unequal crossing-over causing expansion or contraction of the locus (Wicker et al. 2007), while on the other hand the TEs diversity in the same region may prevent unequal crossing-overs (Kuang et al. 2005; Song et al. 1997). As a result of similar TEs copies on different genome locations, sequence exchange can also occur between otherwise non-homologous regions (Meyers et al. 2003). Like shown for the *Bs2* locus in pepper (Mazourek et al. 2009), we observed erosion and truncation of TE related sequences in the *H1* region. This can be explained by multiple insertions followed by sequence drift. More detailed sequence analysis should be performed in the future in order to assess the variety of TEs in the *H1* cluster and to explain their influence on the evolutionary history of this *R* gene locus.

Apart from TEs and RGHs, resistance gene clusters often harbour not many other functional genes. For example, in soybean at the *Rps1-k* locus only a few full-length genes were predicted within a distance of 118 kb, including two *R* genes and four retrotransposons (Gao and Bhattacharyya 2008). Some of the non-*R* genes present in *R* gene clusters have functions related to plant defence such as is reported for the *Mla* locus in barley (Wei et al. 2002), the *Mi-1* locus in tomato (Seah et al. 2007) and the *R1* and *Rx1/Gpa2* loci in potato (Ballvora et al. 2007; Butterbach 2007). At the *H1* locus, the RGH regions in all three haplotypes are interspersed by several copies of genes coding for extensin-like proteins. Although the expression of a potato extensin-like gene is induced by wounding and *Erwinia carotovora* infection of potato tubers (Rumeau et al. 1990), its possible role in plant defence remains to be shown (Dey et al. 1997). Recently, microarray studies have shown that extensins are also up-regulated during cyst nematode infections (Ithal et al. 2007; Khan et al. 2004; Puthoff et al. 2003). They may function in strengthening of the cell wall as an early plant defence response as well as in strengthening the syncytium wall in favour of nematodes (Khan et al. 2004). The presence of multiple copies of extensin-like genes at the *H1* locus may point at a role for these genes in defence responses against potato cyst nematodes. Another plausible explanation for the occurrence of

similar gene copies distributed throughout the three RGH clusters in SH and RH is that they derive from a common ancestor after tandem duplication together with the flanking RGH and subsequent genetic erosion.

At the *H1* locus we found, besides full-length RGHs, ORFs coding for short fragments of resistance gene homologues. The occurrence of such duplicated partial *R* gene fragments inserted upstream or downstream of *R* genes has been reported previously (Mazourek et al. 2009). Although their postulated role in controlling *R* gene expression has not been proven yet, it opens a new field for speculations and functional studies. One possibility is that the translation of such small *R* gene paralogues could have a role in the regulation of *R* gene function through the forming of heteroduplexes (Huang et al. 2005), as it was shown that the intermolecular and intramolecular interactions between R protein domains might function as activation switches upon recognition of the cognate elicitors (Moffett et al. 2002). The need for additional protein components in the formation of functional *R* gene complexes could be an explanation for the phenomena that *R* genes loose their full effectiveness upon introduction in another genetic background (Jacquet et al. 2005). An alternative explanation for the presence of so many *R* gene fragments and truncated homologs at the *H1* cluster is that they are the result of frequent rearrangements occurring at this locus and constitute a reservoir of variation and the generation of new resistance specificities (Michelmore and Meyers 1998).

In this paper, we describe the integration of the physical and genetic map of the *H1* locus, which shows that four *R* gene homologues in SH0 with sequence similarity to the CC-NBS-LRR genes *Mi-1* (Milligan et al. 1998; Vos et al. 1998) from tomato and *Rpi-blb2* from potato (Van der Vossen et al. 2005) are linked to nematode resistance. Therefore, they are considered to be good candidate genes for the potato cyst nematode resistance gene *H1*. The low level of recombination together with the highly repetitive nature of *R* gene loci (Hulbert et al. 2001) is a major obstacle in the positional cloning of *R* genes (Huang et al. 2005). Moreover, the heterozygosity of the potato genome adds to the difficulties associated with this approach (Kanyuka et al. 1999). Hence, it is crucial to support the positional cloning of *R* genes with alternative strategies including the candidate gene approach and comparative genomics (Huang et al. 2005; this paper). Soon, the complete potato genome sequence will be available, which can boost the identification of genes underlying important resistance traits in the future.

References

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

Bakker E, Butterbach P, Rouppe van der Voort J, Van Der Vossen E, Van Vliet J, Bakker J, Goverse A (2003) Genetic and physical mapping of homologues of the virus resistance gene *Rx1* and the cyst nematode resistance gene *Gpa2* in potato. Theor Appl Genet 106:1524–1531

Bakker E, Achenbach U, Bakker J, Van Vliet J, Peleman J, Segers B, Van Der Heijden S, Van Der Linde P, Graveland R, Hutten R, Van Eck H, Coppoolse E, Van Der Vossen E, Bakker J, Goverse A (2004) A high-resolution map of the *H1* locus harbouring resistance to the potato cyst nematode *Globodera rostochiensis*. Theor Appl Genet 109:146–152

Ballvora A, Hesselbach J, Niewohner J, Leister D, Salamini F, Gebhardt C (1995) Marker enrichment and high-resolution map of the segment of potato chromosome VII harboring the nematode resistance gene *Gro1*. Mol Gen Genet 249:82–90

Ballvora A, Ercolano MR, Weiss J, Meksem K, Bormann CA, Oberhagemann P, Salamini F, Gebhardt C (2002) The R1 gene for potato resistance to late blight (*Phytophthora infestans*) belongs to the leucine zipper/NBS/LRR class of plant resistance genes. Plant J 30:361–371

Ballvora A, Jöcker A, Viehöver P, Ishihara H, Paal J, Meksem K, Bruggmann R, Schoof H, Weisshaar B, Gebhardt C (2007) Comparative sequence analysis of *Solanum* and Arabidopsis in a hot spot for pathogen resistance on potato chromosome V reveals a patchwork of conserved and rapidly evolving genome segments. BMC Genomics 8:112

Barone A, Ritter E, Schachtschabel U, Debener T, Salamini F, Gebhardt C (1990) Localization by restriction fragment length polymorphism mapping in potato of a major dominant gene conferring resistance to the potato cyst nematode *Globodera rostochiensis*. Mol Gen Genet 224:177–182

Bennetzen JL (2005) Transposable elements, gene creation and genome rearrangement in flowering plants. Curr Opin Genet Dev 15:621–627

Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 27:573–580

Borm TJA (2008) Construction and use of a physical map of potato. Wageningen University, Wageningen

Bradshaw JE, Ramsay G (2005) Utilisation of the Commonwealth Potato Collection in potato breeding. Euphytica 146:9–19

Butterbach P (2007) Molecular evolution of the disease resistance gene *Rx* in Solanum. Wageningen University, Wageningen

Caromel B, Mugniery D, Lefebvre V, Andrzejewski S, Ellisseche D, Kerlan MC, Rousselle P, Rousselle-Bourgeois F (2003a) Mapping QTLs for resistance against *Globodera pallida* (Stone) Pa2/3 in a diploid potato progeny originating from *Solanum spegazzinii*. Theor Appl Genet 106:1517

Caromel B, Mugniery D, Lefebvre V, Andrzejewski S, Ellisseche D, Kerlan MC, Rousselle P, Rousselle-Bourgeois F (2003b)

Mapping QTLs for resistance against *Globodera pallida* (Stone) Pa2/3 in a diploid potato progeny originating from *Solanum spegazzinii*. Theor Appl Genet 106:1517

Caromel B, Mugniery D, Kerlan MC, Andrzejewski S, Palloix A, Ellisseche D, Rousselle-Bourgeois F, Lefebvre V (2005) Resistance quantitative trait loci originating from *Solanum sparsipilum* act independently on the sex ratio of *Globodera pallida* and together for developing a necrotic reaction. Mol Plant Microbe Interact 18:1186–1194

Dey PM, Brownleader MD, Pantelides AT, Trevan M, Smith JJ, Saddler G (1997) Extensin from suspension-cultured potato cells: a hydroxyproline-rich glycoprotein, devoid of agglutinin activity. Planta 202:179–187

Evans K (1993) New approaches for potato cyst nematode management. Nematropica 23:221–231

Ganal MW, Tanksley SD (1996) Recombination around the *Tm2a* and *Mi* resistance genes in different crosses of *Lycopersicon peruvianum*. Theor Appl Genet 92:101–108

Gao H, Bhattacharyya MK (2008) The soybean–*Phytophthora* resistance locus *Rps1-k* encompasses coiled coil-nucleotide binding-leucine rich repeat-like genes and repetitive sequences. BMC Plant Biol 8

Gebhardt C, Valkonen JPT (2001) Organization of genes controlling disease resistance in the potato genome. Annu Rev Phytopathol 39:79–102

Gebhardt C, Mugniery D, Ritter E, Salamini F, Bonnel E (1993) Identification of RFLP markers closely linked to the *H1* gene conferring resistance to *Globodera rostochiensis* in potato. Theor Appl Genet 85:541–544

Gheysen G, Mitchum MG (2009) Molecular insights in the susceptible plant response to nematode infection. In: Berg RH, Taylor CG (eds) Plant cell monographs. Springer, Berlin, pp 45–81

Huang S, Van Der Vossen EAG, Kuang H, Vleeshouwers VGAA, Zhang N, Borm TJA, Van Eck HJ, Baker B, Jacobsen E, Visser RGF (2005) Comparative genomics enabled the isolation of the *R3a* late blight resistance gene in potato. Plant J 42:251–261

Hughes AL, Yeager M (1998) Natural selection and the evolutionary history of major histocompatibility complex loci. Front Biosci 3:d509–d516

Hulbert SH, Webb CA, Smith SM, Sun Q (2001) Resistance gene complexes: evolution and utilization. Annu Rev Phytopathol 39:285–312

Ithal N, Recknor J, Nettleton D, Maier T, Baum TJ, Mitchum MG (2007) Developmental transcript profiling of cyst nematode feeding cells in soybean roots. Mol Plant Microbe Interact 20:510–525

Jacobs JME, Van Eck HJ, Horsman K, Arens PFP, Verkerk-Bakker B, Jacobsen E, Pereira A, Stiekema WJ (1996) Mapping of resistance to the potato cyst nematode *Globodera rostochiensis* from the wild potato species *Solanum vernei*. Mol Breed 2:51–60

Jacquet M, Bongiovanni M, Martinez M, Verschave P, Wajnberg E, Castagnone-Sereno P (2005) Variation in resistance to the root-knot nematode *Meloidogyne incognita* in tomato genotypes bearing the *Mi* gene. Plant Pathol 54:93–99

Janssen R, Bakker J, Gommers FJ (1991) Mendelian proof for a gene-for-gene relationship between virulence of *Globodera rostochiensis* and the *H1* resistance gene in *Solanum tuberosum* ssp. *andigena* CPC 1673. Rev Nematol 14:213–219

Kaloshian I, Yaghoobi J, Liharska T, Hontelez J, Hanson D, Hogan P, Jesse T, Wijbrandi J, Simons G, Vos P, Zabel P, Williamson VM (1998) Genetic and physical localization of the root-knot nematode resistance locus *Mi* in tomato. Mol Gen Genet 257:376–385

Kanyuka K, Bendahmane A, Rouppe van der Voort JNAM, Van Der Vossen EAG, Baulcombe DC (1999) Mapping of intra-locus duplications and introgressed DNA: Aids to map-based cloning of genes from complex genomes illustrated by physical analysis of the *Rx* locus in tetraploid potato. Theor Appl Genet 98:679–689

Khan R, Alkharouf N, Beard H, MacDonald M, Chouikha I, Meyer S, Grefenstette J, Knap H, Matthews B (2004) Microarray analysis of gene expression in soybean roots susceptible to the soybean cyst nematode two days post invasion. J Nematol 36:241–248

Kohany O, Gentles AJ, Hankus L, Jurka J (2006) Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. BMC Bioinformatics 7:474

Kreike CM, Dekoning JRA, Vinke JH, Vanooijen JW, Gebhardt C, Stiekema WJ (1993) Mapping of loci involved in quantitatively inherited resistance to the potato cyst-nematode *Globodera rostochiensis* pathotype-Ro1. Theor Appl Genet 87:464–470

Kreike CM, de Koning JRA, Vinke JH, van Ooijen JW, Stiekema WJ (1994) Quantitatively-inherited resistance to *Globodera pallida* is dominated by one major locus in *Solanum spegazzinii*. Theor Appl Genet 88:764–769

Kreike CM, Kok-Westeneng AA, Vinke JH, Stiekema WJ (1996) Mapping of QTLs involved in nematode resistance, tuber yield and root development in *Solanum* sp. Theor Appl Genet 92:463–470

Kuang H, Wei F, Marano MR, Wirtz U, Wang X, Liu J, Shum WP, Zaborsky J, Tallon LJ, Rensink W, Lobst S, Zhang P, Tornqvist CE, Tek A, Bamberg J, Helgeson J, Fry W, You F, Luo MC, Jiang J, Robin Buell C, Baker B (2005) The *R1* resistance gene cluster contains three groups of independently evolving, type I *R1* homologues and shows substantial structural variation among haplotypes of *Solanum demissum*. Plant J 44:37–51

Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL (2004) Versatile and open software for comparing large genomes. Genome Biol 5(2):R12

Lim GTT, Wang GP, Hemming MN, McGrath DJ, Jones DA (2008) High resolution genetic and physical mapping of the *I-3* region of tomato chromosome 7 reveals almost continuous microsynteny with grape chromosome 12 but interspersed microsynteny with duplications on Arabidopsis chromosomes 1, 2 and 3. Theor Appl Genet 118:57–75

Mazourek M, Cirulli ET, Collier SM, Landry LG, Kang BC, Quirin EA, Bradeen JM, Moffett P, Jahn MM (2009) The fractionated orthology of *Bs2* and *Rx/Gpa2* supports shared synteny of disease resistance in the solanaceae. Genetics 182:1351–1364

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

Michelmore RW, Meyers BC (1998) Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. Genome Res 8:1113–1130

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

Moffett P, Farnham G, Peart J, Baulcombe DC (2002) Interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death. EMBO J 21:4511

Moloney C, Griffin D, Jones PW, Bryan GJ, McLean K, Bradshaw JE, Milbourne D (2010) Development of diagnostic markers for use in breeding potatoes resistant to *Globodera pallida* pathotype Pa2/3 using germplasm derived from *Solanum tuberosum* ssp. *andigena* CPC 2802. Theor Appl Genet 120:679–689

Mueller LA, Solow TH, Taylor N, Skwarecki B, Buels R, Binns J, Lin C, Wright MH, Ahrens R, Wang Y, Herbst EV, Keyder ER, Menda N, Zamir D, Tanksley SD (2005) The SOL Genomics Network. A comparative resource for Solanaceae biology and beyond. Plant Physiol 138:1310–1317

Noel L, Moores TL, Van Der Biezen EA, Parniske M, Daniels MJ, Parker JE, Jones JDG (1999) Pronounced intraspecific haplotype divergence at the *RPP5* complex disease resistance locus of Arabidopsis. Plant Cell 11:2099–2111

Oerke EC, Dehne HW, Schönbeck F, Weber A (eds) (1994) Crop production and crop protection: estimated losses in major food and cash crops. Elsevier, Amsterdam

Paal J, Henselewski H, Muth J, Meksem K, Menendez CM, Salamini F, Ballvora A, Gebhardt C (2004) Molecular cloning of the potato Gro1–4 gene conferring resistance to pathotype Ro1 of the root cyst nematode *Globodera rostochiensis*, based on a candidate gene approach. Plant J 38:285–297

Panstruga R, Büschges R, Piffanelli P, Schulze-Lefert P (1998) A contiguous 60 kb genomic stretch from barley reveals molecular evidence for gene islands in a monocot genome. Nucleic Acids Res 26:1056–1062

Pineda O, Bonierbale MW, Plaisted RL, Brodie BB, Tanksley SD (1993) Identification of RFLP markers linked to the *H1* gene conferring resistance to the potato cyst nematode *Globodera rostochiensis*. Genome 36:152

Puthoff DP, Nettleton D, Rodermel SR, Baum TJ (2003) Arabidopsis gene expression changes during cyst nematode parasitism revealed by statistical analyses of microarray expression profiles. Plant J 33:911–921

Ramanna MS (1983) First division restitution gametes through fertile desynaptic mutants of potato. Euphytica 32:337–350

Rice SL, Leadbeater BSC, Stone AR (1985) Changes in cell structure in roots of resistant potatoes parasitized by potato cyst-nematodes. Potatoes with resistance gene *H1* derived from *Solanum tuberosum* ssp. *andigena*. Physiol Plant Pathol 27:219–234

Rosso MN, Jones JT, Abad P (2009) RNAi and functional genomics in plant parasitic nematodes. Annu Rev Phytopathol 47:207–232

Rouppe van der Voort J, Wolters P, Folkertsma R, Hutten R, Van Zandvoort P, Vinke H, Kanyuka K, Bendahmane A, Jacobsen E, Janssen R, Bakker J (1997) Mapping of the cyst nematode resistance locus *Gpa2* in potato using a strategy based on comigrating AFLP markers. Theor Appl Genet 95:874–880

Rouppe van der Voort J, Lindeman W, Folkertsma R, Hutten R, Overmars H, Van Der Vossen E, Jacobsen E, Bakker J (1998) A QTL for broad-spectrum resistance to cyst nematode species (*Globodera* spp.) maps to a resistance to gene cluster in potato. Theor Appl Genet 96:654–661

Rouppe van der Voort J, Kanyuka K, Van Der Vossen E, Bendahmane A, Mooijman P, Klein-Lankhorst R, Stiekema W, Baulcombe D, Bakker J (1999) Tight physical linkage of the nematode resistance gene *Gpa2* and the virus resistance gene *Rx* on a single segment introgressed from the wild species *Solanum tuberosum* subsp. *andigena* CPC 1673 into cultivated potato. Mol Plant Microbe Interact 12:197–206

Rouppe van der Voort J, Van der Vossen E, Bakker E, Overmars H, Van Zandvoort P, Hutten R, Klein Lankhorst RK, Bakker J (2000) Two additive QTLs conferring broad-spectrum resistance in potato to *Globodera pallida* are localized on resistance gene clusters. Theor Appl Genet 101:1122–1130

Rumeau D, Maher EA, Kelman A, Showalter AM (1990) Extensin and phenylalanine ammonia-lyase gene expression altered in potato tubers in response to wounding, hypoxia, and *Erwinia carotovora* infection. Plant Physiol 93:1134–1139

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

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

Sobczak M, Golinowski W (2009) Structure of cyst nematode feeding sites. In: Berg RH, Taylor CG (eds) Plant cell monographs. Springer, Berlin, pp 153–187

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

Song WY, Pi LY, Wang GL, Gardner J, Holsten T, Ronald PC (1997) Evolution of the rice *Xa21* disease resistance gene family. Plant Cell 9:1279–1287

Sonnhammer ELL, Durbin R (1995) A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene 167:GC1–GC10

Sullivan MJ, Inserra RN, Franco J, Moreno-Leheude I, Greco N (2007) Potato cyst nematodes: plant host status and their regulatory impact. Nematropica 37:193–201

Tan MYA, Park TH, Alles R, Hutten RCB, Visser RGF, Van Eck HJ (2009) GpaXI$^l_{tar}$ originating from *Solanum tarijense* is a major resistance locus to *Globodera pallida* and is localised on chromosome 11 of potato. Theor Appl Genet 119:1477–1487

Tomczak A, Koropacka K, Smant G, Goverse A, Bakker E (2009) Resistant plant responses. In: Berg RH, Taylor CG (eds) Plant cell monographs. Springer, Berlin, pp 83–113

Trudgill DL (1991) Resistance to and tolerance of plant parasitic nematodes in plants. Annu Rev Phytopathol 29:167–192

Van der Vossen EAG, Rouppe van der Voort JNAM, Kanyuka K, Bendahmane A, Sandbrink H, Baulcombe DC, Bakker J, Stiekema WJ, Klein-Lankhorst RM (2000) Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: a virus and a nematode. Plant J 23:567–576

Van der Vossen EA, Gros J, Sikkema A, Muskens M, Wouters D, Wolters P, Pereira A, Allefs S (2005) The *Rpi-blb2* gene from *Solanum bulbocastanum* is an *Mi-1* gene homolog conferring broad-spectrum late blight resistance in potato. Plant J 44:208–222

Van Os H, Andrzejewski S, Bakker E, Barrena I, Bryan GJ, Caromel B, Ghareeb B, Isidore E, De Jong W, Van Koert P, Lefebvre V, Milbourne D, Ritter E, Rouppe van der Voort JNAM, Rousselle-Bourgeois F, Van Vliet J, Waugh R, Visser RGF, Bakker J, Van Eck HJ (2006) Construction of a 10,000-marker ultradense genetic recombination map of potato: providing a framework for accelerated gene isolation and a genomewide physical map. Genetics 173:1075–1087

Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Vos P, Simons G, Jesse T, Wijbrandi J, Heinen L, Hogers R, Frijters A, Groenendijk J, Diergaarde P, Reijans M, Fierens-Onstenk J, de Both M, Peleman J, Liharska T, Hontelez J, Zabeau M (1998) The tomato *Mi-1* gene confers resistance to both root-knot nematodes and potato aphids. Nat Biotechnol 16:1365–1369

Wei F, Gobelman-Werner K, Morroll SM, Kurth J, Mao L, Wing R, Leister D, Schulze-Lefert P, Wise RP (1999) The *Mla* (powdery mildew) resistance cluster is associated with three NBS-LRR gene families and suppressed recombination within a 240-kb DNA interval on chromosome 5S (1HS) of barley. Genetics 153:1929–1948

Wei F, Wing RA, Wise RP (2002) Genome dynamics and evolution of the *Mla* (powdery mildew) resistance locus in barley. Plant Cell 14:1903–1917

Wicker T, Stein N, Albar L, Feuillet C, Schlagenhauf E, Keller B (2001) Analysis of a contiguous 211 kb sequence in diploid wheat (*Triticum monococcum* L.) reveals multiple mechanisms of genome evolution. Plant J 26:307–316

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel

P, Schulman AH (2007) A unified classification system for eukaryotic transposable elements. Nat Rev Genet 8:973–982

Williamson VM, Hussey RS (1996) Nematode pathogenesis and resistance in plants. Plant Cell 8:1735–1745

Xu Z, Wang H (2007) LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. Nucl Acids Res 35:W265–W268

Yang ZN, Ye XR, Molina J, Roose ML, Mirkov TE (2003) Sequence analysis of a 282-kilobase region surrounding the citrus tristeza virus resistance gene (Ctv) locus in *Poncirus trifoliata* L. Raf. Plant Physiol 131:482–492

Zdobnov EM, Apweiler R (2001) InterProScan—an integration platform for the signature-recognition methods in InterPro. Bioinformatics 17:847–848