Molecular and genetic analyses of potato cyst nematode resistance loci

Promotor:

prof. dr. ir. J. Bakker
Hoogleraar in de Nematologie,
in het bijzonder de fysiologie en moleculaire ecologie van nematoden
Wageningen Universiteit

Co-promotor:

dr. ir. A. Goverse
Universitair docent bij het Laboratorium voor Nematologie,
Departement Plantenwetenschappen, Wageningen Universiteit

Promotiecommissie:

prof. dr. R. G. F. Visser, Wageningen Universiteit
prof. dr. V. M. Williamson, University of California, Davis, Verenigde Staten
dr. C. Gebhardt, Max-Planck-Institut für Züchtungsforschung, Keulen, Duitsland
dr. ir. E. A. G. van der Vossen, Plant Research International, Wageningen

Molecular and genetic analyses of potato cyst nematode resistance loci

Erin Harriët Bakker

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
prof. dr. ir. L. Speelman
in het openbaar te verdedigen
op vrijdag 31 oktober 2003
des middags om halftwee in de Aula

CIP-DATA Koninklijke Bibliotheek, Den Haag

Bakker, E. H.

Molecular and genetic analyses of potato cyst nematode resistance loci

Thesis Wageningen University and Research Centre – with references – with summaries in English and Dutch
ISBN 90-5808-936-3

Keywords: *Globodera* ssp., *Solanum* ssp., disease resistance, *R* gene homologues

W. Bateson (1905): “If the Quick fund were used for the foundation of a Professorship relating to Heredity and Variation the best title would, I think, be “The Quick Professorship of the Study of Heredity”. No single word in common use quite gives this meaning. Such a word is badly wanted, and if it were desirable to coin one, “GENETICS” might do.”

Contents

Chapter 1

General introduction

Parts of this chapter have been accepted for publication in a modified form in

Erin Bakker, Robert Dees, Jaap Bakker and Aska Goverse
Mechanisms involved in plant resistance to nematodes
in: Multigenic and induced systemic resistance in plants
Edited by Sadik Tuzun and Elizabeth Bent
Kluwer Acedemic Press, New York

Potato

Potato (*Solanum tuberosum* ssp. *tuberosum* L.) is an important food crop in temperate climates. With its 300 million tons per year, potato is the fourth food crop in the world (Hawkes 1990). Potato was brought to Europe by Spanish conquistadors in the 16th century from its centre of origin, the Andean region of South America where it was domesticated about 8000 years ago (www.cipotato.org).

Since the rediscovery of genetics (Bateson 1899), plant breeders were able to predict inheritance of important agricultural traits, like for instance resistance to pathogens. This, however, was not an easy task in the case of the highly heterozygous tetraploid potato. Thanks to the development of dihaploid plants using pseudogamy (Hermsen and Verdenius 1973), potato genetics has made a major leap forward. Today, high-density linkage maps of the twelve chromosomes of potato based on molecular markers (Tanksley *et al.* 1992; www.wageningen-ur.nl/uhd) are available. This enables the scientific community to find markers closely linked to single dominant or quantitative traits that contribute to yield and quality. Using these tools, potato breeders can quickly select useful traits, introgressed from wild relatives, while minimising the size of the introgression segment. In this way the unwanted traits (linkage drag) will be minimal.

Potato cyst nematodes

A major yield limiting factor in potato are the two species of potato cyst nematodes (PCN) *Globodera rostochiensis* and *G. pallida* (Evans and Stone 1977). Infection with cyst nematodes causes symptoms like yellowing of the leaves, desiccation and starvation of the plant and perturbation of tuberisation. Worldwide, yield losses due to PCN are estimated to be 10% (Oerke *et al.* 1994). PCN was probably introduced into Europe not long after 1850 from the Andean region, where it co-evolved with potato (Jones and Jones 1974). In that period, many potato accessions were imported from South America in search of resistance to potato late blight (*Phytophthora infestans*), the principle cause of the Irish potato famine.

The life cycle of PCN starts with pre-parasitic juveniles (J2) that hatch from eggs and migrate through the soil in search of a suitable host plant. After penetration of the rhizodermis, the nematode migrates intracellularly through the root. Migration stops when a cell is encountered near the vascular bundle that is suitable as a starting point for feeding site formation. Upon feeding, the second stage parasitic juvenile develops into an adult after three moults. For cyst nematodes, sex is epigenetically determined in the first week of feeding site development (Trudgill 1967). Under favourable conditions, the nematode becomes a female and in case food is limited it becomes a male. The shape of adult females changes from vermiform to spherical and hundreds of eggs are formed inside their bodies, whereupon fertilisation by the vermiform and mobile males takes place. After the females' death, her cuticle hardens and forms a protective cyst that is very persistent in the soil. Within the egg, the first stage

juvenile (J1) moults and the resulting infective juvenile (J2) will hatch in the presence of a suitable host plant under the influence of chemicals secreted by the roots.

Cyst nematodes are able to manipulate plant cells for their own benefit. A multinuclear cell complex is formed inside the root on which the nematode fully depends for nutrition. The nematode-exploited plant cells are metabolically highly active and adapted to withdraw large quantities of nutrient solutions from the vascular system of the host plant (Jones and Northcote 1972). In incompatible potato-PCN interactions, however, feeding cell initiation and development are arrested, resulting in starvation of the nematode. Extensive microscopic observations of PCN feeding sites in plants that harbour the *H1* or the *Gpa2* resistance gene resulted in the identification of two completely different plant responses. The *H1* mechanism is characterised by a rapid hypersensitive response (HR) resulting in necrosis of the feeding site within a couple of days after infection (Rice *et al.* 1985), whereas the *Gpa2* mechanism blocks the development of the feeding cell in a late stage of the infection process (Bakker *et al.* unpublished data).

The HR, which is a common defence mechanism to a wide variety of pathogens including endoparasitic nematodes, is accompanied by an oxidative burst resulting in the production of H_2O_2 (Waetzig *et al.* 1999) and the accumulation of phenylpropanoid compounds (Robinson *et al.* 1988). In case of the *H1* gene, a small feeding cell is induced that becomes encapsulated by a layer of necrotic cells (Rice *et al.* 1985), preventing further expansion of the feeding site. This feeding site arrest occurs before the epigenetic sex determination, and some of the resulting small feeding cells are still able to provide sufficient nutrients for the development of adult males. This could explain the development of increased numbers of adult males on plants that contain the *H1* resistance locus.

The resistance response initiated by the *Gpa2* locus, is characterised by the degeneration of the feeding site around two weeks after infection and by the absence of an HR (Bakker *et al.* unpublished results). Initially, no clear differences can be observed between the compatible and the incompatible interaction. The nematode is able to establish a functional feeding site that allows the development of males and females. It is not until later stages that differences in morphology of the feeding cells are observed. In resistant plants, the proliferation of the feeding site is arrested, resulting in less dense cytoplasm and more vacuoles compared to feeding cells induced in a susceptible plant. Moreover, it is often observed that the connection between the feeding site and the vascular tissue is less pronounced. These features indicate that the metabolic activity of the feeding cell is reduced as the result of the resistance response. Finally, the cells adjacent to the feeding site become necrotic, followed by the degradation of the feeding site itself. This resistance mechanism results in the limitation of nutrients in a late stage of nematode development, after sex determination. This explains the presence of relatively large numbers of females

on plants that show such a slow defence response. In the end the food supply is insufficient, resulting in the arrest of female development and reproduction.

PCN can be controlled by crop rotation, chemical soil disinfestation and the use of resistant cultivars. However, due to the formation of cysts, PCN can survive in the soil for many years in the absence of a host, making crop rotation unattractive for potato farmers. Chemical control of PCN involves very aspecific and extremely harmful pesticides. Due to increasing concern about environmental issues and governmental regulations, this method has been practically abandoned in The Netherlands (Meerjarenplan gewasbescherming, 1989). Therefore, resistant cultivars are becoming increasingly important and hence, the scientific studies on the underlying genes and resistance mechanisms are of great interest.

Resistance to potato cyst nematodes

Over the past decades, nematode resistance has been successfully introgressed into potato. For example, the *H1* gene has, since its discovery in 1952 (Ellenby 1952), been used in many commercially available cultivars. In the United Kingdom, the potato cyst nematode *G. rostochiensis* caused enormous losses in potato yields until *H1* was successfully introgressed from the wild potato subspecies *S. tuberosum* ssp. *andigena.* Even today, after many decades of use, the gene is very effective against *G. rostochiensis*. This makes it one of the most durable resistance genes known. It is noted, however, that most *G. rostochiensis* populations in the United Kingdom have been replaced by *G. pallida,* which is virulent for *H1* (Evans 1993).

Today, many nematode *R* gene loci have been mapped in potato (reviewed by Gebhardt and Valkonen 2001). Table 1 summarises the PCN resistance loci, their genetic background and their genomic location. The quantitative trait locus (QTL) *Gro1.4* is located on chromosome III of *S. spegazzinii* and confers resistance to *G. rostochiensis* (Kreike *et al.* 1996). *Gpa4* on chromosome IV is a QTL from *S. tuberosum* ssp. *tuberosum* that confers resistance to *G. pallida*. *Gro1, Gro1.2* and *Gro1.3* on chromosomes VII, X and XI, respectively, confer resistance to *G. rostochiensis* and originate from *S. spegazzinii* (Barone *et al.* 1990; Ballvora *et al.* 1995; Kreike *et al.* 1996). On chromosome V, four resistance loci (*Gpa, Grp1, GroV1* and *H1*) have been mapped. *Gpa* is a QTL from *S. spegazzinii* and confers resistance to *G. pallida* (Bradshaw *et al.* 1998). *Grp1* is mapped in an interspecific hybrid between *S. tuberosum* and several wild species including *S. vernei, S. vernei* ssp. *ballsii, S. oplocense* and *S. tuberosum* ssp. *andigena* and therefore the exact origin of this locus is unknown. *Grp1* is a single dominant gene conferring resistance to both *G. pallida* and *G. rostochiensis* (Rouppe van der Voort *et al.* 1998b). *GroV1* and *H1* are also single dominant genes. They both confer resistance to *G. rostochiensis. GroV1* originates from *S. vernei* (Jacobs *et al.* 1996) and *H1* from *S. tuberosum* ssp. *andigena* (Pineda *et al.* 1993; Gebhardt *et al.* 1993). The single dominant resistance gene *Gpa2* has been

mapped to the distal end of the short arm of chromosome XII. The *Gpa2* gene confers resistance to the potato cyst nematode *G. pallida* and originates from the wild potato subspecies *S. tuberosum* ssp. *andigena* (Rouppe van der Voort *et al.* 1997a).

The two nematode resistance genes that have been isolated from potato, *Gpa2* (Van der Vossen *et al.* 2000) and *Gro1* (Gebhardt *et al.*, personal communication), both belong to the class of *R* genes that contain a coiled-coil domain, a nucleotide binding site and a leucine rich repeat domain (CC-NBS-LRR). The CC and NBS are thought to be involved in signal transduction, while the LRR domain is most likely responsible for *R* gene specificity (Ellis *et al.* 2000). The *Gpa2* gene has been identified by map-based cloning and is highly similar to *Rx1* (93% nucleotide identity and 88% amino acid homology). The *Rx1* gene, which confers resistance to PVX, has also been identified by map-based cloning and is located in the same cluster (Bendahmane *et al.* 1999). Interestingly, another highly similar *R* gene has been identified from *S. acaule* by the candidate gene approach (Bendahmane *et al.* 2000). This gene, *Rx2*, also confers resistance to PVX, but it is located on chromosome V. Besides *Gpa2* and *Rx1,* the complex locus on chromosome XII contains two other resistance gene homologues (RGHs): *RGH1* - a putative *R* gene with unknown specificity - and RGH3 - probably a pseudogene with a truncated effector domain (Van der Vossen *et al.* 2000).

Table 1 Overview of mapped loci that harbour resistance to the potato cyst nematodes *G. pallida* and *G. rostochiensis*

Gene	Pathogen	SD[1] or QTL[2]	Origin	Isolated	LG[3]	Reference
Gro1.4	*G. ros.*	QTL	*S. spegazzinnii*	no	III	Kreike *et al.* 1996
Gpa4	*G. pal.*	QTL	*S. tuberosum* ssp. *tuberosum*	no	IV	Bradshaw *et al.* 1998
Gpa	*G. pal.*	QTL	*S. spegazzinnii*	no	V	Kreike *et al.* 1996
Grp1	*G. pal./G. ros.*	QTL	*S. spp.*[4]	no	V	Rouppe van der Voort *et al.* 1998b
H1	*G. ros.*	SD	*S. tuberosum* ssp. *andigena*	no	V	Pineda *et al.* 1993; Gebhardt *et al.* 1993
GroV1	*G. ros.*	SD	*S. vernei*	no	V	Jacobs *et al.* 1996
GroI	*G. ros.*	SD	*S. spegazzinnii*	YES	VII	Ballvora *et al.* 1995; Barone *et al.* 1990
Gro1.2	*G. ros.*	QTL	*S. spegazzinnii*	no	X	Kreike *et al.* 1996
Gro1.3	*G. ros.*	QTL	*S. spegazzinnii*	no	XI	Kreike *et al.* 1996
Gpa2	*G. pal.*	SD	*S. tuberosum* ssp. *andigena*	YES	XII	Rouppe van der Voort *et al.* 1997a; Van der Vossen *et al.* 2000

1) SD = single dominant *R* gene
2) QTL = quantitative trait locus
3) LG = linkage group
4) Mapped in (tetraploid) clone that is an interspecific hybrid between *S. tuberosum* and several wild potato species including *S. vernei, S. vernei* ssp. *ballsii* , *S. oplocense* and *S. tuberosum* ssp. *andigena*

Outline of this thesis

Chapter 2: Two additive QTLs conferring broad-spectrum resistance in potato to *Globodera pallida* are localised on resistance gene clusters. In this chapter, it has been shown that with the use of a selected set of PCN test populations, broad spectrum resistance to *G. pallida* can be fully ascribed to the action of two loci: *Gpa5* and *Gpa6*. Both loci were readily mapped to the potato genome by use of an online catalogue of AFLP markers covering a substantial part of the potato genome (http://www.dpw.wau.nl/pv/aflp/catalog.htm) and CAPS markers based on sequenced RFLP probes. Locus *Gpa5* maps to chromosome V and explains at least 61% of the genetic variation. The *Gpa6* locus exhibits a minor effect on the resistance (24%) and acts additively to *Gpa5*. The *Gpa6* locus maps to chromosome IX.

Chapter 3: A high-resolution map of the *H1* locus, harbouring resistance to the potato cyst nematode *Globodera rostochiensis*. To employ a map-based cloning strategy, a high-resolution map of the *H1* region has been constructed. For marker saturation a bulked segregant analysis and the ultra-dense genetic map of the diploid potato clone SH have been employed. This resulted in two closely linked (< 0.1 cM) AFLP markers. One of the markers is in repulsion and the other in coupling phase. This situation will serve as a basis for the identification of the *H1* gene.

Chapter 4: Genetic and physical mapping of homologues of the virus resistance gene *Rx1* and the cyst nematode resistance gene *Gpa2* in potato. Study of the three *S. tuberosum* ssp. *tuberosum* haplotypes, homeologous to the *S. tuberosum* ssp. *andigena* haplotype that harbours the *Gpa2/Rx1* resistance gene cluster, in the two diploid potato clones (SH and RH) revealed nine additional *Gpa2/Rx1* homologues. The RGHs were identified with a specific primer pair based on conserved motifs of the LRR domain from *Gpa2* and *Rx1*. Sequence analysis of the RGHs revealed that they are highly similar to *Gpa2* and *Rx1* with sequence identities ranging from 93% to 100%. A modified AFLP method was used to facilitate the genetic mapping of the RGHs. They are all located in the *Gpa2/Rx1* cluster on chromosome XII.

Chapter 5: Structural diversity and evolutionary relationships of *Gpa2/Rx1* homologues in potato. In chapter 5, sequence analysis has been carried out on the C-terminal parts of 14 highly similar members of a family of *R* gene homologues. The family belongs to the class of CC-NBS-LRR *R* genes and includes the nematode *R* gene *Gpa2,* and the two virus *R* genes *Rx1* and *Rx2*. In this study, it is revealed that most sequence variation occurs in the LRR β-strand/β-turn motifs of the genes. These motifs are also under diversifying selection. Furthermore, a patchwork of sequence identities for both paralogues and orthologues suggests an evolutionary history involving unequal crossing-over and gene conversion.

Chapter 2

Two additive QTLs conferring broad-spectrum resistance in potato to *Globodera pallida* are localised on resistance gene clusters

This chapter is published as:

Jeroen Rouppe van der Voort, Edwin van der Vossen, Erin Bakker, Hein Overmars, Peter van Zandvoort, Ronald Hutten, René Klein Lankhorst and Jaap Bakker (2000) Two additive QTLs conferring broad-spectrum resistance in potato to *Globodera pallida* are localised on resistance gene clusters. Theoretical and Applied Genetics 101: 1122-1130

Abstract

Broad-spectrum resistance in potato to the potato cyst nematode (PCN) is commonly regarded as a complex inheriting trait. Yet, in this paper we show that by use of a selected set of PCN test populations, broad spectrum resistance to the species *Globodera pallida* can be fully ascribed to the action of two loci: *Gpa5* and *Gpa6*. These loci were readily mapped by means of a strategy based on two steps. Firstly, the chromosomal localisation of both loci was assessed by use of an online catalogue of AFLP markers covering a substantial part of the potato genome (http://www.dpw.wau.nl/pv/aflp/catalog.htm). Subsequently the chromosomal regions of both loci were identified by means of CAPS markers based on sequenced RFLP probes. Locus *Gpa5* explains at least 61% of the genetic variation. This locus maps to chromosome V on a region, which has previously been shown to harbour resistance factors to viral (*Nb, Rx2*), fungal (*R1*) and nematodal (*Gpa, Grp1*) pathogens. The *Gpa6* locus exhibits a minor effect on the resistance (24%) and acts additively to *Gpa5*. Interestingly, the *Gpa6* locus maps to a region on chromosome IX where in the homeologous tomato genome, the virus resistance gene *Sw-5* resides as part of a resistance gene cluster. In potato, resistance to potato virus X has been reported in the vicinity of this region. The map location of *Gpa6* indicates the presence of a resistance gene cluster at the end of the long arm of chromosome IX of potato.

Introduction

Quantitative resistance calls up the image of complex, polygenic inheritance patterns with multiple genes having small additive effects on the resistant phenotypes. According to Van der Plank (1982) these minor genes may be equally effective to different pathogen isolates (pathotypes) and may be relatively durable in the field. However, quantitative resistance based on several minor genes is less desirable for breeding purposes given its complexity to introgress this type of resistance in existing cultivars. Breeders rely, therefore, on major genes for resistance. Monogenic resistance is usually referred to as resistance, which complies with a gene-for-gene interaction. It contrasts to quantitative resistance by its pathotype specificity; for each major gene in the plant host there exists an avirulence gene in the pathogen (Flor 1956; Flor 1971). Major resistance (*R*) genes are often clustered in the genome and molecular analyses have shown that the gene-for-gene model holds true for several plant-pathogen systems (reviewed in Hammond-Kosack and Jones 1997). In these systems, the cloned *R* genes are thought to encode receptor molecules which perceive a pathogen (avirulence) signal whereupon a resistance response is triggered (Baker *et al.* 1997). As a result of the strong selective force of monogenic field resistance, the

site of recognition in the pathogen may be altered in a short time span. Qualitative resistance is therefore considered to be less durable than quantitative resistance.

The image of multiple minor genes underlying quantitative resistance is inconsistent with the present outcomes of most quantitative trait loci (QTL) mapping studies in different plant-pathogen combinations. Although some examples exist where quantitative resistance is explained by more than five QTLs (Geiger and Heun 1989; Bubeck *et al.* 1993; Qi *et al.* 1998; Leonards-Schippers *et al.* 1994; Wang *et al.* 1994) it is more common to find two to five loci (summarised in Young 1996) underlying complex disease resistance. In these cases, quantitative resistance is polygenic in the sense that a few genes have large effects on the phenotype. Evidence is accumulating that a wide variety of molecular and genetic mechanisms may underlie quantitative resistance. For example, insect resistance in potato is mediated by the glandular secretions of two types of leaf trichromes. QTLs associated with the resistance are co-localised with loci linked with phenol oxidase, trichrome density and sucrose ester production (Bonierbale *et al.* 1994). The expression of a resistance phenotype may also be developmentally regulated; *e.g.* while a single gene was found to control resistance to the spot blotch pathogen in the seedling stage, two QTLs were detected for resistance in the adult plant stage (Steffenson *et al.* 1996). Quantitative resistance may also fit in a classical gene-for-gene relationship. A major *R* gene product, which triggers a specific HR response upon recognition of the appropriate avirulence gene product (Baker *et al.* 1997), may have a reduced effect once a mutated version of this avirulence gene product is encountered. This phenomenon has been observed for the rice resistance gene *Xa4* which acts as a QTL against a virulent strain of the fungus *Xanthomonas oryzae* pv. *oryzae* (Li *et al.* 1999).

In the present study, quantitative resistance in potato to the potato cyst nematode *Globodera pallida* can be fully ascribed to two additive QTLs, which are likely to be explained by a gene-for-gene mechanism in which two distinct avirulence determinants in the nematode population interact (directly or indirectly) with two *R* gene products. The potato cyst nematodes (PCN) *G. rostochiensis* and *G. pallida* are serious pests in potato and especially *G. pallida* is a threat for Western-European potato culture. Broad-spectrum monogenic resistance to this species is not available and the supposed heterogeneity of the species would obstruct resistance-breeding efforts (Schnick *et al.* 1990; Folkertsma *et al.* 1996a; Folkertsma *et al.* 1996b). By use of a selected set of PCN test populations (Folkertsma 1997), quantitative resistance to *G. pallida* is dissected into the resistance loci *Gpa5* and *Gpa6*. These resistance loci map to two *R* gene clusters in potato and one of these clusters appears to be novel for the potato genome. Although it is likely that these *R* loci act in a pathotype specific way, it is argued that, as a result of the population genetics of potato cyst nematodes, the combination of *Gpa5* and *Gpa6* may confer durable resistance to *G. pallida*.

Materials and methods

Plant material
The broadness of PCN resistance was tested in potato clones AM78-3704, 3704-76 and 3704-27. The tetraploid clone AM78-3704 contains resistance to both PCN species and is an interspecific hybrid between *Solanum tuberosum* and several wild potato species including *S. vernei* 24/20, *S. vernei* ssp. *ballsii* 2/1, *S. vernei* LGU 8, *S. oplocense* EBS 1786, *S. tuberosum* ssp. *andigena* CPC 1673 and *S. spegazzinii* 440. Clones 3704-76 and 3704-27 are dihaploids (2n = 2x = 24) produced by prickle pollination of clone AM78-3704 with haploid inducer *S. phureja* clones (Hutten *et al.* 1993).

A mapping population of diploid potato was obtained from a cross between the clones 3704-76 × RH89-039-16. Clone 3704-76 is the resistant parent in our population and is referred to as JP. The susceptible male parent, clone RH89-039-16, will be referred to as RH. The mapping population F_1JP×RH consisted of 103 vigorous F_1 genotypes. Leaf material for DNA isolation was collected in the greenhouse at the seedling stage. Tubers for nematode tests were produced by first year clones on the field.

The origin of the resistance segregating from genotype JP was compared with the resistance segregating from genotype 3778-16. With the exception of *S. spegazzinii* 440, genotype 3778-16 has a similar wild species background. Locus *Grp1* conferring resistance to *G. rostochiensis* and *G. pallida* has been identified by analysing a population derived from a cross between clones 3778-16 and RH (Rouppe van der Voort *et al.* 1998b).

Nematodes
The nematode populations used in this study are listed in Table 1. *G. rostochiensis* and *G. pallida* populations are indicated by a "Ro" and "Pa" prefix respectively, followed by the pathotype designation and the code of the population. Resistance screening of the segregating potato population was carried out by use of the populations "Pa2-D383" and "Rookmaker". Population Pa2-D383 is unrelated to population Rookmaker (which is referred to as population Pa3-Rook). Population Pa3-Rook is selected out of a set of 226 *G. pallida* populations found in The Netherlands as being a representative test population for the majority of these populations (Bakker *et al.* 1992; Folkertsma *et al.* 1996a; Folkertsma *et al.* 1996b). This population appeared to contain the highest number of virulent genotypes when tested with a range of resistant clones and cultivars.

Table 1 Level of PCN resistance in the tetraploid genotype AM78-3704 and the AM78-3704 derived dihaploid clones. Presented are average numbers of cysts recovered from closed container tests.

Potato clone	Ro1-19	Ro5-22	Pa2-D383	Pa2-D350	Pa2-HPL1	Pa3-1097	Pa2-Rook	Pa3-74.768.20
AM78-3704	1	86	0	1	0	2	4	2
3704-76	6	42	0	2	0	1	2	2
3704-27	n.d.	n.d.	8	4	3	130	109	52
RH89-039-16	189	98	101	132	146	240	153	166

n.d.: not determined.

Resistance testing and data collection
Preparation of the PCN inoculum was as described (Rouppe van der Voort *et al.* 1997a). The resistance spectrum assay was performed in a closed container test (Phillips *et al.* 1980) using 125 cc plastic beakers filled with silver sand using one tuber per assay and inoculated with nematodes to a final density of 5 eggs/J_2 per gram soil. The containers were maintained in the dark at approximately 20°C for at least three months.

The inheritance of the resistance to populations Pa_2-D383 and Pa_3-Rook was analysed in mapping population F_1JP×RH in three replications. Resistant standards were cv. Multa (resistant to Pa_2-D383), *S. vernei* hybrid cv. Nika (resistant to Pa_2-D383) and cvs. Florijn and Seresta (resistant to Pa_2-D383 and Pa_3-Rook). As susceptible standard cv. Eigenheimer was used. Plant growing conditions and design of the experiment were as described (Rouppe van der Voort *et al.* 1997a). Resistance data were only recorded when at least three well-rooted plants of a single plant genotype were available.

DNA markers analysis and linkage map construction
The isolation of genomic DNA, the generation of AFLP markers and data analysis were done as described previously (Van Eck *et al.* 1995a; Vos *et al.* 1995). A genetic map was constructed only from the resistant genotype JP. This map was constructed on the basis of segregating AFLP markers which were generated by use of the AFLP primer combinations E+aaa/M+acg, E+aac/M+cac, E+aac/M+cag, E+aca/M+cac, E+aga/M+cat and E+atg/M+cta (the restriction enzymes *Eco*RI and *Mse*I, abbreviated "E" and "M" in the primer designation were used for template preparation). These primer combinations have previously been used to build up a catalogue of chromosome specific AFLP markers for a set of different potato genotypes (Rouppe van der Voort *et al.* 1997b; Rouppe van der Voort *et al.* 1998a). Alignment of the JP map was achieved by the identification of common AFLP markers between any of the potato genotypes included in the catalogue. The genetic reference maps are available from URL: http://www.dpw.wau.nl/pv/aflp/catalog.htm.

CAPS markers (cleaved amplified polymorphic sequences; Konieczny and Ausubel 1993) for potato RFLP loci GP21 and GP179 (Meksem *et al.* 1995), GP186 and TG432 (De Jong *et al.* 1997) and CT220 (Ganal *et al.* 1998) were applied to determine the chromosomal regions of the QTLs. Segregating JP alleles were detected after digestion of the amplification products using the restriction endonucleases *Dra*I for marker GP21, *Rsa*I for marker GP179 and *Nci*I for GP186 and TG432. The primer sequences for locus CT220 were derived from Folkertsma *et al.* (1999) and were designed from the published tomato cDNA sequence of tomato RFLP probe CT220 (Ganal *et al.* 1998; sequences can be retrieved via the SolGenes database (Paul *et al.* 1994), which can be accessed via http://ukcrop.net/perl/ace/search/SolGenes). Primers CT220F (aag cga att atc tgt caa c) and CT220R (gtt cct gac cat tac aaa agt ac) were used to amplify a genomic DNA fragment of approximately 220 bp from potato DNA by applying the following temperature cycle file: 3 min. 93°C followed by 30 sec. 93°C, 45 sec. 60°C, 90 sec. 72°C for 35 cycles and finished by 10 min. elongation at 72°C. The amplification product was digested with *Mse*I for identification of a segregating JP allele

The grouping of markers and determination of marker orders were calculated using the software package JoinMap 1.4 (Stam 1993). Maps were drawn by means of the graphical package Drawmap (Van Ooijen 1994).

Statistical analysis

Analysis of variance (ANOVA) were carried out on $^{10}\log(x+1)$ transformed average cyst counts per plant genotype according to the model: $\sigma^2_{tot} = \sigma^2_{plant} + \sigma^2_{rep}$. The phenotypic variance (σ^2_{tot}), genetic variance among the plant genotypes (σ^2_{plant}) and environmental variance among replications (σ^2_{rep}) were estimated using Genstat 5 release 3.2 (Payne *et al.* 1987) and the broad sense heritabilities were estimated based on the appropriate mean squares from the ANOVA. QTL analyses were performed by using the program MapQTL (Van Ooijen and Maliepaard 1996b; Van Ooijen and Maliepaard 1996a). Both a non-parametric rank-sum test of Kruskal-Wallis (see *e.g.* Sokal and Rohlf 1995) and two parametric methods; interval mapping (Lander and Botstein 1989) and multiple-QTL mapping (MQM) (Jansen 1993; Jansen and Stam 1994) were applied. These procedures are implemented in the MapQTL software. Threshold values for assigning a QTL to a map position are $P < 0.001$ for the Kruskal Wallis test and a LOD score of 3.0 for interval and MQM mapping. The percentage of the total variance explained by a marker genotype (R^2) was calculated by $(\sigma^2_m - \sigma^2_{res})/ \sigma^2_m$ where σ^2_m represents the variance between the two marker genotype classes (*i.e.* presence or absence of a particular marker) and σ^2_{res} the residual variance.

Results

Broadness of PCN resistance from AM78-3704

The level of PCN resistance present in the tetraploid genotype AM78-3704 and two AM78-3704 derived dihaploid clones 3704-76 and 3704-27 was assessed by a resistance screen with a diverse set of PCN populations. These PCN populations are known to harbour differences in molecular as well as virulence characteristics (Folkertsma *et al.* 1996a; Folkertsma 1997). Comparison of the tetraploid genotype with the dihaploid clones showed that resistance against *G. pallida* has been retained in clone 3704-76 (Table 1). However, loss of resistance to the Pa3 populations was observed in clone 3704-27. No resistance to *G. rostochiensis* Ro_5-22 was found. The diploid clone RH89-039-16 was susceptible to all populations tested. From these data it was decided to analyse a segregating population derived from a cross between 3704-76 × RH89-039-16 (F_1JP×RH). Since large differences exist in molecular and virulence characteristics of *G. pallida* populations Pa2-D383 and Pa3-Rook, these populations were used for further resistance testing.

Inheritance of resistance

The average numbers of cysts developed on the parental genotypes as assessed in pot tests for *G. pallida* populations Pa2-D383 and Pa3-Rook are shown in Table 2. The average number of newly developed cysts in the progeny ranged from 0 to 1361 for population Pa2-D383 and from 9 to 1047 for population Pa3-Rook. For population Pa2-D383 susceptible plants could be clearly distinguished from plants assigned to the resistant class. The numbers of cysts developed on resistant plants varied from 0 to 42, whereas these numbers ranged from 202 to 1361 on the susceptible plants.

The Pa3-Rook cyst numbers showed a continuous distribution among the plant genotypes; no clear distinction between a resistant and susceptible class could be made. Analysis of variance on normalised cyst counts showed significant differences in the level of resistance among the plant genotypes ($0.001 < P < 0.0001$). No significant differences in cyst numbers were found among the blocks of replicates. The broad sense heritability, listed in Table 2, indicated that the variation in the numbers of cysts developed was genetically determined. Skewness and kurtosis values were –0.68 and 0.07 respectively showing that the assumption on normally distributed $^{10}\log(x+1)$ transformed cyst numbers was justified (Snedecor and Cochran 1967).

Table 2 Results of the quantitative analysis of PCN resistance as measured by counted cyst numbers of the respective PCN populations. Presented are the heritability, the most flanking marker to the resistance loci, the test values indicating the appropriate map position and the R^2 at the QTL position.

Trait	#cysts[a] JP	# cysts RH	h^{2b}	Marker[c]	Chrom. Location[d]	P value[e]	R^{2f}	LOD[g]
Pa2-D383	1	856	-	TG432	Chr. 5	< 0.0001	-	21
Pa3-Rook	23	635	0.91	TG432	Chr. 5	< 0.0001	61%	12
				CT220	Chr. 9	< 0.001	24%	3

[a] # cysts: average number of cysts, [b] h^2 : heritability, [c] Marker : nearest marker to resistance locus, [d] Chrom. Location: chromosomal localisation of resistance locus, [e] *P* value of the nearest marker in a Kruskal Wallis test, [f] R^2 : percentage of the total variance explained by the marker genotypes, [g] LOD: logarithm of odds score which indicates the likelihood that two loci are linked (Lander and Botstein 1989).

AFLP mapping

A genetic map of genotype JP was produced in order to find association between molecular markers and the resistance trait(s). The AFLP map of genotype JP consisted of 149 AFLP markers and spanned a total map length of 603 cM. The sizes of the linkage groups varied between 13 cM (for chromosome IV) and 69 cM (for chromosome V). Comparison of the positions of common AFLP markers present on the JP map and the maps of five other potato genotypes revealed that parts of the maps of chromosomes I, IV, VIII, X and XI were not fully covered with AFLP markers. It was therefore anticipated that QTLs may be unresolved as a result of the relative incompleteness of the JP map.

Mapping of Pa2-D383 resistance

Linkage analysis showed a significant association between the resistance to Pa2-D383 and AFLP markers which were localised at the top of chromosome V. This map position is well known for its presence of a complex *R* gene cluster in potato harbouring resistance to different pathogen species (summarised in Leister *et al.* 1996). Therefore CAPS markers GP186, GP21, TG432 and GP179, residing at the short arm of potato chromosome V, were tested in population F_1JP×RH for a more precise mapping of this locus. Three (GP186, GP21 and TG432) of the four markers segregated in F_1JPxRH. Marker TG432 showed the highest linkage with the resistance locus (recombination fraction = 3/102; LOD = 21). As two different resistance loci to *G. pallida* have previously been assigned to this map position (*Gpa* (Kreike *et al.* 1994) and *Grp1* (Rouppe van der Voort *et al.* 1998b)) the origin of the resistance was further investigated.

Mapping of Pa3-Rook resistance

The presence of the TG432 marker allele in the plant genotypes appeared to be associated with a reduced number of Pa3-Rook cysts also. A total of 32 out of 42 plant genotypes which harboured the TG432 allele contained < 100 cysts whereas among the remaining 61 genotypes (without the TG432 allele) 58 genotypes were

found which contained > 100 cysts (Fig. 1). Subsequent analysis using MapQTL software revealed a significant effect of marker TG432 on the resistance; both in the Kruskal Wallis test and by interval mapping, the highest significance levels were found at marker locus TG432 (Table 2, Fig. 2). This QTL explained about 61% of the total variance, which is in the same order of magnitude as the amount of variation in Pa3-Rook resistance explained by locus *Grp1* (Rouppe van der Voort *et al.* 1998b). Interestingly, the LOD profiles for Pa3-Rook resistance of *Grp1* and the present QTL coincide, indicating that similar gene(s) underlie *G. pallida* resistance in genotypes 3778-16 and JP. However, *Grp1* differs from the JP QTL at chromosome V as it confers resistance to both species of the potato cyst nematode. In analogy to the nomenclature of other *G. pallida* resistance loci in potato (Rouppe van der Voort *et al.* 1997a; Kreike *et al.* 1994) we propose to name the QTL on chromosome 5, *Gpa5*.

An additional effect on Pa3-Rook resistance was found at AFLP markers on the long arm of chromosome IX. The position of this locus could be more precisely defined by application of the most distal RFLP marker CT220 at this chromosomal region as a CAPS marker. Locus CT220 yielded the highest significance level for Pa3-Rook resistance (Fig. 3). Since CT220 is localised at the end of a linkage group only one neighbouring marker was used in the interval mapping. The addition of more neighbouring markers (to a maximum of 5) in the analysis showed an increase in the LOD score to a maximum of 9.1 at this end of chromosome IX but resulted in a less precise localisation of the QTL.

The QTL at chromosome IX explained 24% of the total variance (Table 2) and acts addititively with *Gpa5*. The additive effect of the chromosome V and chromosome IX loci is shown in Fig. 1. A significant reduction in the numbers of Pa3-Rook cysts was observed in the class of plants harbouring the TG432 and CT220 alleles as compared to the three other marker genotype classes. In the TG432/CT220 class, one plant genotype was found with an average number of 550 cysts. This plant most likely lacks *Gpa5* as a result of a recombination event in the TG432 – E+aaa/M+acg-244.0 region. The respective plant genotype contains all JP marker alleles in the GP186 – TG432 region and lacks all subsequent JP alleles onwards which are in coupling with *Gpa5* resistance. Since no resistance specificity to PCN species has been assigned to chromosome IX, we propose to name the QTL at chromosome IX *Gpa6*.

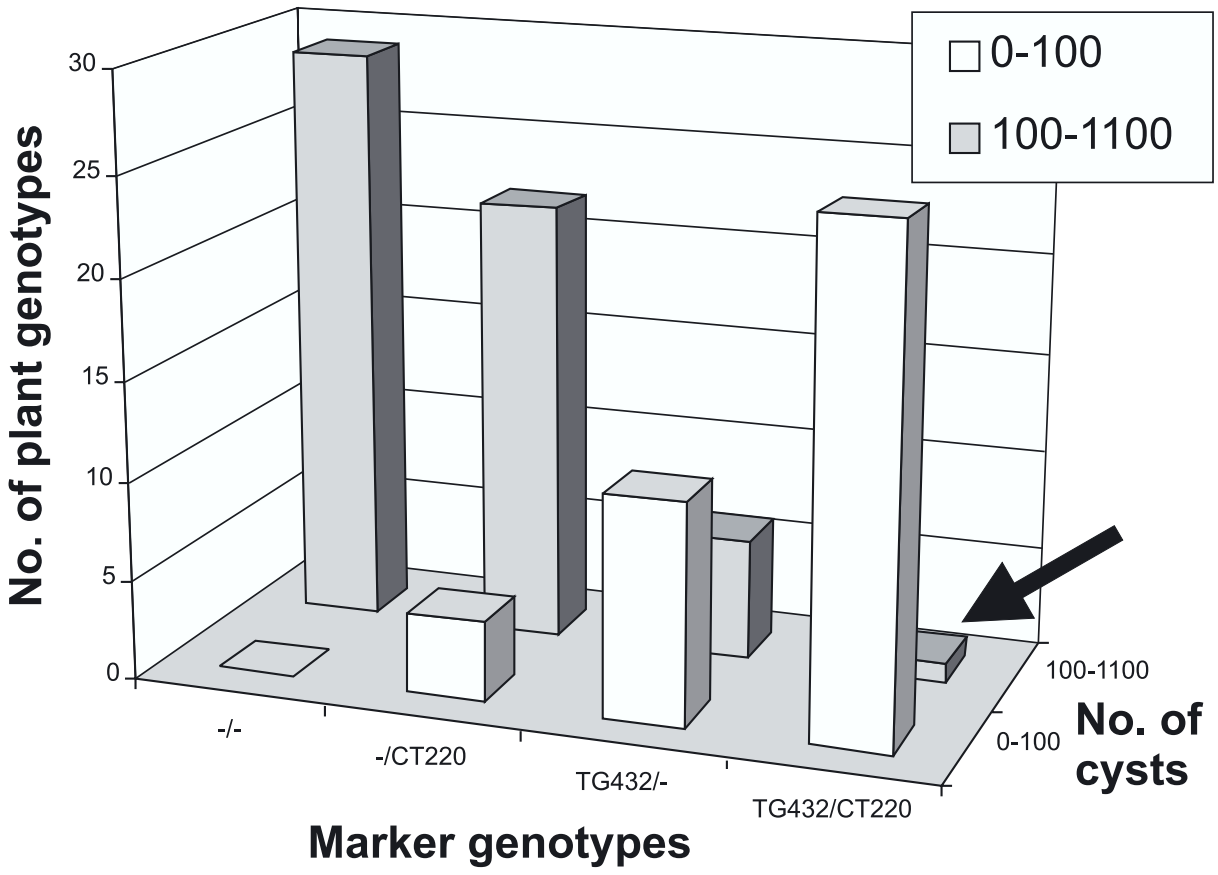


Figure 1 Resistance to Pa$_3$-Rook is conferred by the combined action of *Gpa5* and *Gpa6*. Shown is the distribution of the number of *G. pallida* Pa$_3$-Rook cysts recovered within the TG432/CT220 offspring classes TG432/CT220 (presence of both marker alleles), TG432/- (presence of the TG432 allele, absence of the CT220 allele), -/CT220 (absence of the TG432 allele, presence of the CT220 allele) and -/- (absence of both marker alleles). The genotype indicated by an arrow most likely harbours a recombination event between TG432 and *Gpa5* as this genotype lacks all subsequent distal marker alleles in coupling with the TG432 allele.

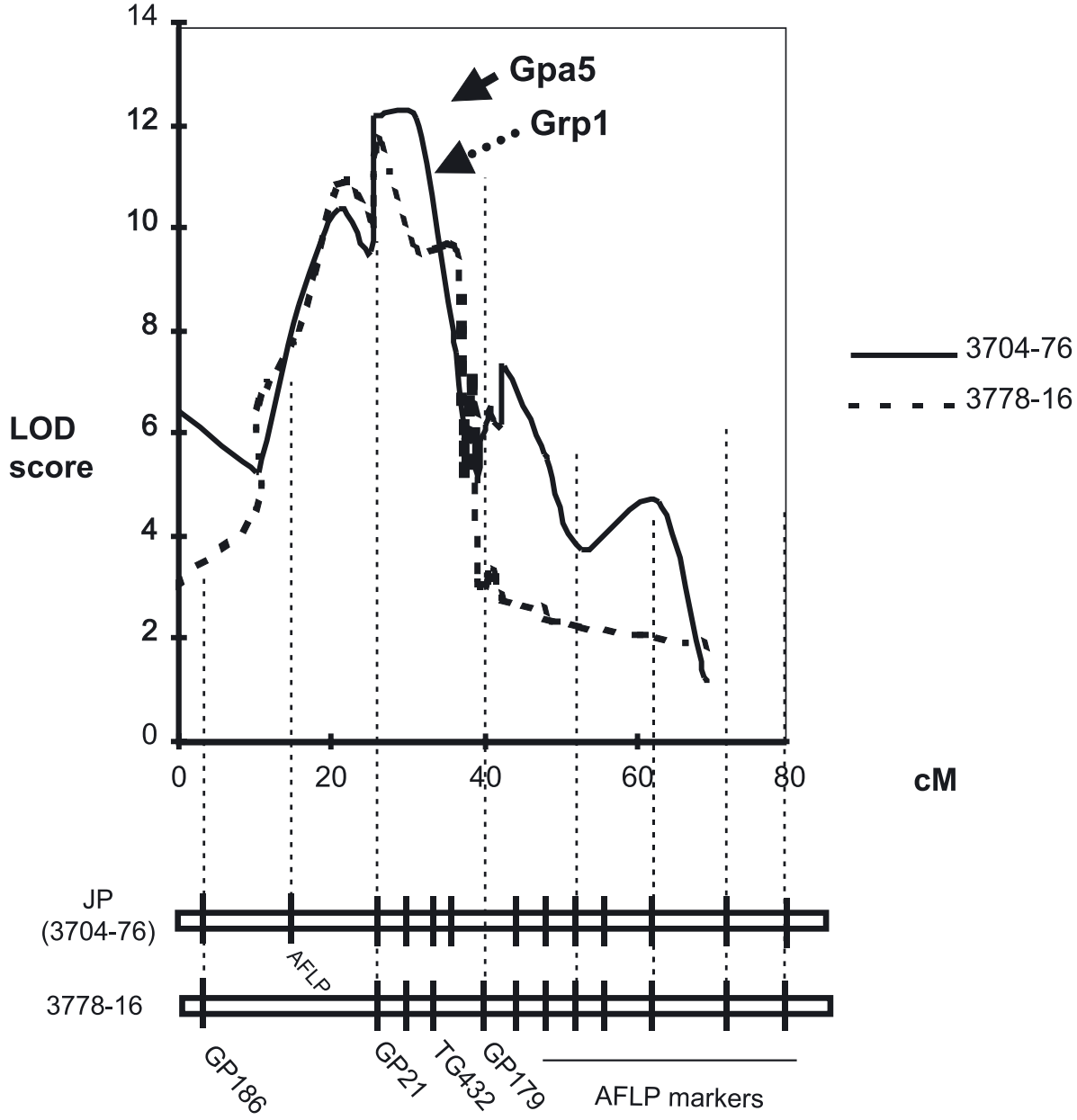


Figure 2 LOD plots for *G. pallida* Pa3-Rook resistance on chromosome V in genotypes JP and 3778-16, which show that the effect on Pa3-Rook resistance is localised on the same map region. To the Pa3-Rook resistance locus on chromosome V segregating from clone JP is referred to as *Gpa5* (see text). The PCN resistance in 3778-16 is ascribed to the action of locus *Grp1*. For reasons of simplicity, only schematic maps are drawn. More details on the AFLP markers mapped on chromosome V are presented in Fig. 4.

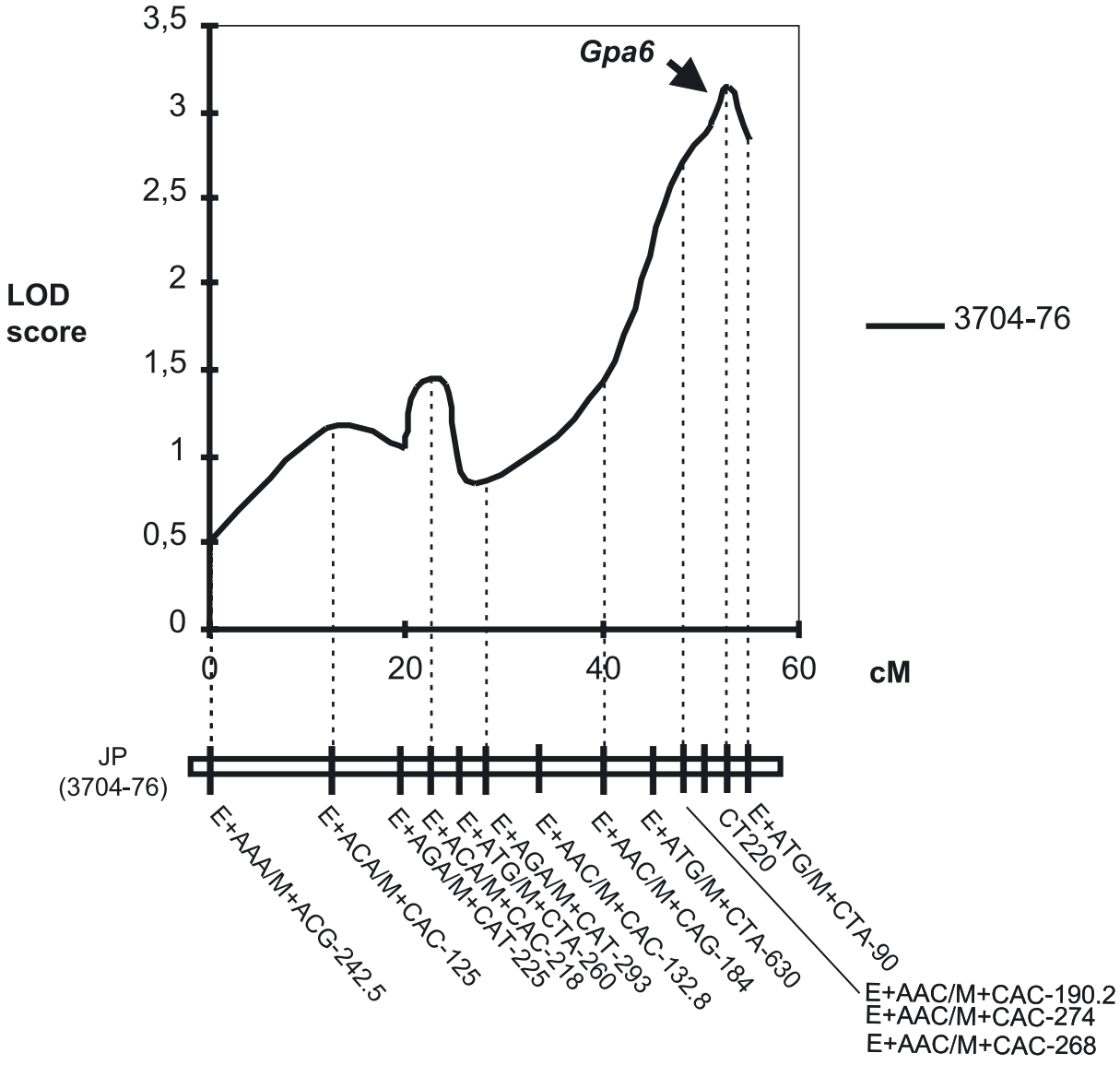


Figure 3 LOD plot for *G. pallida* Pa3-Rook resistance on chromosome IX. To this resistance locus is referred as *Gpa6* (see text).

Gpa5 may be part of the Grp1 resistance locus

To address the question whether the same gene(s) underlie the *G. pallida* resistance conferred by *Gpa5* and *Grp1*, we compared the occurrence of common marker alleles linked in coupling with the QTLs at chromosomes V of clones JP and 3778-16. The result of this analysis is shown in Fig. 4. At the short arm of chromosome V, identical marker alleles were identified for CAPS loci GP186, GP21, TG432 and AFLP loci E+atg/M+cta-350, E+aga/M+cat-88.9. This similarity is lost from marker GP179 onwards. Marker GP179 and all subsequent AFLP markers, which were identified in clone 3778-16 were either absent in JP or present in repulsion with *Gpa5*. The occurrence of two sets of common marker alleles (intervals GP21-TG432 and E+aca/M+cac-320 – E+aac/M+cac-250.1) indicates the presence of similar introgression segments in clones JP and 3778-16. Within these intervals, the DNA sequences between both clones may be highly similar. Since *G. pallida* resistance has been an important trait for selection to generate both breeding lines we argue that the same gene(s) underlie the *G. pallida* resistance. However, the linkage between TG432 and GP179 and all subsequent marker alleles on the 3778-16 chromosome harbouring *Grp1* may have been lost during the course of the breeding process of clone JP. We therefore hypothesise that the recombination between loci TG432 and GP179 in a progenitor of clone JP has been associated with loss of *G. rostochiensis* resistance in clone JP.

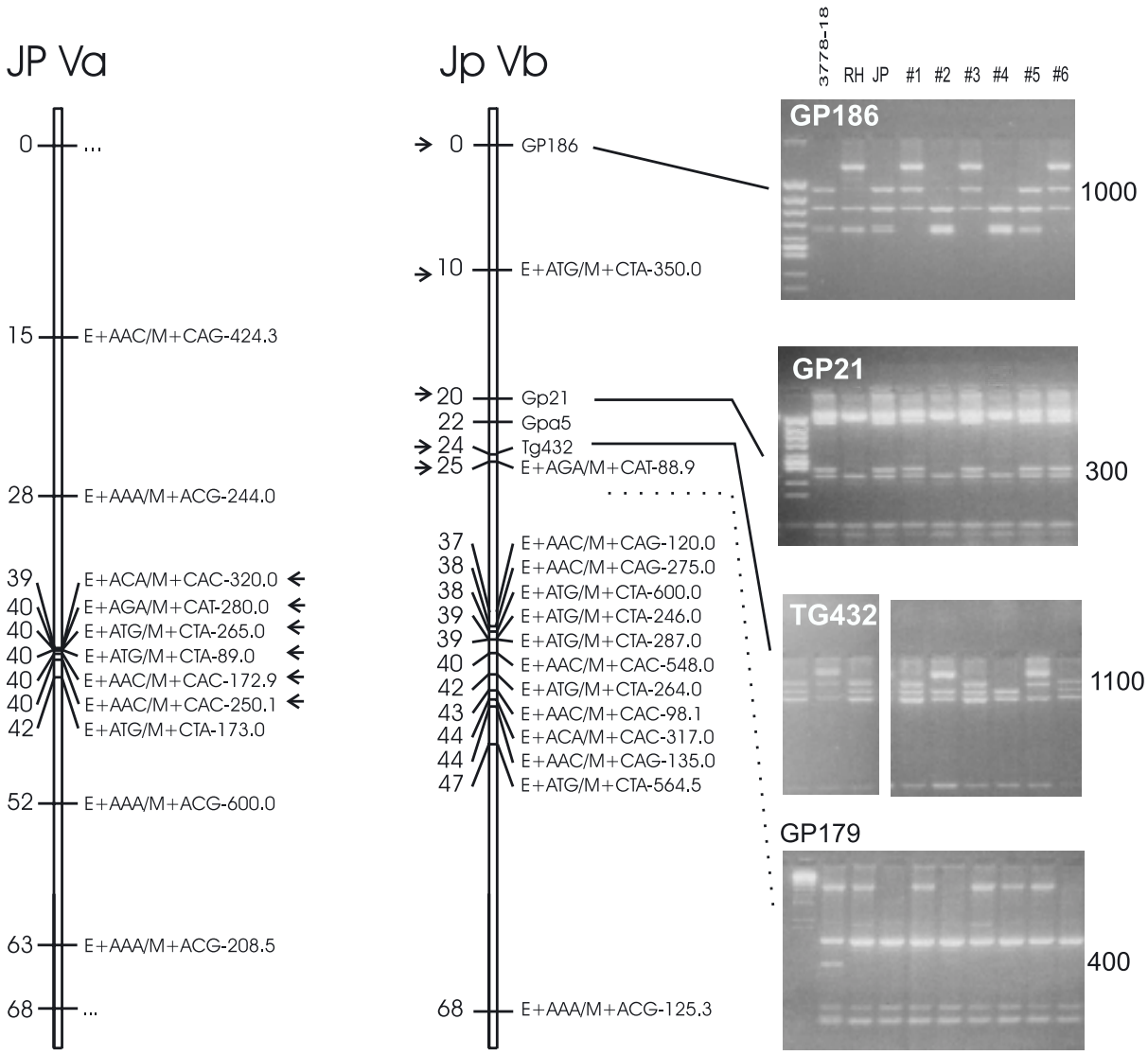


Figure 4 Comparison of the presence of common marker alleles between clones JP and 3778-16 indicates that *Gpa5* is likely to be introgressed by a common ancestor of both parental clones. Shown are the maps of the two homologous chromosomes V of clone JP, which harbour linkage blocks of marker alleles, which are also present in 3778-16 (indicated by arrow heads). The comparative analysis is exemplified by the profiles of the CAPS markers localised on top of chromosome V. Next to a marker lane, amplification products are loaded of clones 3778-16, RH, JP and a series of six progeny clones. Markers GP186, GP21 and TG432 show the similar profiles of clones 3778-16 and JP. For marker GP179 (presented is the profile which contains the mapped 3778-16 allele which is identified after *Rsa*I digestion) no segregating JP allele is identified even after the use of a series of 20 restriction enzymes. The corresponding map position of locus GP179 on the potato map is indicated by a dashed line. The numbers the right indicate the molecular weights (in bp) of the segregating alleles.

Discussion

This study has shown that the map position of two resistance loci *Gpa5* and *Gpa6* coincides with two resistance gene clusters in potato. *Gpa5* maps close to marker TG432 on a region where resistance to potato virus X (*Rx2*; Ritter *et al.* 1991, *Nb;* De Jong *et al.* 1997) and to the fungus *Phytophthora infestans* (*R1;* Leonards-Schippers *et al.* 1992, *Pi-QTL*; Leonards-Schippers *et al.* 1994) have been mapped. In addition, two resistance loci to PCN are located in this region. *Gpa* which acts against *G. pallida* (Kreike *et al.* 1994) and *Grp1* which confers resistance to both PCN species (Rouppe van der Voort *et al.* 1998b) reside within the GP21 – GP179 marker interval. This prompted us to the question whether the resistance mapped in clone JP could be ascribed to either of these loci. However, combining various data revealed that *Gpa5* is different from *Gpa* and *Grp1*. Locus *Gpa* has been mapped in an intraspecific population of the wild potato species *S. spegazzinii* and this wild species is not included in the pedigree of the resistant genotype JP. Moreover, the specificity of *Gpa* extends to that of *Gpa5*; *Gpa* confers complete resistance to population Pa_3-Rook (Wolters, personal communication) whereas 61% of the Pa_3-Rook resistance is explained by *Gpa5*. Locus *Gpa5* can also be distinguished from *Grp1* as it confers only partial resistance to *G. pallida*. Since the level of resistance of both loci is in the same order of magnitude, it may be possible that the *G. pallida* resistance is mediated by the same gene(s). This hypothesis is corroborated by the fact that the pedigrees of the parental clones JP and 3778-16 share the major part of the progenitors, which have been used to introgress wild species resistance. In addition, it has previously been argued that *Grp1* is a compound locus since *G. rostochiensis* and *G. pallida* diverged millions of years ago (Bakker *et al.* 1992) and the multiple specificity of *Grp1* is unlikely to be the result of the action of a single gene (Rouppe van der Voort *et al.* 1998b). Circumstantial evidence for the occurrence of *Gpa5* in the genome of 3778-16 was found by a comparative analysis of the maps of both genotypes. On the basis of common marker alleles within a 5 cM interval at the short arm of chromosome V (Fig. 4), it can be argued that the same gene(s) underlie the *G. pallida* resistance. Recombination in the TG432 – GP179 interval during the process of generating clone JP may thus be associated with loss of *G. rostochiensis* resistance in clone JP.

Locus *Gpa6* explains only a minor part of the genetic variation of the resistance to population Pa_3-Rook. However, the finding of *Gpa6* is significant because the combined action of *Gpa5* and *Gpa6* results in the same level of resistance as has been transmitted by clone JP to the progeny. *Gpa6* is not able to resist population Pa_2-D383 which indicates that the resistance at *Gpa6* acts in a pathotype specific way. The pathotype specificity of *Gpa6* indeed suggests that a major gene(s) underlie the resistance of *Gpa6*. The involvement of a major gene(s) in *Gpa6* resistance is further corroborated by the map location of this locus. *Gpa6* maps close to marker CT220, an RFLP marker locus that targets a resistance gene cluster in tomato containing the *Sw-*

5 resistance gene in tomato (Brommonschenkel and Tanksley 1997; Folkertsma *et al.* 1999). The *Sw-5* gene confers broad spectrum resistance to tomato spotted wilt virus and belongs to the class of resistance genes which encodes a putative gene product harbouring a leucine zipper (LZ), a nucleotide binding site (NBS) and a leucine rich region (LRR) (Brommonschenkel *et al.* 2000; Folkertsma *et al.* 1999). The gene cluster to which the *Sw-5* gene belongs is comprised of at least 5 resistance gene analogues (Folkertsma *et al.* 1999; Folkertsma and Prins personal communication) and the location of *Gpa6* suggests the presence of a similar cluster on the homeologous potato genome.

On the potato genome, no additional resistance specificities have been mapped in the vicinity of marker CT220. The closest resistance locus to marker CT220 is the Nx_{phu} locus which confers resistance to potato virus X (Tommiska *et al.* 1998). This locus maps at 11 cM from marker CT220 and the map positions of these loci should therefore be considered different. The question whether any molecular relationship exists between *Gpa6*, Nx_{phu} and *Sw-5* can only be resolved when more information is available on the genes, which underlie *Gpa6* and Nx_{phu} resistance.

The partial resistance phenotype of *Gpa5* and *Gpa6* does not conflict with the hypothesis that major genes are responsible for both resistance specificities considered that the *G. pallida* test population Pa_3-Rook is likely to be heterogeneous at the respective (a)virulence loci. In case avirulence alleles are not fixed in a population, a single *R* gene, which operates on the basis of a gene-for-gene relationship, will confer partial resistance against the population as a whole whereas on the level of the individual, the *R* gene will confer absolute resistance against the matching avirulent genotype. The potato cyst nematode is an obligate outbreeding species and field populations, such as Pa_3-Rook, are often composed of a mixture of virulent and avirulent genotypes (Bakker *et al.* 1993). Formal proof for the occurrence of a gene-for-gene relationship has so far only been demonstrated for the interaction between *G. rostochiensis* and the *H1* gene from *S. tuberosum* ssp. *andigena* CPC1673 (Janssen *et al.* 1991), but evidence is accumulating that various other PCN resistance genes operate on the same basis. For example, the *Gpa2* gene which is part of a resistance gene cluster at chromosome XII of potato (Kanyuka *et al.* 1999) and which encodes a LZ-NBS-LRR type of protein (Van der Vossen *et al.* 2000) confers specific resistance against population Pa_2-D383. This population is on the molecular level clearly separated from all other *G. pallida* populations, which are found in the Netherlands (Folkertsma *et al.* 1996a; Folkertsma *et al.* 1996b). Also the *H2* gene (Dunnet 1961) which map location is presently unknown, confers resistance against a set of *G. pallida* populations pathotyped "Pa_1" which are well defined at the molecular level (Blok *et al.* 1997; Folkertsma 1997).

The availability of broad-spectrum resistance against *G. pallida* is of practical importance. Combining *e.g. Gpa6* with *Grp1* brings complete resistance against both

PCN species into focus. Compared to major gene resistance against various other pathogen species, it is expected that resistance against potato cyst nematodes may be relatively durable. PCN are endemic in the Andean region of South-America and are thought to be introduced into Europe in the 19th century. Only a limited part of the genetic variation (and thus the variation at virulence loci also) present in their centre of origin has been introduced. Because the multiplication rate of PCN is low, their active spread is limited and the time between generations is often 2-4 years in a normal crop rotation, it seems highly unlikely that the broad spectrum resistance obtained by combining *Gpa5* and *Grp1* is broken down rapidly.

In summary, this study reports on the QTL mapping of the resistance loci *Gpa5* and *Gpa6* in potato. These QTLs coincide with the map location of other resistance factors, which are known to be involved in gene-for-gene relationships. On the basis of the occurrence of common motifs in previously cloned *R* genes, it is hypothesised that *R* genes are evolutionary related components of a recognition system and that the mechanisms generating variation are entailed with the clustered distribution on the genome (Baker *et al.* 1997). Recent advances in the identification and cloning of resistance gene candidates (RGCs) has shown that these type of sequences abundantly occur in plant genomes (Botella *et al.* 1997; Shen *et al.* 1998). With the aid of the increasing amount of sequence information different RGCs in the genome will be discriminated and resistance specificities will be assigned to the genes. This may eventually lead to a detailed dissection of the resistance gene clusters harbouring *Gpa5* and *Gpa6* and may confirm our hypothesis that these two additive QTLs operate on the basis of a classical gene-for-gene relationship.

Chapter 3

A high-resolution map of the *H1* locus harbouring resistance to the potato cyst nematode *Globodera rostochiensis*.

This chapter has been submitted for publication as:

Erin Bakker, Ute Achenbach, Jeroen Bakker, Joke van Vliet, Johan Peleman, Bart Segers, Stefan van der Heijden, Piet van der Linde, Herman van Eck, Edwin van der Vossen, Jaap Bakker and Aska Goverse
A high-resolution map of the *H1* locus harbouring resistance to the potato cyst nematode *Globodera rostochiensis*. Submitted to Theoretical and Applied Genetics

Abstract

The resistance gene *H1* confers resistance to the potato cyst nematode *Globodera rostochiensis* and is located at the distal end of the long arm of chromosome V of potato. For marker enrichment of the *H1* locus a bulked segregant analysis (BSA) was carried out using 704 AFLP primer combinations. A second source of markers tightly linked to *H1* is the ultra-high density (UHD) genetic map of the potato cross SH×RH. This map has been produced with 387 AFLP primer combinations and consists of 10,365 AFLP markers in 1,118 BINs (www.dpw.wageningen-ur.nl/uhd/). Comparing these two methods revealed that BSA resulted in one marker/cM and the UHD map in four markers/cM in the *H1* interval. Subsequently, a high-resolution genetic map of the *H1* locus has been developed using a segregating F1SH×RH population consisting of 1209 genotypes. Two PCR based markers were designed at either side of the *H1* gene to screen the 1209 genotypes for recombination events. In the high-resolution genetic map, two of the four co-segregating AFLP markers could be separated from the *H1* gene. Marker EM1 is located at a distance of 0,2 cM and marker EM14 is located at a distance of 0,8 cM. The other two co-segregating markers CM1 (in coupling) and EM15 (in repulsion) could not be separated from *H1*.

Introduction

The potato cyst nematode (PCN) species *Globodera rostochiensis* and *G. pallida* cause serious yield losses in potato crops worldwide (Ross 1986). An effective and durable way to control PCN is the use of resistant potato cultivars. Twelve PCN resistance loci have been mapped in potato on chromosomes III, IV, V, VII, IX, X, XI and XII (reviewed by Gebhardt and Valkonen 2001). Eight resistance traits (*Gro1.4, Gpa4, Gpa, Gpa5, Grp1, Gpa6, Gro1.2* and *Gro1.3*) confer partial resistance, while four of them (*H1, GroV1, Gro1 and Gpa2*) confer absolute resistance. Interestingly, many PCN resistance loci map to a region where at least one other (and often more) single dominantly inherited resistance gene is present (so called hot spots for resistance). This holds not only for the single dominantly inherited PCN resistance genes *Gpa2, Gro1* and *H1*, but also for quantitative trait loci (QTL) like *Grp1*, *Gpa3*, *Gpa5* and *Gpa6*. So far, the only nematode resistance genes that have been identified in potato are *Gpa2* and *Gro1* (Van der Vossen *et al.* 2000; Gebhardt, personal communication).

H1 resistance has been discovered in 1952 in *Solanum tuberosum* ssp. *andigena* in the Common Wealth Potato collection (Ellenby 1952). Since then, the *H1* gene has been introgressed in many commercially available cultivars. Even today, after many decades of use, the gene is very effective against *G. rostochiensis* in the UK. This makes it one of the most durable resistance genes known (Evans 1993). The *H1* gene is the only nematode resistance gene with proof for a gene-for-gene interaction

(Janssen *et al.* 1991) and confers resistance to *G. rostochiensis* pathotypes Ro_1 and Ro_4 (reviewed by Jones *et al.* 1981) by the activation of an HR. The feeding site that is initiated by the nematode becomes encapsulated by a layer of necrotic cells and degenerates in the course of a week (Rice *et al.* 1985). For cyst nematodes, sex is epigenetically determined in this first week of feeding site development, depending on the amount of food available (Trudgill 1967). Therefore, the *H1* resistance response, results in starvation of the nematodes, while a limited number of individuals will develop into males.

The *H1* locus has been mapped in potato on the distal end of the long arm of chromosome V and is closely linked to RFLP markers CP113 and CD78 (Gebhardt *et al.* 1993; Pineda *et al.* 1993). In the same chromosomal region of *S. vernei* another PCN resistance gene has been mapped (*GroV1*) that might be allelic to the *H1* locus (Jacobs *et al.* 1996). In this study, a combination of bulked segregant analysis and the ultra dense genetic map of potato has been used to identify markers closely linked to the *H1* gene. CAPS markers $CT51_{CAPS}$ and $239E4left_{CAPS}$ have subsequently been used to screen for recombination events in a population of 1209 genotypes segregating for the *H1* locus. Finally, AFLP markers and nematode resistance assays have been used to construct a high-resolution map of the *H1* locus. The markers EM1 and CM1 are in coupling with and tightly linked to the *H1* gene (0,2 and 0 cM respectively) and will form the basis for a positional cloning strategy.

Materials and methods

Plant material
A mapping population of 136 F1 genotypes from the cross between the diploid potato clones SH83-92-488 (SH)×RH89-039-16 (RH) was available (Rouppe van der Voort *et al.* 1997a). Additional offspring extended the population to a total of 1209 genotypes. The female parent (SH) harbours the *H1* locus introgressed from *S. tuberosum* ssp. *andigena* CPC 1673 and the male parent (RH) is fully susceptible to all *G. rostochiensis* populations tested. A second population was produced comprising 120 F1 progeny from the cross between the diploid potato clones DH84-25-2389 (DH; resistant) and KW84-16-2396 (KW; susceptible). 45 potato cultivars (Table 1) were used to test markers for marker-assisted selection (MAS).

DNA extraction
Genomic DNA from SH, RH and progeny was extracted from the plants as described (Stewart and Via 1993). The method was adjusted for 96-well format using 1 ml tubes (Micronic BV, Lelystad, The Netherlands), supplemented with two steel balls (ø 2 mm). Samples were homogenised using a Retsch 300 mm shaker at maximum speed

(Retsch BV, Ochten, The Netherlands). The genomic DNA of 45 potato cultivars, DH, KW and progeny was isolated as described (Van der Beek *et al.* 1992).

Bulked segregant and marker analysis
For bulked segregant analysis (BSA), 704 *Eco*RI/*Mse*I primer combinations with 3 selective nucleotides at each primer were performed on bulks containing DNA of 10 susceptible F1 genotypes and bulks containing DNA of 10 resistant F1 genotypes derived from the cross DH×KW. AFLP analysis for BSA and for the markers derived from the ultra high-density (UHD) genetic map of SH (Table 2) was performed according to Vos *et al.* (1995). The UHD map and the AFLP primer combinations used to produce the map are available at www.dpw.wageningen-ur.nl/uhd/

Two CAPS markers (Konieczny and Ausubel 1993) were designed. $CT51_{CAPS}$ has been derived from the sequenced RFLP probe CT51 (Tanksley *et al.* 1992). The sequence could be retrieved from the SolGenes database (Paul *et al.* 1994), which can be accessed via http://ukcrop.net/perl/ace/search/SolGenes. Marker $239E4left_{CAPS}$ has been derived from the sequenced left border of SH BAC clone 239E4 (kindly provided by E. Coppoolse). The DNA sequences of the PCR primers and the corresponding thermal cycling conditions are presented in Table 3.

Resistance assays
PCN resistance assays were performed on plants derived from stem cuttings, tubers or *in vitro* stocks. The assays were performed as described (Rouppe van der Voort *et al.* 1997a; Rouppe van der Voort *et al.* 1999). Two potato cyst nematode lines, homozygous in their virulence character, have been used in the *H1* resistance assays in the SH×RH population. *G. rostochiensis* line 19 is avirulent on potato genotypes harbouring the *H1* gene, while line 22 is virulent (Janssen *et al.* 1990). For the resistance assay of the DH×KW population, the Ro_1 field population A57 was used. Plants with <5 cysts are considered to be resistant.

Table 1 Eleven candidate markers resulting from the BSA are tested on 45 potato cultivars for co-segregation with the *H1* locus. Status S is susceptible to G. rostochiensis pathotype Ro_1 and status R is resistant to G. rostochiensis pathotype Ro_1 (+ = present, – = absent, n.t. = not tested).

Cultivar	Status	EM1	EM2	EM3	EM4	EM5	EM6	EM7	EM8	EM10	EM11	EM12
Bartina	S	-	-	-	-	-	-	-	-	-	-	-
Concurrent	S	-	-	-	+	-	-	n.t.	-	-	-	-
Lizen	S	-	-	-	+	+	n.t.	n.t.	-	-	-	-
Mansour	S	-	-	-	-	+	n.t.	-	-	-	-	-
Starletta	S	-	-	-	+	+	n.t.	-	-	n.t.	-	n.t.
Alpha	S	-	-	-	-	-	-	-	-	-	-	-
Bea	S	-	-	-	-	+	-	-	-	-	-	-
Blanka	S	-	-	-	-	-	-	-	-	-	-	-
Bintje	S	-	-	-	-	-	-	-	n.t.	-	-	-
Burmania	S	-	-	-	-	-	-	-	-	-	-	-
Cleopatra	S	-	-	-	-	+	n.t.	-	-	-	-	-
Eigenheimer	S	-	-	-	+	-	-	-	n.t.	-	-	n.t.
Furore	S	-	-	-	-	-	-	-	n.t.	-	-	-
Humalda	S	-	-	-	-	+	-	-	-	-	-	+
Saskia	S	-	-	-	-	+	-	-	-	-	-	+
Baraka	S	-	-	-	-	+	-	-	-	-	-	-
Desiree	S	-	-	-	+	-	-	-	-	-	-	-
Edzina	S	-	-	-	-	-	-	-	-	-	-	n.t.
Favorita	S	-	-	-	-	+	-	-	-	-	-	-
Joakla	S	-	-	-	+	+	-	-	-	-	-	n.t.
Manna	S	-	-	-	-	+	-	-	-	-	-	n.t.
Aphrodite	S	-	-	-	+	+	-	-	-	-	-	n.t.
Monalisa	S	-	-	-	-	+	n.t.	-	-	-	-	-
Sirtema	S	-	-	-	+	n.t.	-	-	-	-	-	+
Provita	R	+	+	+	-	-	n.t.	-	+	-	+	-
Fresco	R	+	+	+	-	-	-	-	-	-	+	-
Premiere	R	+	+	+	-	-	-	-	+	-	+	+
Revelino	R	+	+	+	+	+	+	+	+	+	+	+
Amigo	R	+	+	+	+	+	+	+	+	+	+	n.t.
Anosta	R	+	+	+	+	+	+	+	+	+	+	+
Prevalent	R	+	+	+	-	-	-	-	-	-	+	-
Prior	R	+	+	+	-	+	n.t.	-	-	-	+	-
Producent	R	+	+	+	-	+	-	-	-	-	+	-
Asterix	R	+	+	+	+	+	+	+	+	+	+	+
Cardinal	R	+	+	+	-	-	-	-	-	-	+	+
Lutetia	R	+	+	+	-	+	n.t.	-	-	-	+	+
Marijke	R	+	+	+	+	-	-	-	-	-	+	+
Obelix	R	+	+	+	-	-	-	-	-	-	+	+
Sante	R	+	+	+	+	+	+	+	+	+	+	+
Ukama	R	+	+	+	+	+	-	-	-	-	+	+
Van Gogh	R	+	+	+	+	+	+	+	+	-	+	+
Vebeca	R	+	-	-	+	+	+	+	-	-	-	+
Sinaeda	R	+	+	n.t.	+	+	+	+	+	+	n.t.	+

Table 2 AFLP markers derived from the UHD genetic map of potato. Marker names, UHD map marker names, restriction enzymes, selective nucleotides and band sizes are indicated.

Marker	UHD map	Restriction enzymes		Selective nucleotides		Band size
CM1	Caga/Mcac_233S5	*Sac*I	*Mse*I	aga	cac	233 bp
EM1	Eatg/Mgca_40S5	*Eco*RI	*Mse*I	atg	gca	152 bp
EM14	Eacc/Maac_16	*Eco*RI	*Mse*I	acc	aac	Not sized
EM15	Eagt/Mcac_98S5	*Eco*RI	*Mse*I	agt	cac	98 bp
EM16	Eacc/Maag_22	*Eco*RI	*Mse*I	acc	aag	Not sized

Table 3 Primer sequences, thermal cycling conditions and the polymorphic restriction site for CAPS markers used to screen for recombination events in the *H1* region of potato.

Marker	Primers	PCR conditions	Restriction enzyme
$CT51_{CAPS}$	5'-gca gga ttc cat ttg ctt gc 5'-gtt att gtc taa cca cct cgg	94ºC 3 min 94ºC 30 s, 51ºC 30 s, 72ºC 120 s 35 cycles 72 ºC 5 min	*Alu*I
$239E4left_{CAPS}$	5'-ggc ccc aca aac aag aaa ac 5'-agg tac ctc cat ctc cat ttt gta ag	94ºC 3 min 94ºC 30 s, 51ºC 30 s, 72ºC 90 s 35 cycles 72 ºC 5 min	*Alu*I

Results

Markers linked to H1 generated by bulked segregant analysis

The 120 F1 progeny of DH×KW were tested for resistance to *G. rostochiensis* pathotype Ro_1 resulting in 79 susceptible and 41 resistant genotypes. BSA was carried out on pools of 10 susceptible and 10 resistant genotypes. Testing 704 AFLP primer combinations resulted in 13 candidate markers, which were subsequently tested on the individual genotypes of the DH×KW population. This resulted in, respectively, 6 markers (EM1 to EM6) located at a distance of 0.95 cM, 4 markers (EM7 to EM10) at 1.9 cM, 2 markers (EM11 and EM12) at 2.9 cM and 1 marker (EM 13) at 6.3 cM from the *H1* locus. The 13 candidate markers, their extensions, the band sizes and their distance to *H1* are presented in Table 4.

Table 4 Candidate markers derived from BSA on pools of susceptible and resistant F1 genotypes of the DH×KW population. The marker names, selective nucleotides, sizes and distances to the *H1* locus are indicated.

Marker	Selective nucleotides	Size (bp)	Distance to *H1* (cM)
EM1	atg/gca	152	0.95
EM2	act/cgt	137	0.95
EM3	atg/cac	132	0.95
EM4	aag/gag	239	0.95
EM5	ccc/cac	98	0.95
EM6	caa/cct	87	0.95
EM7	acg/gga	53	1.9
EM8	aga/cac	233	1.9
EM9	aga/cgc	273	1.9
EM10	caa/caa	115	1.9
EM11	atg/cca	234	2.9
EM12	aac/cca	121	2.9
EM13	aca/caa	132	6.3

In addition, 11 out of 13 candidate markers (EM1 to EM8 and EM10 to EM12) were tested on 45 potato cultivars. Table 1 shows the presence or absence of a marker and susceptibility or resistance of each potato cultivar for nematode resistance to *G. rostochiensis* pathotype Ro_1. Markers EM1 to EM6 have the same distance to the *H1* locus in the DH×KW population, but EM1 is the only marker that co-segregates with nematode resistance in all potato cultivars. Therefore, it is concluded that marker EM1 is more closely linked to the *H1* gene.

Chromosomal position of the H1 gene in SH

Forty-seven F1 genotypes, randomly selected from the original mapping population F1SH×RH, were screened for resistance to the potato cyst nematode *G. rostochiensis* line 19. Twenty plants were resistant and twenty-seven were susceptible. The results of the resistance assay were integrated into the UHD genetic map (www.dpw.wageningen-ur.nl/uhd/), resulting in the mapping of the *H1* gene at the distal end of the long arm of chromosome V, where also the markers CP113 and CD78 are located. In the UHD mapping population, the *H1* gene co-segregates with four AFLP markers (Table 2). Markers EM1, EM14 and CM1 are in coupling with *H1*, whilst EM15 is in repulsion. EM16 flanks the *H1* locus at the telomeric side and is separated by one recombination event (Fig. 1).

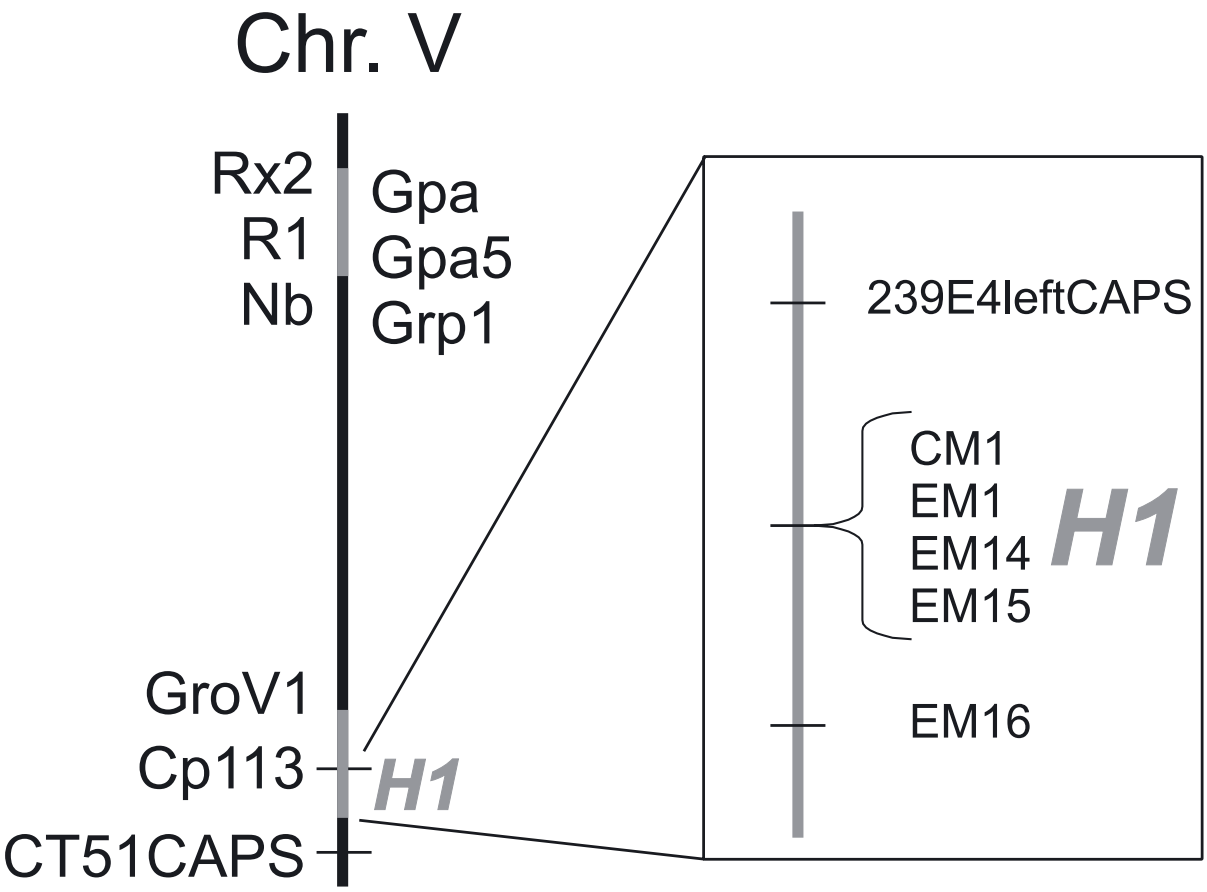


Figure 1 Schematic representation of chromosome V of potato, which contains two *R* gene regions. *Rx2* (Ritter *et al.* 1991), *R1* (Leonards-Schippers *et al.* 1992), *Nb* (DeJong *et al.* 1997), *Gpa* (Kreike *et al.* 1994), *Gpa5* (Rouppe van der Voort *et al.* 2000), and *Grp1* (Rouppe van der Voort *et al.* 1998b) are located at the short arm of chromosome V, while *GroV1* (Jacobs *et al.* 1996) and *H1* (Gebhardt *et al.* 1993; Pineda *et al.* 1993) are located at the long arm of chromosome V. The box at the right contains a close-up of the *H1* locus in the UHD genetic map. The *H1* gene co-segregates with EM1, EM14, EM15 and CM1 and is flanked by EM16 and 239E4left$_{CAPS}$.

The high-resolution map

To fine-map the *H1* locus, a population of 1073 plants was screened for recombination events using CAPS markers. At the telomeric side of the *H1* gene CAPS marker CT51$_{CAPS}$ (Fig. 2A) was designed, which was located at a distance of 10 cM from the *H1* gene. Instead of CAPS marker CP113 (Niewohner *et al.* 1995), an alternative CAPS marker 239E4left$_{CAPS}$ (Fig. 2B) was designed that mapped to a position 0.8 cM from the *H1* gene at the centromeric side.

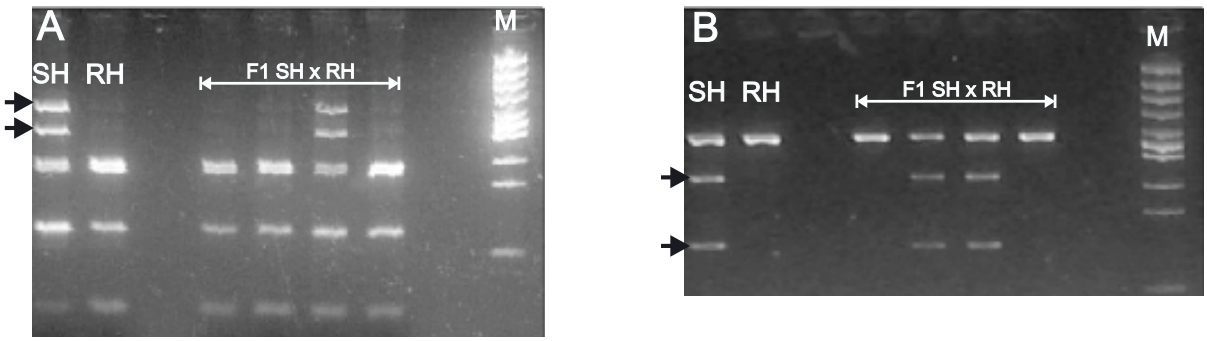


Figure 2 Performance of the CAPS markers $CT51_{CAPS}$ (A) and $239E4left_{CAPS}$ (B) on the resistant parent (SH), the susceptible parent (RH) and 4 descendants. Digestion of the PCR product with *Alu*I resulted for both markers in two polymorphic bands, segregating from the resistant parent as indicated with arrows.

Screening the 1073 F1SH×RH genotypes resulted in 129 plants showing a recombination event between the PCR-based markers $CT51_{CAPS}$ and $239E4left_{CAPS}$. Screening these 129 genotypes with EM16, which flanks the *H1* interval, resulted in fifty-six genotypes with a recombination event between this marker and $239E4left_{CAPS}$. To map these recombination events more precisely, a third marker screening was carried out with the four AFLP markers that co-segregate with *H1* in the UHD population. For nine genotypes, AFLP analysis resulted in missing values and they were omitted from further analysis. Subsequently, a successful resistance assay on forty genotypes resulted in the identification of twenty-three susceptible and seventeen resistant plants. Integrating the results of the marker screening and the resistance assay resulted in a high-resolution map as shown in Figure 3. In a total of 1209 offspring, there are no recombination events between the *H1* gene and the markers CM1 (in coupling) and EM15 (in repulsion), demonstrating complete linkage of these two markers. Furthermore, the *H1* gene is separated from EM1, EM14 and $239E4left_{CAPS}$ by 2, 8 and 31 recombination events respectively, equating to 0.2, 0.8 and 3,0 cM distance.

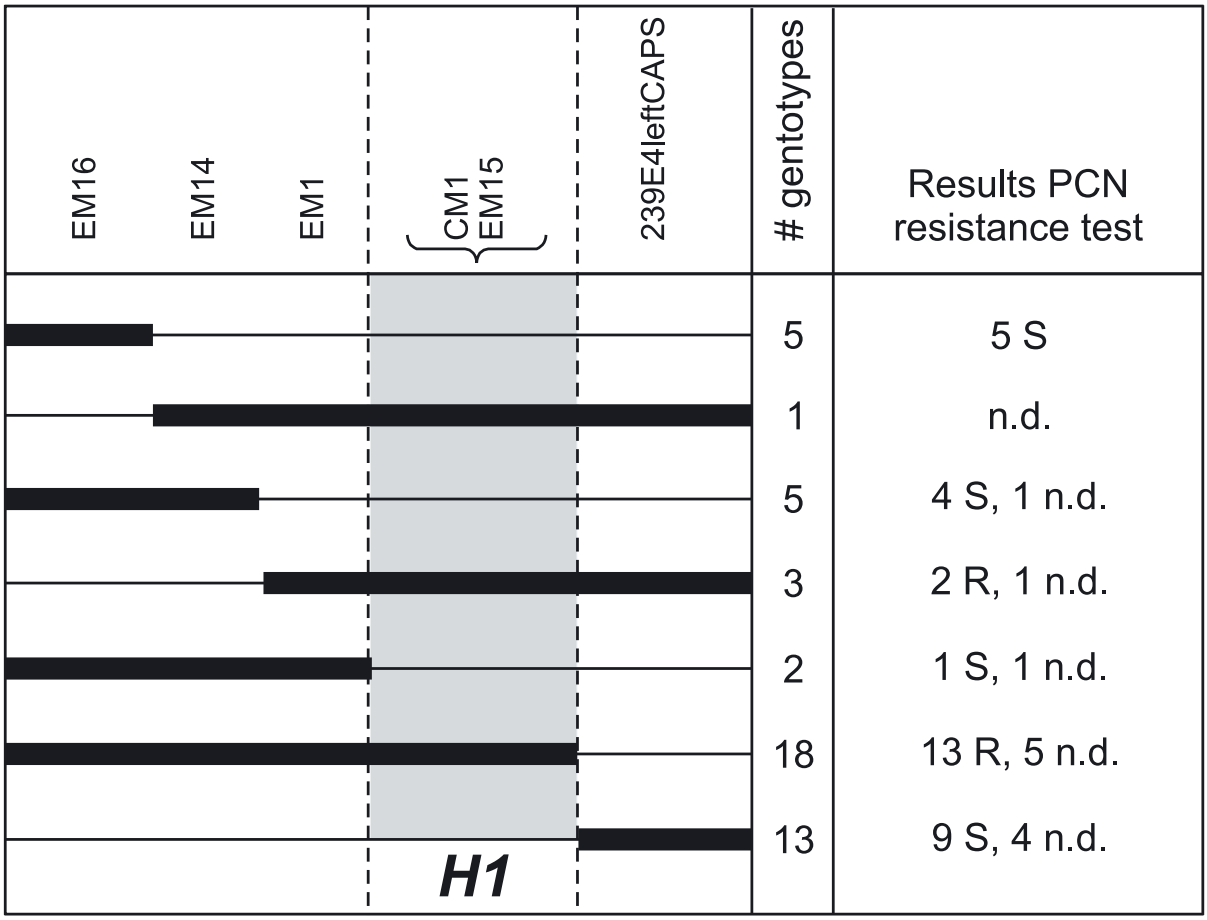


Figure 3 The high-resolution map of the *H1* locus in potato. Markers EM1, EM14, EM16, CM1 and 239E4left$_{CAPS}$ (in coupling) and EM15 (in repulsion) are presented in the correct genetic order. Markers EM14, EM1, CM1 and EM15 co-segregated in the population used to produce the UHD map. Horizontal lines depicted in bold represent chromosomal regions derived from the resistant haplotype of SH, thin horizontal lines represent chromosomal regions derived from the susceptible haplotype of SH. The region were the *H1* gene is genetically located is indicated by a grey column between dotted lines. In the second column, the number of genotypes for each recombination event is given and in the third column, the results of the PCN resistance assays are listed (n.d. = not done, S = susceptible, R = resistant).

Discussion

In this study, a high-resolution map of the *H1* locus has been made in potato linking the *H1* gene with the AFLP markers CM1 (coupling) and EM15 (repulsion). To identify markers closely linked to the *H1* gene, a BSA has been carried out using 704 AFLP primer combinations. In addition, markers closely linked to the *H1* locus have been selected from the UHD genetic map of potato that was produced with 387 AFLP

primer combinations (*www.dpw.wageningen-ur.nl/uhd/*). Of the markers identified in the BSA, only one marker could be mapped at a distance of less than 1 cM from the *H1* gene in the high-resolution map, while 4 markers derived from the UHD map are located at a distance of less than 1 cM from the *H1* gene. Apparently, with BSA some markers linked to the *H1* gene are not recognised. The UHD genetic map of potato is a globally saturated map and will therefore be suitable for mapping purposes of traits located anywhere in the potato genome. With the high-resolution map of *H1* we have shown the usefulness of a globally saturated UHD genetic map for high-resolution mapping.

The majority of the 129 recombination events between the markers EM1 and 239E4left$_{CAPS}$ were observed between *H1* and the marker 239E4left$_{CAPS}$. This could be explained by a large physical distance, or by the presence of a so-called "hot spot" for recombination. In case of a "hot spot" for recombination, a small physical distance results in a large genetic distance. This phenomenon has been reported for resistance gene loci previously. For instance, the tomato resistance gene loci *Asc* (Mesbah et al. 1999) and *I2* (Segal et al. 1992) measure 125 kb/cM and 43 kb/cM, respectively. These physical distances are much shorter than the average distance of 750 kb/cM for the tomato genome (Tanksley et al. 1992). A physical map of the *H1* locus is needed to draw further conclusions on this issue.

Marker assisted selection (MAS) can be very useful to select for interesting traits. In case of resistance to potato cyst nematodes, the selection can already take place at the seedling stage and laborious nematode resistance tests can be avoided. Of course, for successful MAS, it is essential to have markers at a very short distance from the gene of interest. Marker EM1 has been tested in 45 potato cultivars and always co-segregated with resistance. This marker could be a suitable candidate for MAS. In the high-resolution genetic map, however, EM1 is not the closest marker. Marker CM1 co-segregates with resistance in a progeny of 1209 genotypes and in theory should be a better candidate. Markers EM1 and CM1 are AFLP markers and for MAS, cheaper and easy to handle CAPS markers are preferred. Therefore, these markers should first be converted to CAPS markers before they can be used in MAS.

Forty-five potato cultivars were selected to test the candidate markers resulting from the BSA for their suitability for MAS. The cultivars were especially selected for resistance derived from S. tuberosum ssp. andigena (*H1*). The marker CP113 is also closely linked to *H1* and this marker has been tested on twenty resistant and thirty-three susceptible cultivars (Gebhardt et al., 1993). Marker CP113, however, could not be detected in any of the resistant cultivars and the most probable explanation was that these cultivars harbour GroV1 (derived from S. vernei) instead of *H1* (Gebhardt et al., 1993). Cultivar "Producent" was tested both for marker CP113 and for marker EM1. This cultivar harbours resistance derived from S. tuberosum ssp. andigena and was tested positive for marker EM1, while marker CP113 was absent. A possible

explanation for this could be that marker CP113 is farther removed from the *H1* gene than marker EM1.

In this study, six genotypes have been tested for nematode resistance in vitro as well as with tubers (data not shown). The results of both resistance assays were identical for all genotypes. Testing for nematode resistance in vitro is quicker than testing with tubers. The resistance response of the *H1* triggers an HR already within the first week of nematode infection (Rice et al. 1985). In addition, virulent and avirulent G. rostochiensis lines were used to perform the nematode resistance tests, reducing the number of nematodes that are able to develop into adult females on a resistant genotype to approximately 1 in 600 (Janssen et al. 1990). Therefore, the results of the in vitro assay can be observed already at three weeks after inoculation, when numerous developing females emerge on the roots of susceptible plant, while the results of a tuber assay are ready at ten to twelve weeks after inoculation. Furthermore, the in vitro assay can be performed on shoots derived from a seedling, while at least four tubers are required for a reliable tuber assay. Therefore, we conclude that the faster in vitro assay is a reliable method to test genotypes for the presence or absence of the *H1* gene.

The development of the high-resolution genetic map is a crucial step in the positional cloning of the *H1* gene. Marker CM1 is both located at a distance of less than 0.1 cM from the *H1* locus and in coupling phase. Therefore, this is a suitable marker for screening an SH BAC library for the construction of a physical map and subsequent cloning of the *H1* gene. The gene-for-gene interaction between the *H1* gene and G. rostochiensis (Janssen et al. 1991) will provide us with a useful model system to study the molecular mechanisms underlying cyst nematode resistance.

Chapter 4

Genetic and physical mapping of homologues of the virus resistance gene *Rx1* and the cyst nematode resistance gene *Gpa2* in potato

This chapter has been published as:

Erin Bakker, Patrick Butterbach, Jeroen Rouppe van der Voort, Edwin van der Vossen, Joke van Vliet, Jaap Bakker and Aska Goverse (2003)
Genetic and physical mapping of homologues of the virus resistance gene *Rx1* and the cyst nematode resistance gene *Gpa2* in potato. Theoretical and Applied Genetics 106: 1524-1531

Abstract

Nine resistance gene homologues (RGHs) were identified in two diploid potato clones (SH and RH) with a specific primer pair based on conserved motifs in the LRR domain of the potato cyst nematode resistance gene *Gpa2* and the potato virus X resistance gene *Rx1*. A modified AFLP method was used to facilitate the genetic mapping of the RGHs in the four haplotypes under investigation. All nine RGHs appeared to be located in the *Gpa2/Rx1* cluster on chromosome XII. Construction of a physical map using bacterial artificial chromosome (BAC) clones for both the *Solanum tuberosum* ssp. *tuberosum* and the *S. tuberosum* ssp. *andigena* haplotype of SH showed that the RGHs are located within a stretch of less than 200 kb. Sequence analysis of the RGHs revealed that they are highly similar (93 to 95%) to *Gpa2* and *Rx1*. The sequence identities among all RGHs range from 85 to 100%. Two pairs of RGHs are identical or nearly so (100 and 99,9%), with each member located in a different genotype. Southern blot analysis on genomic DNA revealed no evidence for additional homologues outside the *Gpa2/Rx1* cluster on chromosome XII.

Introduction

Plants are constantly under attack by a wide range of pathogens and pests. To defend themselves, plants have evolved an innate surveillance system encoded by a large set of resistance genes. Most resistance genes are single dominant and confer resistance in a gene-for-gene specific manner (Flor 1971). More than 30 resistance genes (*R* genes) have been cloned from different plant species and they can be divided into four classes based on common structural motives (Takken and Joosten 2000). The majority of the *R* genes are characterised by a leucine rich repeat (LRR) and a nucleotide binding site (NBS) domain. *R* genes that belong to this superfamily confer resistance to completely unrelated taxonomic groups like bacteria, fungi, viruses and nematodes (Mindrinos *et al.* 1994; Whitham *et al.* 1994; Lawrence *et al.* 1995; Vos *et al.* 1998; Milligan *et al.* 1998; Bendahmane *et al.* 1999; Van der Vossen *et al.* 2000).

In potato, nineteen *R* genes have been mapped to eleven chromosomal regions (Gebhardt and Valkonen 2001). Five of them, *Rx1*, *Gpa2*, *Rx2* and *R1* and *Gro1,* have been isolated and they all belong to the NBS-LRR class. *Rx1* and *Gpa2* originate from *Solanum tuberosum* ssp. *andigena* and have been identified by map based cloning (Bendahmane *et al.* 1999; Van der Vossen *et al.* 2000). *Rx2* has been isolated from *S. aucaule* using a PCR based approach (Bendahmane *et al.* 2000). *R1* and *Gro1* have been isolated from *S. demissum* and *S. spegazzinni* respectively by using a combination of map based cloning and the candidate gene approach (Ballvora *et al.* 2002; Gebhardt peronal communication). *Rx1* and *Gpa2* are highly homologous, yet they confer resistance to two completely unrelated pathogens *viz.* potato virus X and

the potato cyst nematode *Globodera pallida*. The genes are tightly linked on chromosome XII of potato (Bendahmane *et al.*, 1997; Rouppe van der Voort *et al.*, 1999). Sequencing a 187 kb region revealed that *Gpa2* and *Rx1* are part of a complex locus containing at least two other closely related resistance gene homologues (RGHs): *RGC1* and the pseudogene *RGC3* (Bendahmane *et al.* 1997; Rouppe van der Voort *et al.* 1999; Bendahmane *et al.* 1999; Van der Vossen *et al.* 2000). In the LRR domain, more variation is observed between the RGHs than in the NBS domain. The mean K_a/K_s ratio for the LRR region is larger than one, whereas in the NBS regions the ratio is smaller than one, indicating that the LRR domain is subject to diversifying selection and that specificity is determined by this domain (Van der Vossen *et al.* 2000).

In this paper we describe the use of LRR-specific primer combinations for both the identification and mapping of nine *Gpa2/Rx1* homologues in two diploid potato clones (SH and RH). The value of using LRR-specific primers for the dissection of *R* gene clusters in different haplotypes will be discussed.

Materials and methods

Plant material and DNA extraction
A mapping population for *S. tuberosum* ssp. *tuberosum* was available from the diploid potato clones SH83-92-488×RH89-039-16 (Rouppe van der Voort *et al.* 1997a). The female parent SH83-92-488 contains an introgression segment originating from the wild accession *S. tuberosum* ssp. *andigena* CPC1673 on which the *Globodera pallida* resistance gene *Gpa2* and the PVX resistance gene *Rx1* are located. The male parent RH89-039-16 has been selected for its fertility and the production of vigorous offspring. SH83-92-488 will be referred to as SH and RH89-039-16 will be referred to as RH. The mapping population F_1SH×RH consists of 136 vigorous F1 genotypes. Genomic DNA was extracted from frozen leaf tissue of *in vitro* plants as described (Van der Beek *et al.* 1992).

BAC library and DNA extraction
The construction of the BAC library from the diploid potato clone SH83-92-488 has been described (Rouppe van der Voort *et al.* 1999). The library has been extended with 30,000 clones and now comprises 90,000 clones. Pooling and preparation for screening was done as described (Kanyuka *et al.* 1999) and resulted in 255 plate pools. DNA extracted from these BAC clones was used as template for PCR and sequence analysis. For this, clones were cultured overnight at 37°C in 500 ml LB medium supplemented with 170 μg/ml chloramphenicol for selection. Plasmid DNA was isolated using the "very low copy plasmid DNA purification protocol" of the plasmid midi kit according to manufacturers instructions (Quiagen, Hilden,

Germany). For each BAC, 1 µl of DNA was digested with *Hind*III to check the concentration and the purity.

PCR analysis
A cluster-specific primer pair was designed based on conserved DNA sequences in the LRR domains of *Gpa2*, *Rx1* and *SH-RGH1* (previously designated as *RGC1* (Van der Vossen *et al.* 2000)). The forward primer (LRR-F: ttg gtg tcg taa cag tga gg) starts at position +1533 of *Gpa2* and the reverse primer (LRR-R: ctg gct agt cct cag aac ac) at position +3192 of *Gpa2* (UTR). A PCR was performed with this primer pair using the Expand High Fidelity PCR System according to manufacturer instructions (Roche, Mannheim, Germany). The following PCR cycle file was applied: 3 min 94°C followed by 30 s 94°C, 30 s 55°C, 90 s 72°C for 10 cycles, 30 s 94°C, 30 s 55°C, 90 s 72°C with an extension of 5 s/cycle for 25 cycles and finally 5 min elongation at 72°C.

Cloning and digestion analysis of the PCR products
PCR products were cloned into the PCR2.1TOPO-vector for transformation of *E. coli* TOP10 cells according to the manufacturer instructions (Invitrogen, San Diego, CA, USA). For each PCR product, 12 positive clones were cultivated overnight at 37°C in 2 ml liquid LB medium with 100 µg/ml ampicillin for selection. Each clone was reamplified with the same LRR primer pair and the resulting PCR products were digested with *Taq*I. PCR was performed using an adjusted PCR buffer containing 100 mM TrisHCl pH 8.3, 500 mM KCl, 25 mM MgCl, 10% Triton X-100 to avoid additional cleaning steps of the PCR products prior to digestion. The DNA fragments were separated on a 4% agarose gel consisting of 1% ultra pure agarose (Life Technologies, Breda, The Netherlands) and 3% NuSieve ® GTG ® agarose (FMC, Philadelphia, PA, USA) in 1× TAE buffer at 120 V. Clones were selected for further analysis based on differences in the digestion patterns.

Sequence analysis
Sequence analysis of the PCR products was carried out by Greenomics, PRI, Wageningen, The Netherlands. BAC-end sequencing was performed using approximately 1 µg template DNA in a cycle sequencing reaction using either 100 ng sp6 or t7 primer, 8 µl Big Dye terminator mix (PerkinElmer, Wellesley, MA, USA) in a total volume of 20 µl. The PCR protocol consisted of 25 cycles of 30 s 96°C, 15 s 50°C and 4 min 60°C. After ethanol precipitation at room temperature for 10 min and recovery, the labelled DNA was dissolved in 3 µl formamide. DNA was then heated for 2 min at 96°C and directly cooled on ice. Approximately 1.5 µl was then loaded on a 6% TBE (pH 8.3) polyacrylamide gel. Sequence electrophoresis was carried out on either an ABI 373XL or ABI 377 sequencer. DNA sequence analysis and comparisons were carried out using DNAstar™ software.

CHEF gel electrophoresis
BAC insert sizes were determined with CHEF electrophoresis on a 1% agarose gel (Seakem ® Gold, FMC, Philadelphia, PA, USA) in 0.5×TBE buffer at 4°C using a BIO-RAD CHEF DR II system (Bio-Rad Laboratories, Hercules, CA, USA) at 200 V with a pulse time of 5 to 15 s for 18 h.

Southern blot analysis
Approximately 20 ng BAC DNA and 3 µg genomic DNA was digested with *Hind*III. The digested DNA was separated on a 1% agarose (Agarose NA, Pharmacia, Peapack, NJ, USA) gel in 1×TAE buffer at 50 V. The gels were blotted overnight on hybridisation filters using capillary forces (Sambrook *et al.* 1989). Southern analysis was performed based on the DIG Application Manual for filter hybridisation (Roche, Basel, Switzerland). An LRR based probe (position +1533 to position +1936 (403 bp)) and an NBS and LRR based probe (position +1367 to position +1552 (185 bp)), were amplified from *Gpa2* BAC DNA. Primer pairs used are PLRR-F (ttg gtg tcg taa cag tga gg) and PLRR-R (gtt ctc tgt agg ctc atg ac) at an annealing temperature of 60°C, and PNBSLRR-F (gtg gaa tgc atg atg tga cc) and PNBSLRR-R (ctc act gtt acg aca cca ag) at an annealing temperature of 55°C. The results were visualised on an autoradiogram after 5 min up to 2 hours exposure.

Genetic mapping
Mapping was performed using a modified method (Fig. 1) based on the AFLP technique (Vos *et al.* 1995). The template used was the PCR product generated from gDNA using the iPM4 primers as described (Kanyuka *et al.* 1999). The PCR product was digested with *Taq*I and a *Taq*I adapter as described (Vos *et al.* 1995) was ligated to the digestion products. This was followed by fragment amplification on 1:25 diluted template using a labelled *Taq*I primer (gat gag tcc tga ccg a (Vos *et al.* 1995)) and the original iPM4R primer. The AFLP thermal cycle conditions were applied as described (Vos *et al.* 1995). Separation of labelled fragments and autoradiography was done as described (Van Eck *et al.* 1995b).
Linkage analysis of pair wise recombination frequencies between segregating RGHs and markers were performed using JOINMAP 1.4 (Stam 1993).

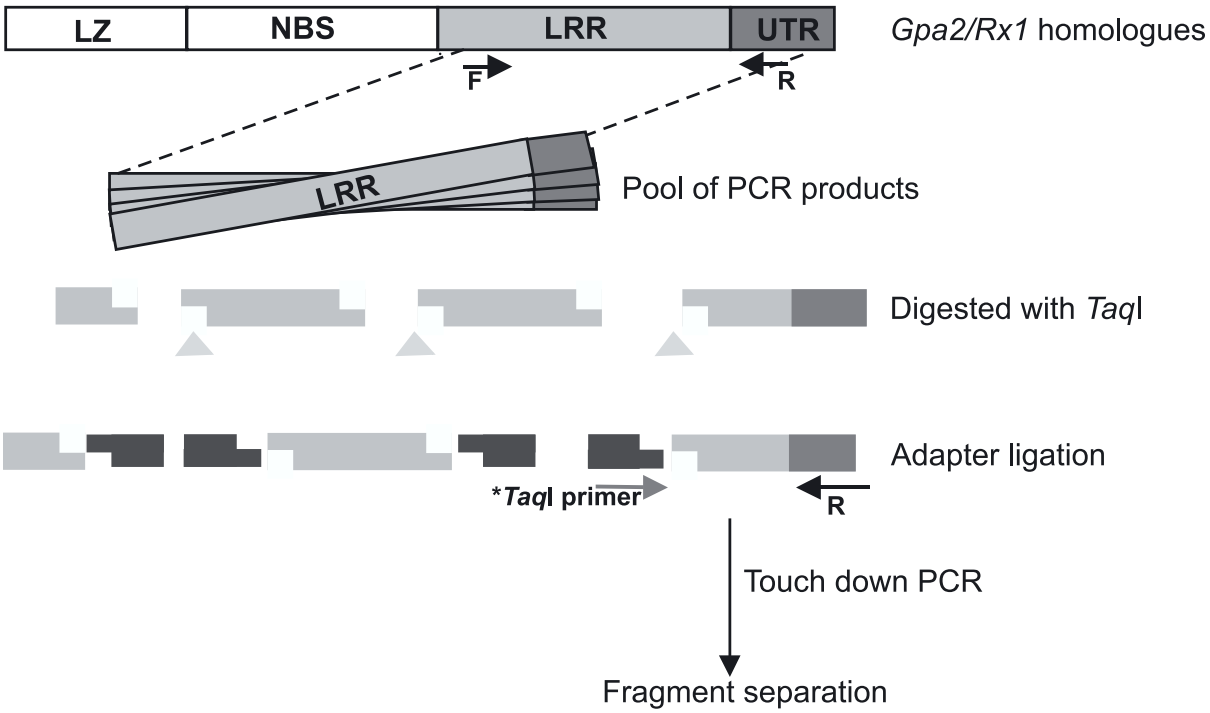


Figure 1 Schematic depiction of the method used to genetically map the *Gpa2/Rx1* homologues. The LRR-specific primers iPM4F and iPM4R are used to amplify a pool of RGH sequences from either genomic DNA or BAC DNA. Subsequently, the PCR products are used as template in a modified AFLP reaction. After digestion with the restriction enzyme *Taq*I the products are ligated to a *Taq*I adapter. A second touch down PCR reaction is performed using the *Taq*I primer and the original iPM4R primer.

Physical mapping

BAC-end sequences were employed to design PCR primers to study overlap between BAC clones. The primers designed for each BAC-end are listed in Table 1. PCR conditions used in the amplifications are equal for all primer pairs and are as follows: 3 min 94°C, followed by 30 s 94°C, 30 s 60°C, 90 s 72°C for 35 cycles and 5 min 72°C. For the alignment of the contigs in SH, we used primers as described (Kanyuka *et al.* 1999) and 187 kb sequence information derived from 4 overlapping BAC clones harbouring *Gpa2*, *Rx1, SH-RGH1* and *SH-RGH3* (previously designated *RGC3* (Van der Vossen *et al.* 2000)).

Table 1 Primer sequences based on BAC end sequences to detect overlap between BAC clones

BAC end	primers
7E16sp6	5'- cgg ggt gta atg tga tga gc-3' 5'- ggc ctg caa gtc tgt gca c-3'
7E16t7	5'- gtt cgt atg agc gag tat gg-3' 5'- tcc acg atg gtc tcc tcg-3'
25G18sp6	5'-cca att tca agc ttc ttc ata g-3' 5'-cag tca agg tgc ttt gga gg-3'
25G18t7	5'- gtt acc tgc tat gtg agc tc-3' 5'- cat cag ctg cct tgc agt tg-3'
36G3sp6	5'- gcc caa cat gat agg tcg c-3' 5'- ctt ggt atc aga gca cag ag -3'
36G3t7	5'-tgt atg aat tgg gtc att ccg-3' 5'- gcc caa tat tcc tcc atc tg-3'

Results

Identification of Gpa2/Rx1 homologues in SH and RH

To identify homologues of *Gpa2* and *Rx1*, a specific primer set was designed based on conserved regions in the LRR domain. This primer set was used to screen a BAC library of the diploid potato clone SH, which harbours the *R* genes *Gpa2* and *Rx1* on a *S. tuberosum* ssp. *andigena* introgression segment. For each positive BAC clone, amplification products were cloned and 12 transformants were used for reamplification. After *Taq*I digestion, resistance gene homologues (RGHs) were selected based on their unique digestion pattern (Fig. 2a). This resulted in the identification of three RGHs: *SH-RGH5* (BAC clone 7E16), *SH-RGH6* and *SH-RGH 7* (both located on BAC clones 25G18 and 36G3). Also *Gpa2* (BAC 85N1), *Rx1* (BAC 43) and *SH-RGH1* (BAC 60K18) were detected based on their expected *Taq*I digestion pattern (Bendahmane *et al.* 1999; Van der Vossen *et al.* 2000). No additional RGHs were found after screening genomic DNA of SH.

The diploid potato clone RH was studied using genomic DNA. PCR analysis resulted in three bands at 800 bp, 1300 bp and 1600 bp. The complete PCR product was cloned and analysis of the *Taq*I patterns of the reamplified PCR products of 25 clones resulted in the identification of six *Gpa2/Rx1* homologues: *RH-RGH1* to *RH-RGH6* (Fig. 2b). Interestingly, *RH-RGH4* and *RH-RGH5* resembled the digestion patterns of *SH-RGH5* and *SH-RGH6*, respectively. In total, twelve RGHs were identified in potato including *Gpa2, Rx1* and *SH-RGH1* using a *Gpa2/Rx1* cluster-specific primer set (Table 2). Another *Gpa2/Rx1* homologue present in the cluster, the pseudogene *SH-RGH3*, was not amplified with the cluster primer set. The annealing site of the

reverse primer is not present in *SH-RGH3,* because the 3′ end of its sequence is truncated.

The *Gpa2/Rx1* homologues of SH and RH were sequenced to determine their sequence resemblance with *Rx1* and *Gpa2*. Comparison of the nucleotide sequence of the LRR domains of the RGHs showed that they are highly similar to each other with sequence identities ranging from 85.5% to 100% (Table 3). The sequences of *RH-RGH4* and *RH-RGH5* were indeed virtually identical to those of *SH-RGH5* and *SH-RGH6,* respectively (99.9% and 100% sequence identity).

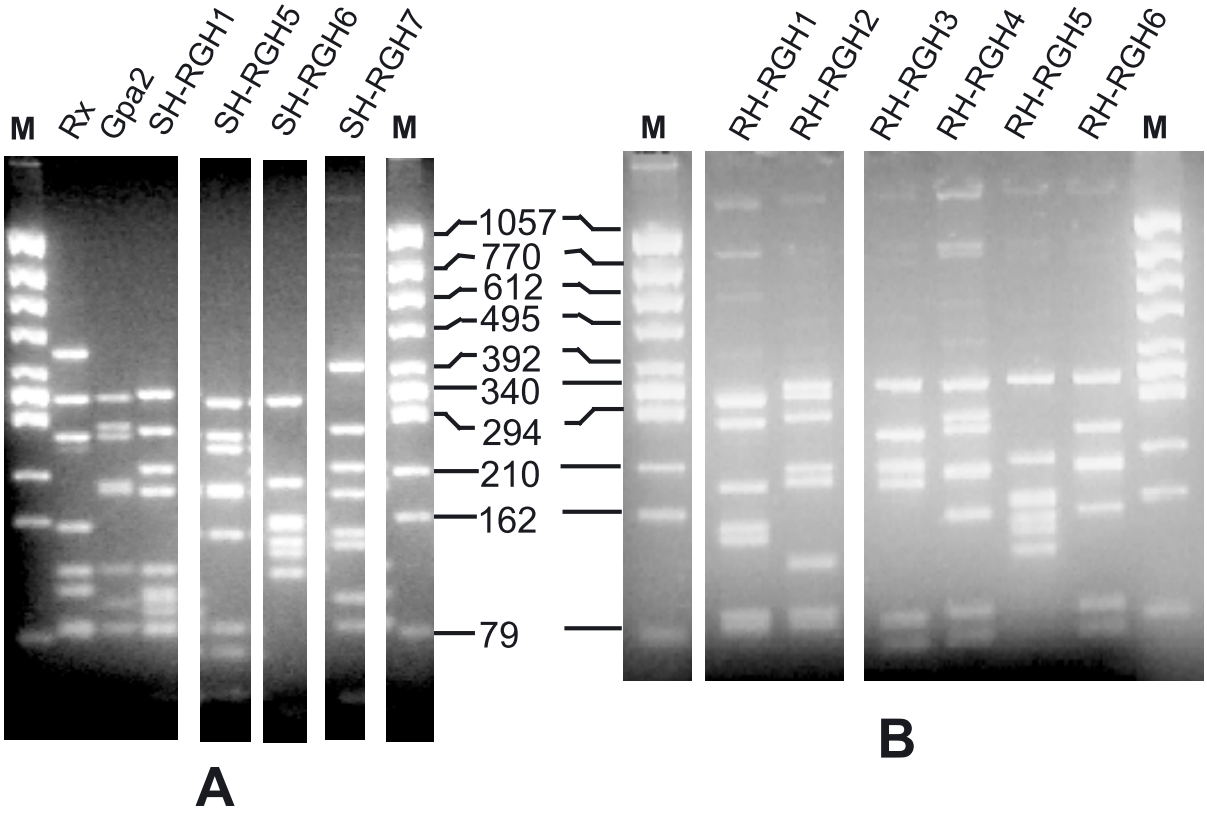


Figure 2 Identification of *Gpa2/Rx1* resistance gene homologues (RGHs) in the diploid potato clones SH (2A) and RH (2B). RGHs were selected based on differences in their *Taq*I digestion pattern after PCR amplification with specific LRR primers. The *Taq*I digestion patterns are related to a DNA base pair ladder (M). Bands larger then 612 bp are the result of partially digested DNA.

Table 2: Resistance gene homologues (RGHs) identified in BAC and genomic DNA of the diploid potato clones SH and RH using a primer pair based on the LRR domain of *Gpa2* and *Rx1*.

RGH	PCR product (kb)	BAC clones	gDNA
SH-RGH1	1.6	85N1	+
Gpa2	1.6	85N1, 60K18	-[a)]
Rx1	1.6	BAC 43	-
SH-RGH5	1.5	7E16	+
SH-RGH6	1.3	25G18, 36G3	+
SH-RGH7	1.6	25G18, 36G3	-
RH-RGH1	1.6	-	+
RH-RGH2	1.6	-	+
RH-RGH3	1.6	-	+
RH-RGH4	1.5	-	+
RH-RGH5	1.3	-	+
RH-RGH6	0.8	-	+

a) not detected in genomic DNA

Table 3: Sequence identities (%) of the LRR domains of the RGHs in SH and RH

RGH	*Gpa2*	*Rx1*	*SH-RGH1*	*SH-RGH3*	*SH-RGH5*	*SH-RGH6*	*SH-RGH7*	*RH-RGH1*	*RH-RGH2*	*RH-RGH3*	*RH-RGH4*	*RH-RGH5*
Rx1	93.5											
SH-RGH1	93.6	93.2										
SH-RGH3	86.5	86.8	87.3									
SH-RGH5	92.9	92.8	93.5	87.4								
SH-RGH6	94.3	93.2	93.1	92.6	92.4							
SH-RGH7	92.8	93.3	91.9	86.5	91.9	93.1						
RH-RGH1	92.9	93.8	91.8	87.0	91.6	92.8	98.5					
RH-RGH2	93.7	94.8	93.1	88.3	93.5	94.8	93.4	92.2				
RH-RGH3	94.1	93.7	93.9	87.7	93.1	94.5	91.2	93.3	93.9			
RH-RGH4	93.1	92.9	93.5	87.5	99.9	92.6	92.1	91.7	93.4	93.1		
RH-RGH5	94.3	93.2	93.1	92.6	92.4	100.0	93.0	92.8	94.7	94.5	92.5	
RH-RGH6	93.6	94.0	85.5	85.5	93.3	90.8	92.3	92.2	94.1	94.3	93.3	90.8

Genetic mapping of the Gpa2/Rx1 homologues in SH×RH

Gpa2/Rx1-specific primers were used to perform PCR on the parents SH and RH and a progeny of 100 individuals to determine the genetic position of the *Gpa2/Rx1* homologues in the potato genome. The selected BAC clones harbouring *Gpa2/Rx1* homologues were included as a control. On the resulting pools of PCR products, a modified AFLP analysis was performed that makes use of sequence polymorphisms in the last *Taq*I restriction sites of the PCR products. Using the sequence information of the RGHs we calculated the length of the expected fragments from the *Taq*I restriction site closest to the 3' end of the PCR products. In this way we were able to determine the positions in the gels of the RGHs. Figure 3 shows the autoradiogram with segregating bands for *Gpa2, Rx1, SH-RGH5, SH-RGH6* and *RH-RGH1-RH-RGH6* at the expected positions in the gel (summarised in Table 4). *SH-RGH1* and *SH-RGH7* cannot be linked to a segregating band, because they co-migrate with a thick band in the gel. However, *SH-RGH1* is present on the same BAC as *Gpa2*, and *SH-RGH7* on the same BAC as *SH-RGH6*. All the segregating RGHs were linked to the *Gpa2/Rx1* resistance gene cluster with a logarithm of odds (LOD) score between 8.2 and 12.6. Other DNA fragments visible on the autoradiogram either co-segregate with RGHs, or do not segregate at all. They may be artefacts caused by the double PCR and the limited complexity of the template used.

SH-RGH5 and *SH-RGH6* are in repulsion with *Gpa2*. This implies that *SH-RGH5*, *SH-RGH6* and *SH-RGH7* are all derived from the *S. tuberosum* ssp. *tuberosum* haplotype and that no additional RGHs have been found on the *Gpa2/Rx1* cluster of the *S. tuberosum* ssp. *andigena* introgression segment. Furthermore, *RH-RGH1, 2, 3* and *6* are in coupling with each other and in repulsion with *RH-RGH4* and *5* (Table 4). From these data we can conclude that all the identified homologues are genetically linked to the *Gpa2/Rx1* cluster on chromosome XII in SH and in RH. Based on these and other data (see also below), we postulated the position of the RGHs as shown in Figure 4.

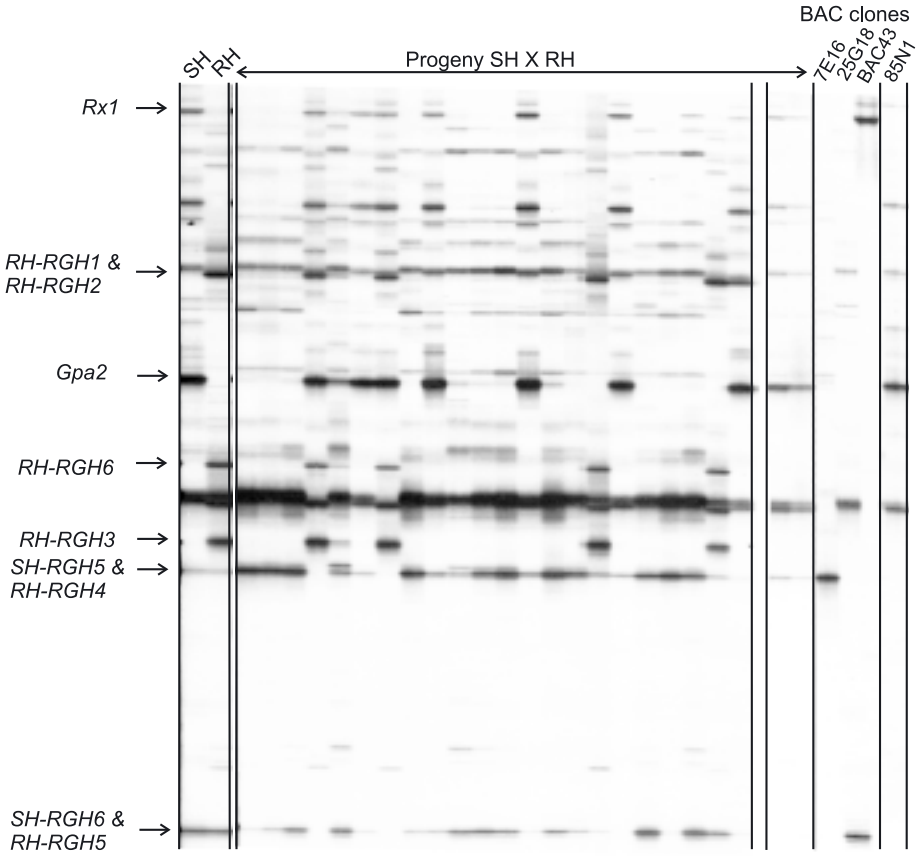


Figure 3 Mapping of RGHs in a SH×RH cross and in a number of BAC clones using a modified AFLP method on a pool of RGH sequences. All the indicated RGHs map on chromosome XII and are linked to *Gpa2* with a LOD score between 8.2 and 12.6. *Rx1* was also identified by the co-migrating band on BAC43 from which *Rx1* was originally cloned (Bendahmane *et al.* 1999). Likewise, we were also able to verify the expected bands for *Gpa2* and *SH-RGH1* by comparing the pattern of the mapping population with the bands in BAC 85N1. Furthermore, we could relate the bands present on BAC clones 7E16 and 25G18 with bands in the progeny. The majority of the remaining, unassigned bands are artefacts of the PCR procedure.

Table 4: Segregation of the RGHs in a progeny of 136 individuals. RGHs with the same mark (++ or --) are in coupling.

SH-RGH	segregation	RH-RGH	segregation
Gpa2	++	*RH-RGH1*	++
Rx1	++	*RH-RGH2*	++
SH-RGH5	--	*RH-RGH3*	++
SH-RGH6	--	*RH-RGH6*	++
		RH-RGH4	--
		RH-RGH5	--

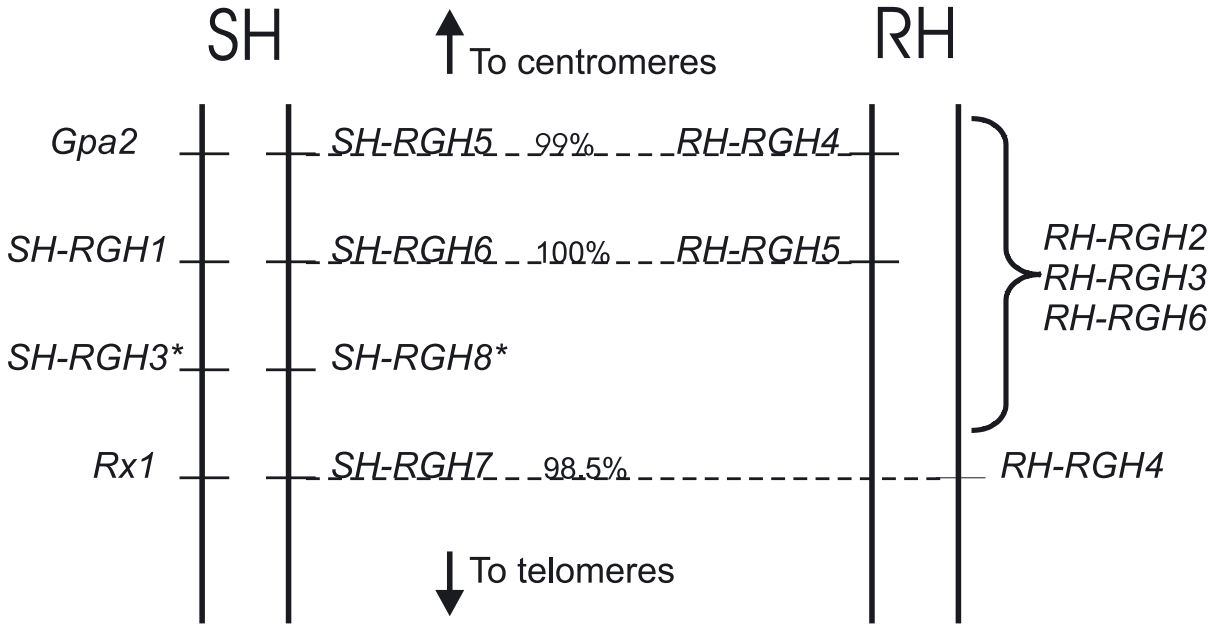


Figure 4 Schematic drawing of the *Gpa2/Rx1* homologues on the four chromosomes of SH and RH. The relative order and orientation of a number of RGHs has been postulated as follows (see also results). *SH-RGH5* is the only RGH present on a BAC clone positioned at the top of the cluster. *RH-RGH4* has a 99.9% identity to *SH-RGH5* and therefore we assume that this RGH is also positioned at the top of the cluster in RH. Because the intron positions of *SH-RGH7* are the same as in *Rx1* we assume that this RGH, like *Rx1*, is at the bottom of the cluster. *RH-RGH1* and *SH-RGH7* have a sequence identity of 98.5 and therefore we assume that *RH-RGH1* is also at the bottom of the cluster in RH. We also assume that the two pseudogenes *SH-RGH3* and *SH-RGH8* that could not be amplified in the PCR (marked with an asterisk) occupy the same position in the cluster. Finally, we assume, based on their 100% sequence identity, that *RH-RGH5* has the same position in the cluster as *SH-RGH6*. No relative order for *RH-RGH2, RH-RGH3* and *RH-RGH6* could be postulated.

Physical mapping of the Gpa2/Rx1 cluster in SH

The BAC clones containing RGHs were used to construct a physical map of the *Gpa2/Rx1* cluster for both the *S. tuberosum* spp. *tuberosum* and the *S. tuberosum* spp. *andigena* haplotypes of SH. Genetic mapping revealed that all these RGHs are located on chromosome XII and hence, the BAC inserts correspond with genomic fragments of this chromosomal region of the potato genome. *SH-RGH5* and *SH-RGH6* co-migrate with markers that are in repulsion with *Gpa2* implicating that they are located on the homologous chromosome. Based on this information we could assign the origin of the BAC clones to any of the two haplotypes: BAC 43, BAC 85N51 and BAC 34F16 were derived from the homologous chromosome in *S. tuberosum* ssp. *andigena* introgression segment and BAC clones 25G18, 7E16, and 36G3 were derived from *S. tuberosum* ssp. *tuberosum*. The BAC clones derived from *S. tuberosum* ssp. *andigena* formed a closed contig and could be easily aligned with the original physical map based on the presence or absence of *Gpa2*, *SH-RGH1,* and *Rx1* and the

markers 73L, 111R, 111L, 221R, 45L, 77R, and 77L (Kanyuka *et al.* 1999). In order to make a contig of the *S. tuberosum* ssp. *tuberosum* haplotype, the BAC ends were sequenced and primers were designed to perform PCR on the other BAC clones of this haplotype. Additionally, BACs 25G18 and 36G3 harbour the same RGHs. These data enabled us to construct a closed contig of these BAC clones (Fig. 5). CHEF electrophoresis of the BAC inserts (data not shown) showed that BAC 7E16 is approximately 120 kb. Both 36G3 and 25G18 are approximately 90 kb in size resulting in a physical map of *S. tuberosum* ssp. *tuberosum* of about 200 kb containing *SH-RGH5, 6* and *7*.

Moreover, we were also interested in comparing the genomic organisation of the *Gpa2/Rx1* cluster derived from the S. *tuberosum* spp. *tuberosum* haplotype and the one derived from the *S. tuberosum* ssp. *andigena* introgression segment. A segment of 187 kb of the *S. tuberosum* ssp. *andigena* haplotype has previously been sequenced for cloning *Rx1* and *Gpa2* (Van der Vossen *et al.* 2000). This sequence information was used to align the BAC end sequences of the BAC clones derived from *S. tuberosum* ssp. *tuberosum* with the contig of the *S. tuberosum* ssp. *andigena* haplotype. The approximate 80% sequence identity for the BAC ends of the left arms of BAC clones 7E16, 25G18 and 36G3 with regions between *Gpa2* and *SH-RGH1* led to the orientation of the BAC contig as proposed in Figure 5. The order of the homologues present on BAC clones 25G18 and 36G3 is not clear. However, comparison of intron positions between the RGHs showed that *SH-RGH7* has similar intron positions as *Rx1* and *SH-RGH6* has not (Bakker *et al.*, unpublished data). Therefore we presented the order of the homologues as depicted in Figure 5.

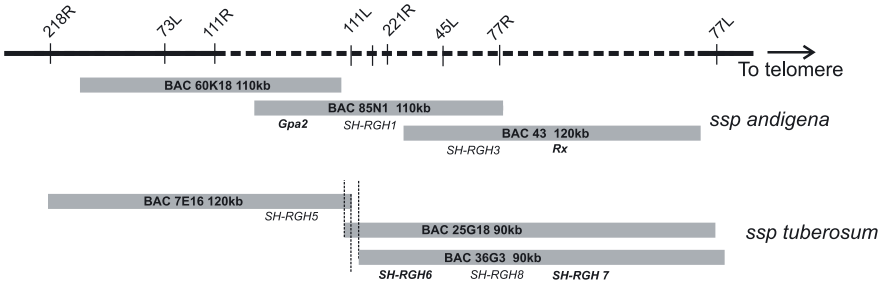


Figure 5 Physical map of the *Gpa2/Rx1* cluster in the diploid potato clone SH. The region (187 kb) around the resistance genes *Gpa2*, *Rx1* and *SH-RGH1* and the pseudogene *SH-RGH3* on the haplotype derived from *S. tuberosum* ssp *andigena* has been sequenced. On the *S. tuberosum* ssp. *tuberosum* haplotype also four *Gpa2/Rx1* homologues are present: *SH-RGH5, 6,7* and *8*. Like *SH-RGH3,* the homologue *SH-RGH8* could not be identified by PCR, but was detected by Southern blot analysis. The BAC clones are represented in grey rectangles and RGHs are indicated alongside the BACs on which they were detected. RGHs are indicated in bold when intron positions are known. Vertical lines indicate the marker positions used to align the BAC contig from the *S. tuberosum* ssp. *tuberosum* haplotype with the *S. tuberosum* ssp. *andigena* introgression segment. Vertical dotted lines indicate overlap between BAC ends. Dotted horizontal line in the *S. tuberosum* ssp. *andigena* chromosome indicates the position of the sequenced region of 187 kb.

Southern blot analysis
To confirm the specificity of the *Gpa2/Rx1*-specific primer set, southern blot analysis was carried out at high stringency conditions on *Hind*III digested DNA of the BAC clones from SH harbouring the RGHs (data not shown). Hybridisation with a 400 bp fragment derived from the 5'-end of the LRR domain of *Gpa2* resulted in a single band for 7E16 and four bands for 25G18 and 36G3. *SH-RGH7* has a *Hind*III site in the probe region which accounts for one of the two additional bands detected in 25G18 and 36G3. The other additional band indicates the presence of an extra RGH in the *Gpa2/Rx1* cluster in the haplotype derived from *S. tuberosum* ssp. *tuberosum.* Similar to the pseudogene *SH-RGH3*, this RGH is not amplified with the cluster specific primer pair. Based on this information together with the information on the positions of BAC 25G18 and 36G3, we assume that this additional RGH (designated *SH-RGH8*) is also a pseudogene and that it is most likely positioned between the homologues *SH-RGH6* and *SH-RGH7* (Fig. 4 and Fig. 5). It is noted that the sequenced homologous region in the haplotype of *S. tuberosum* ssp. *andigena* contains exactly the same number of homologues. In addition, a 185 bp fragment was used to determine the number of RGHs with an NBS domain. This probe reached 73 bp into the NBS domain and 70 bp into the LRR domain and gave identical results as the LRR probe. These data show that all *Gpa2/Rx1* homologues identified in the PCR based method also possess at least part of an NBS domain similar to *Gpa2* and *Rx1*.

The total number of *Gpa2/Rx1* homologues in the potato genome was also determined by southern blotting on genomic DNA of SH and RH. Hybridisation with the LRR and NBS probe confirmed our previous data and resulted in the detection of eight bands for SH and six for RH. All the bands for SH were also present in the BAC clones harbouring *Gpa2/Rx1* homologues. The observation that Southern blotting resulted in the same number of RGHs as the PCR approach indicates that no additional, slightly modified RGHs are present in RH.

Discussion

In this paper, we describe the identification and mapping of nine resistance gene homologues (RGHs) of a single complex locus in four homologous chromosomes using a cluster-specific primer combination based on the LRR domain of *Gpa2* and *Rx1*. Eight of the nine RGHs could be mapped on chromosome XII in a single step procedure using a modified AFLP method. Physical mapping revealed that the remaining homologue *(SH-RGH7)* was also located in the *Gpa2/Rx1* cluster on chromosome XII. The segregation of all six RH-RGHs in the mapping population shows that the two haplotypes of RH are heterozygous at the *Gpa2/Rx1* locus. This is confirmed by the fact that all six RH-RGHs have different sequences. Only two pairs of RGHs are (virtually) identical: *SH-RGH5* and *RH-RGH4* (99.9%) and *SH-RGH6* and

RH-RGH5 (100%). In both cases these identical RGHs are derived from different genotypes.

Gpa2 and *Rx1* are highly similar. At the amino acid level, they have a homology of 88% and at the nucleotide level their identity is even 93% (Van der Vossen *et al.* 2000). *SH-RGH1*, an RGH present on the same haplotype as *Gpa2* and *Rx1*, has also a similar sequence identity to the LRR domains of *Gpa2* and *Rx1,* respectively 93.6 and 93.5%. A complete open reading frame was detected for this putative *R* gene based on the 187 kb sequence of this region (Van der Vossen *et al.* 2000). This indicates that *SH-RGH1* could be a functional homologue of *Gpa2* and *Rx1* with unknown specificity. The nine RGHs identified in this study are all closely related to *Gpa2* and *Rx1* with sequence identities ranging from 93 to 95%. Southern analysis showed that all these RGHs have at least part of an NBS domain.

Our results indicate that all homologues are located in the *Gpa2/Rx1* cluster on chromosome XII. Surprisingly, in *S. aucaule* a functional *Gpa2/Rx1* homologue with the same specificity as *Rx1* is found (Bendahmane *et al.* 2000). This gene (*Rx2*), however, is genetically linked to RFLP marker Gp21 on chromosome 5 (Ritter *et al.* 1991; Bendahmane *et al.* 2000). Sequence identity between *Rx1* and *Rx2* is so high (97,9% for the complete genes and 99.4% for the LRR domains) that another *Rx1/Gpa2*-like cluster on chromosome V, if present in SH or RH, would certainly have been identified in this study. An explanation for this remarkable phenomenon that highly homologous genes are located on different chromosomes in two closely related species could be a recent translocation event after the speciation of *S. tuberosum* and *S. aucaule*. However, the synteny between the more distantly related species *S. tuberosum* and *Lycopersicon esculentum* is very high (Grube *et al.* 2000) indicating that such translocation events are rare within the genus *Solanum*.

In this paper, a PCR based method has been used to identify *Gpa2/Rx1* homologues. Remarkably, some of the amplified RGHs gave conflicting results, among others, with regard to physically mapping. For example, PCR analyses of nearly completely overlapping BAC clones (25G18 and 36G3) resulted in totally different RGHs. Sequence alignments of these RGHs with other RGHs (two by two) revealed that they were the result of a chimaeric PCR product derived from two distinct RGHs. These artificial RGHs consisted of two extraordinary stretches of several hundred nucleotides each. One stretch was 100% identical to one RGH and the other stretch was 100% identical to another, completely different RGH. These results were obtained with BAC DNA and genomic DNA as well. Fortunately, chimaeric PCR products were typically less frequent than genuine RGHs and occurred in most cases only once in a series of reamplified colonies. Although PCR techniques are commonly used to identify and map RGHs, this phenomenon has to our knowledge not been described before in the literature.

Despite the importance of potato as a food crop and its vulnerability to various pests and diseases only few *R* genes have been cloned. To facilitate cloning genes of interest an Ultra High Dense genetic map comprising 10,365 AFLP markers has been constructed (www.dpw.wag-ur.nl/uhd/). This has been accomplished by using the mapping population of SH×RH. However, not for all potato species that harbour interesting *R* genes a dense genetic map and BAC library will become available. The production of these tools is still laborious and costly, and the possibility to dissect *R* gene clusters with specific primer combinations in different species is a promising alternative. Comparative analysis has shown that the genomes of members of the *Solanaceae* family have a large synteny (Grube *et al.* 2000). The results described in this paper indicate that characterising *R* genes from other potato (sub)species with PCR-based approaches may be feasible.

Chapter 5

Structural diversity and evolutionary relationships of *Gpa2/Rx1* homologues in potato

This chapter is in preparation for publication as:
Erin Bakker, Patrick Butterbach, Martijn Holterman, Jaap Bakker and Aska Goverse
Structural diversity and evolutionary relationships of *Gpa2/Rx1* homologues in potato

Abstract

Sequence analysis was carried out on the C-termini of fourteen highly similar members of an *R* gene family in potato. The fourteen sequences were derived from five different haplotypes, and include the two virus resistance genes *Rx1* and *Rx2,* and the potato cyst nematode resistance gene *Gpa2*. Apart from these three functional *R* genes, cDNA was obtained for two other members. Using gDNA and cDNA sequences in a multiple sequence alignment enabled us to identify intron splice sites and putative open reading frames for ten members. Splice site positions appeared to be conserved, while intron sizes were variable. The acidic tail, consisting of two direct sequence repeats is present in *Rx1* and *Rx2,* but absent in *Gpa2*. For seven members, only one repeat of the acidic tail was observed and in three members the acidic tail was absent. Comparative sequence analysis of all *Gpa2/Rx1* homologues revealed that most sequence variation occurs in the LRR β-strand/β-turn motifs. K_a/K_s ratio analyses indicated that these motifs are also under stronger positive selection pressure than the complete C-terminal part of the genes. Comparing the *Gpa2/Rx* homologues revealed a patchwork of sequence identities for both paralogues and orthologues suggesting an evolutionary history that involved unequal crossing-over and gene conversion.

Introduction

Plants have evolved a surveillance system consisting of a large array of genes, which confer resistance to a wide range of pests and pathogens. According to the gene-for-gene model, an elicitor produced by the pathogen is specifically recognised (directly or indirectly) by an *R* gene product upon infection of the plant (Flor 1971). This results in the activation of a defence response. Over thirty *R* genes have been isolated from different plant species and based on their structural domains, they can be divided into distinct classes (reviewed by Hulbert *et al.* 2001). The majority of *R* genes encode for proteins with a leucine rich repeat (LRR) and a nucleotide binding site (NBS) domain. This superfamily of NBS-LRR proteins can be subdivided into a class of R proteins with a coiled-coil (CC) domain or a Toll/Interleukin-1 receptor domain (TIR) at the N-terminus.

R proteins are composed of functional modules involved in pathogen recognition and triggering of a disease resistance response. Several lines of evidence point at a role for the LRR domain in the determination of *R* gene specificity. Sequence homology with the LRR domain of the porcine ribonuclease inhibitor suggested that the three-dimensional structure of the LRR domain of *R* genes is horse shoe shaped with at the inside a series of parallel β sheets consisting of conserved structural amino acid residues forming the back bone and variable solvent exposed residues

that play a role in protein-protein interactions (Jones and Jones 1997). Comparative analysis of *R* gene sequences revealed that the LRR domain is the most variable part of the R protein suggesting a role in resistance specificity. This is supported by the observation that the solvent exposed amino acid residues of the LRR β-strand/β-turn motifs are hyper-variable and subject to diversifying selection (Parniske *et al.* 1997). Furthermore, combining the LRR domain of the flax resistance gene *L2* with the TIR-NBS regions of the *L6* and *L10* alleles resulted in a chimaeric gene product with *L2* specificity (Ellis *et al.* 1999), which demonstrates that the LRR domain is the main specificity determinant of the R protein.

Some *R* genes like the flax *L* gene reside in simple loci, but mostly they are found on complex loci harbouring several tandemly repeated *R* gene homologues (reviewed by Hulbert *et al.* 2001). In potato, a small gene family of four closely related *R* genes has been identified on a complex locus on chromosome XII including the virus resistance gene *Rx1* and the nematode resistance gene *Gpa2* (Bendahmane *et al.* 1999; Van der Vossen *et al.* 2000). A similar cluster of four additional *Gpa2/Rx1* homologues was identified on chromosome XII of the complementary susceptible haplotype of the diploid potato clone SH. In the susceptible diploid potato clone RH, respectively two and four additional *Gpa2/Rx1* homologues were identified on both haplotypes derived from *S. tuberosum* ssp *tuberosum* (Bakker *et al.* 2003). This multigene nature of resistance loci may facilitate meiotic instability in a heterozygous state. Unequal crossing-over and gene conversion has been suggested to play a major role in the generation of new *R* gene specificities (Pryor and Ellis 1993; Richter *et al.* 1995; Hammond-Kosack and Jones 1997; Hulbert 1997; Parniske *et al.* 1997). These processes, however, tend to homogenise the paralogues, which does not comply with the findings at the *Dm* and *Cf* clusters where orthologues are more similar than paralogues (Meyers *et al.* 1998; Parniske *et al.* 1997). Therefore, Michelmore and Meyers (1998) suggest that *R* genes mainly evolve by divergent evolution and a birth-and-death process.

In general, members of a complex resistance gene cluster like the *P* locus in flax (Dodds *et al.* 2001) and the *Cf* locus in tomato (Parniske *et al.* 1997) confer resistance to different pathotypes of the same pathogen species. In contrast, the *Gpa2/Rx1* locus harbours two *R* genes that confer resistance to two completely unrelated pathogens. *Gpa2* confers resistance to the potato cyst nematode *Globodera pallida*, while *Rx1* recognises the coat protein of the potato virus X. Interestingly, a third highly similar *R* gene (*Rx2*) has been identified on chromosome V of *S. acaule*, which confers resistance to PVX, like *Rx1* (Bendahmane *et al.* 2000). Sequence comparison of the three *R* genes revealed that *Rx1* and *Rx2* have only 5 amino acid differences in the LRR domain, whereas 69 amino acid differences were found for *Gpa2* (Bendahmane *et al.* 2000; Van der Vossen *et al.* 2000). Moreover, the ratio between non-synonymous and synonymous amino acid substitutions (K_a/K_s) suggested that the LRRs of *Gpa2* and *Rx1* are subject to diversifying selection (Van der Vossen *et al.* 2000). These data

suggest that the LRR domain of *Gpa2* and *Rx1* is involved in determining nematode and virus recognition specificity, respectively.

To gain more insight in the structural and evolutionary relationship of the *Gpa2/Rx1* cluster in potato, the C-terminal sequences of fourteen *Gpa2/Rx1* homologues have been analysed. First, the structural diversity of the homologues was examined based on the intron-exon composition, the presence or absence of InDels, the identification of hyper-variable amino acid positions, and other structural motifs like the acidic tail. Second, evolutionary relationships between the homologues were determined based on a similarity dendrogram, the K_a/K_s ratios, and informative polymorphic sites (IPS). A patchwork of sequence identities was found in the sequences of both paralogues and orthologues suggesting an evolutionary history that involved unequal crossing-over and gene conversion.

Materials and methods

cDNA preparation and PCR screening
Total RNA was isolated from potato clone SH using the TRIzol LS reagent (Invitrogen, San Diego, CA, USA) according to manufacturers instructions. From the total RNA, mRNA was subtracted using the Oligotex mRNA minikit (Quiagen, Hilden, Germany) according to manufacturers instructions. With the Marathon cDNA Amplification Kit (Clonetech, CA, USA) a 3' RACE was carried out as described (Bendahmane *et al.* 1999). In the first amplification round a *Gpa2/Rx1* LRR specific forward primer (Bakker *et al.* 2003) was used together with adapter primer AP1, provided with the kit. In this round, the following cycle file was applied: 2 min 94ºC followed by 5 s 94ºC, 4 min 72ºC for 4 cycles, then 5 s 94ºC, 4 min 70ºC for 4 cycles, and finally 5 s 94ºC, 4 min 68ºC for 24 cycles. The second amplification round was done using the LRR-F primer and a *Gpa2/Rx1* LRR specific reverse primer (Bakker *et al.* 2003) and the following PCR cycle file: 3 min 94ºC followed by 30 s 94ºC, 30 s 55ºC, 90 s 72ºC for 10 cycles, 30 s 94ºC, 30 s 55ºC, 90 s 72ºC with an extension of 5 s/cycle for 25 cycles and finally 5 min elongation at 72ºC.

The resulting PCR product was cloned into the PCR2.1TOPO-vector for transformation of *E. coli* TOP10 cells according to manufacturers instructions (Invitrogen, San Diego, CA, USA). Seventy positive clones were cultivated overnight at 37ºC in 2 ml liquid LB medium supplemented wit 100 µg/ml ampicillin for selection. Each clone was then reamplified and digested with *Taq*I. The DNA fragments were separated on a 4% agarose gel consisting of 1% ultra pure agarose (Life Technologies, Breda, The Netherlands) and 3% NuSieve ® GTG ® agarose (FMC, Philadelphia, PA, USA) in 1× TAE buffer at 120 V. Clones were selected for further analysis based on differences in the digestion patterns. To avoid sequencing

of PCR artefacts (Bakker *et al.* 2003), 3 clones were selected for each pattern represented by more than one clone. Sequencing was carried out by Greenomics, Wageningen, The Netherlands and sequence assembling was done with the software package Vector NTI (Informax, Inc. Bethesda, Maryland, USA).

Sequence analysis
Sequences of the RGHs *SH-RGH5*, *6* and *7* and *RH-RGH1* to *6* are obtained as described (Bakker *et al.* 2003). Sequences of the resistance gene homologues *RGC1* (designated *SH-RGH1*), *RGC3* (designated *SH-RGH3*) and the resistance genes *Gpa2* (Van der Vossen *et al.* 2000), *Rx1* (Bendahmane *et al.* 1999) and *Rx2* (Bendahmane *et al.* 2000) are obtained from genbank (accession numbers AF266747, AF266746, AF195939, AJ011801, AJ249448, respectively).

Sequences were aligned using Clustal W (Thompson *et al.* 1994) and K_a/K_s ratios were calculated using NewDiverge (Wisconsin Package version 8, September 1994, Genetics Computer Group, Madison, Wisconsin, USA). Significance of the K_a/K_s ratios was calculated using a G test of independence.

The appropriate substitution model and parameter values were determined with the program Modeltest v.3.06 (Posada and Crandall 1998). Using the results from Modeltest, a tree was constructed with the program PAUP* v.4.0b10 (Swofford 1998) implementing the Neighbour-Joining (NJ) method. Trees were bootstrapped 1000 times.

Results

Gene structure of the Gpa2/Rx1 homologues
The structure of the three resistance genes *Gpa2, Rx1* and *Rx2* is characterised by the presence of two introns located at the 3' end of the gene (Bendahmane *et al.* 2000; Bendahmane *et al.* 1999; Van der Vossen *et al.* 2000). To determine the intron and exon structure of the additional *Gpa2/Rx1* homologues, a PCR screening was carried out on cDNA from the diploid potato clone SH using a cluster specific primer set based on the sequence of the LRR region. Restriction analysis and sequencing resulted in the identification of cDNA fragments of *SH-RGH6* and *SH-RGH7* indicating expression of two putative resistance genes derived from the susceptible haplotype of SH. No cDNA from *SH-RGH3* could be obtained as the reverse primer does not anneal to the C-terminal end of the gene (Bakker *et al.* 2003). *SH-RGH1* and *SH-RGH5*, however, were also not retrieved from this cDNA pool. The constitutive expression level of the resistance genes *Gpa2* and *Rx1* is known to be very low and therefore, the cDNA of these genes could only be amplified from a cDNA pool that was made with an additional amplification step (Bendahmane *et al.* 1999; Van der

Vossen *et al.* 2000). Although the same procedure was used in this study, *Rx1* was identified eight times, but *Gpa2* could not be obtained after screening seventy clones. This suggests that the cDNA pool or the PCR screening was unsaturated. Therefore, it cannot be excluded that the *Gpa2/Rx1* homologues *SH-RGH1* and *SH-RGH5* are also expressed.

Comparison of the genomic sequence of *Gpa2, Rx1, Rx2 SH-RGH6,* and *SH-RGH7* with their cDNA sequences revealed the putative open reading frames of *SH-RGH6* and *SH-RGH7.* These analyses showed that the positions of the intron splice sites are very conserved. Therefore, a nucleotide alignment of genomic sequences was produced for all fourteen *Gpa2/Rx1* homologues to determine the most likely donor and acceptor sites for the remaining homologues *SH-RGH1, SH-RGH3, SH-RGH5,* and *RH-RGH1-6*. At the intron splice sites, the alignment was unambiguous and the donor and acceptor sites of *SH-RGH1, SH-RGH5* and *RH-RGH1-6* could be predicted (Fig. 1). The majority of the genes (*SH-RGH5* and *7,* and *RH-RGH1, 2, 3* and *4*) have splice sites at the same positions as *Rx1, Rx2,* and *Gpa2* resulting in the presence of two introns. Intron I starts at +1176 and ends at +1419 and Intron II starts at +1546 and ends at +1661 of the consensus sequence. There are, however, some exceptions resulting in unique gene structures. *SH-RGH1* has only intron I spliced, because the donor site of intron II is not present. *SH-RGH6* and *RH-RGH5*, which are virtual identical, both have a 158 bp deletion spanning the acceptor site of intron I and a 71 bp deletion spanning the donor site of intron II resulting in the fusion of intron I and intron II. *RH-RGH6* has an 834 bp deletion spanning intron I, so only intron II is spliced. *SH-RGH3* only has the donor site of intron I in common with the other homologues. At position +1418, 1 bp upstream of the expected acceptor site, the sequence of *SH-RGH3* has a 320 bp deletion compared to *Rx1* and from that position onwards, it shows homology to the 5' UTR of *Gpa2, Rx1* and *SH-RGH1.* Although the positions of the splice-sites are very conserved in the *Gpa2*/Rx homologues, the size of the introns varies substantially (Fig. 1). The sizes of Intron I range between 205 bp (*SH-RGH4* and *RH-RGH5)* and 241 bp (*RH-RGH1* and *2*), whereas the sizes of Intron II range from 92 bp (*SH-RGH5* and *RH-RGH4*) to 113 bp (*SH-RGH7* and *RH-RGH3*).

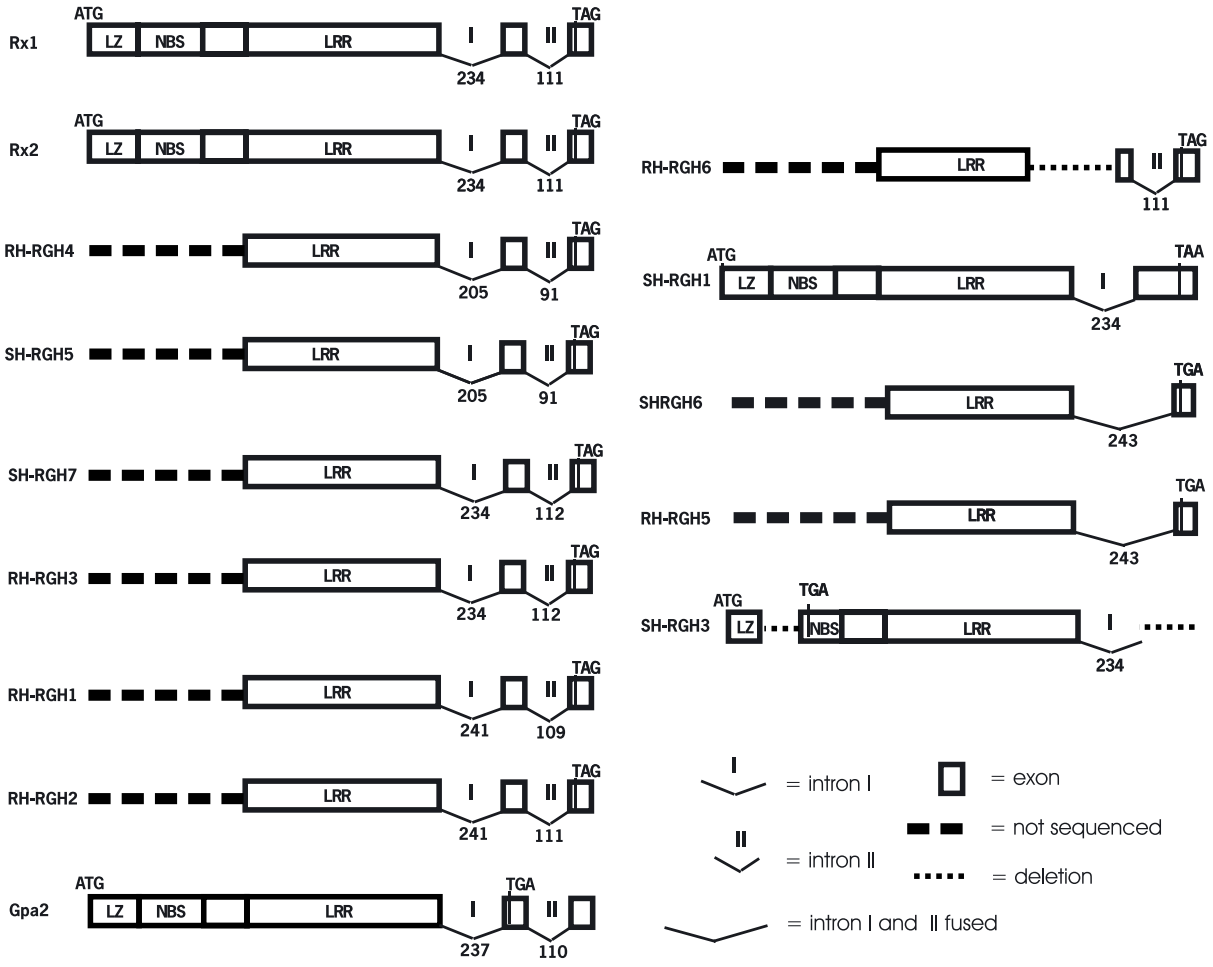


Figure 1 Schematic picture of the *Gpa2/Rx1* homologues derived from five potato haplotypes with their different domains (LZ = leucine zipper, NBS = nucleotide binding site, LRR = leucine rich repeat). The intron sizes (in bp), stop codons (TAG, TGA or TAA) and start codons (ATG) are indicated. The horizontal dashes indicate parts that are not sequenced. The horizontal dotted lines indicate sequence gaps.

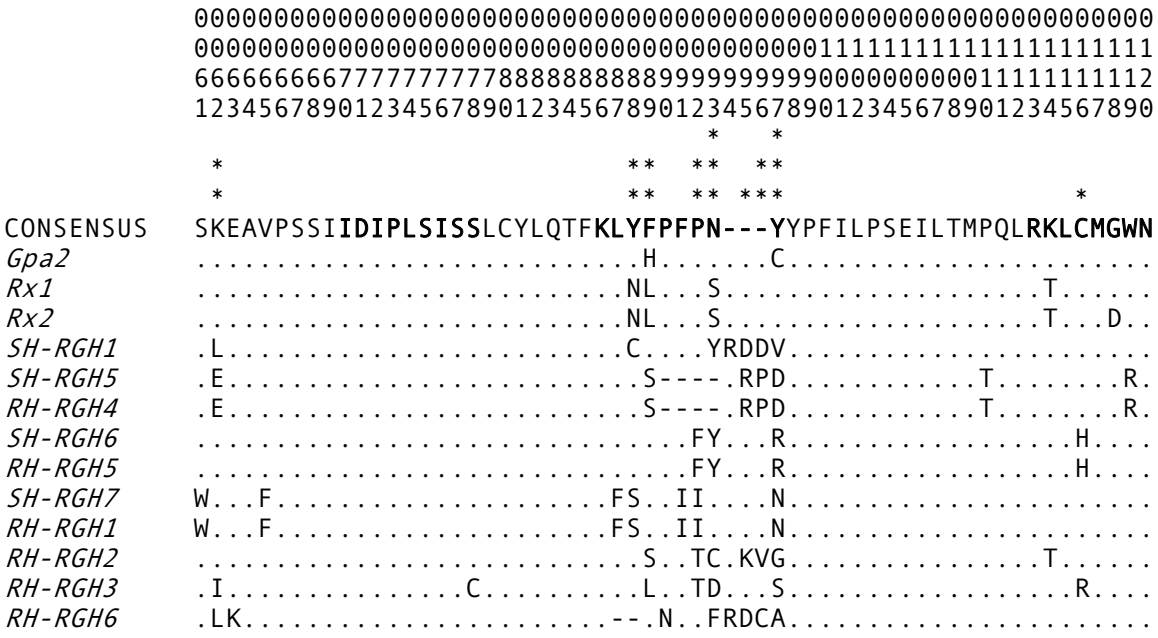
Structural diversity

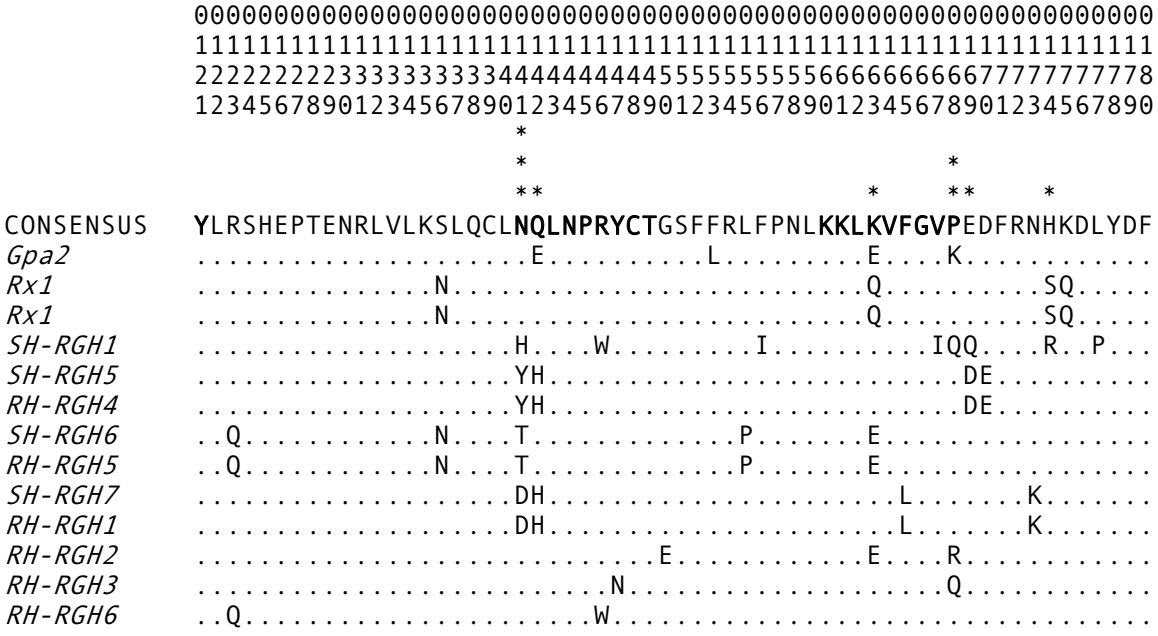
The genes *Gpa2, Rx1* and *Rx2* encode for a predicted polypeptide of 912, 937, and 938 amino acids, respectively (Bendahmane *et al.* 1999; Van der Vossen *et al.* 2000; Bendahmane *et al.* 2000). *SH-RGH1* could be translated in a putative amino acid sequence of 1048 residues. A frameshift in *SH-RGH3,* caused by an 89bp deletion from position +443, resulted in a short polypeptide due to multiple stop codons starting from residue 156. For the additional *Gpa2/Rx1* homologues *SH-RGH5, 6* and

7 and *RH-RGH1-6*, the amino acid sequence of the C-terminal part was predicted by translation of the putative open reading frames. Comparison of the *Gpa2/Rx1* homologues showed that *SH-RGH5* and *7* and *RH-RGH1-4* and *6* have the same stop codon (TAG) as *Rx1* and *Rx2* (Fig. 1). For *SH-RGH6* and *RH-RGH5*, this stop codon has a point mutation resulting in a lysine and the alternative stop codon encoded by TGA is located 7 amino acids downstream. In *SH-RGH1,* a frame shift due to the missing donor site of intron II resulted in a stop codon encoded by TAA (+1576). Finally, an 11 bp insertion directly downstream of Intron I in *Gpa2* results in a frame shift and the use of an alternative stop codon TGA, which is located 2 amino acids downstream of this insertion (Van der Vossen *et al.* 2000).

Alignment of the thirteen putative amino acid sequences of the C-terminal parts of the proteins resulted in a consensus sequence of 440 amino acids (Fig. 2). They exhibit a high degree of homology, with levels ranging from 75% to 100%. The most striking differences can be found at the C-terminal end of the sequence, were the two virus resistance genes *Rx1* and *Rx2* harbour an acidic tail (Bendahmane *et al.* 1999; Bendahmane *et al.* 2000) consisting of two direct repeats composed of SVTT(V/D)EDDDD (Fig. 2). Comparison with the other *Gpa2/Rx1* homologues revealed that *SH-RGH1, SH-RGH5, RH-RGH4, SH-RGH7, RH-RGH1, RH-RGH2,* and *RH-RGH6* have a single acidic repeat with some minor differences (S406T in *RH-RGH6*, T409I in *SH-RGH7,* D410Y in *RH-RGH1* and *RH-RGH2,* and D415 is deleted in *SH-RGH7*, *RH-RGH1* and *RH-RGH2*). The acidic tail is completely absent in *Gpa2, SH-RGH6, RH-RGH5* and *RH-RGH3* (Fig. 2).

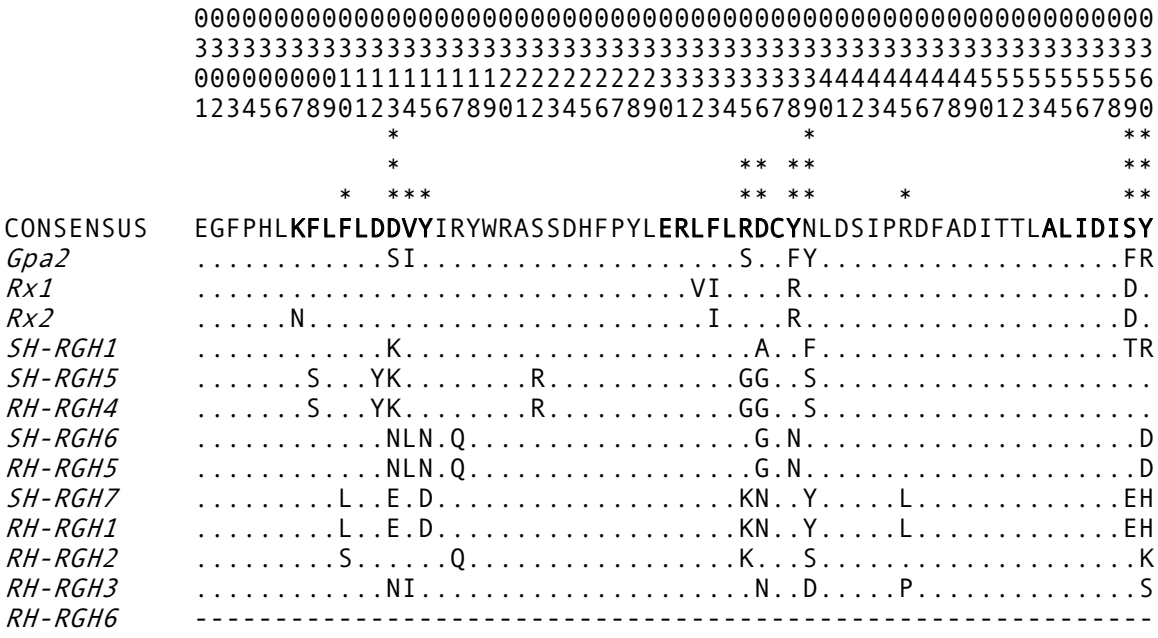
```
                000000000000000000000000000000000000000000000000000000000000000000
                000000000000000000000000000000000000000000000000000000000000000000
                000000001111111111222222222233333333334444444445555555556
                123456789012345678901234567890123456789012345678901234567890
                       *                    **
                    **** * *              * **                    * * *  **
CONSENSUS       HSIIMLRGFKCVTLELSFKLVRVLDLGLTRCPIFPSGVLSLIHLRYLSLCFYPCLQQYRG
Gpa2            ....T.CI.....................T...................R.N.R......
Rx1             ......G..E..................NTW..................R.N......Q.
Rx2             ......G..E..................NTW..................R.N......Q.
SH-RGH1         ....TFC...F................S.I.....................N...K....
SH-RGH5         ....TFCR..FF........L.....H...Y.....................F..L.L.
RH-RGH4         ....TFCR..FF........L.....H...Y.....................F..L.L.
SH-RGH6         .....S.R.....................F.....................K..L.
RH-RGH5         .....S.R.....................F.....................K..L.
SH-RGH7         ....V..L.E...P.............W.P..S................R.T.......E
RH-RGH1         ....V..L.E...P.............W.P..S................R.T.......E
RH-RGH2         .......L.....................I................C......VL..LR
RH-RGH3         .........N.I...............W.P.....................N...L....
RH-RGH6         ......G.L..F................NIY.....................K....
```

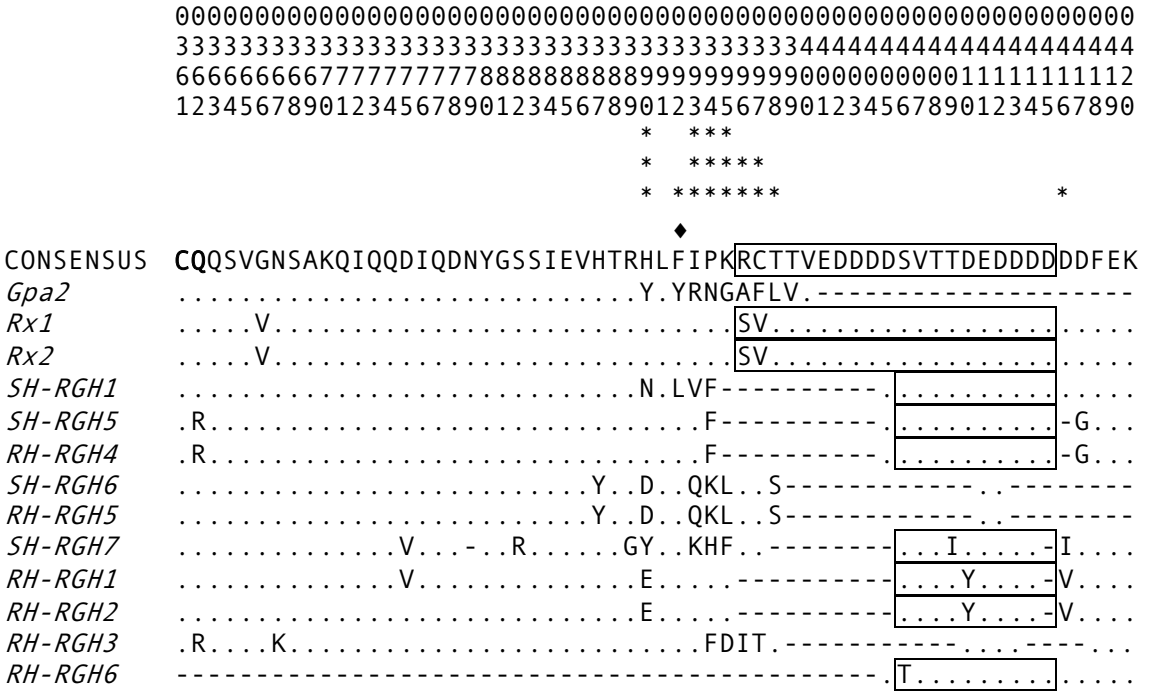

CONSENSUS SKEAVPSSIIDIPLSISSLCYLQTFKLYFPFPN---YYPFILPSEILTMPQLRKLCMGWN
Gpa2 ..........H.......C.....
Rx1 ..........NL...S.....T......
Rx2 ..........NL...S.....T...D..
SH-RGH1 .L..........C....YRDDV.....
SH-RGH5 .E..........S----.RPD.....T........R.
RH-RGH4 .E..........S----.RPD.....T........R.
SH-RGH6 ..........FY...R.....H....
RH-RGH5 ..........FY...R.....H....
SH-RGH7 W...F.....FS..II....N.....
RH-RGH1 W...F.....FS..II....N.....
RH-RGH2 ..........S..TC.KVG.....T......
RH-RGH3 .I.....C.....L..TD...S.....R....
RH-RGH6 .LK.....--.N..FRDCA.....


CONSENSUS YLRSHEPTENRLVLKSLQCLNQLNPRYCTGSFFRLFPNLKKLKVFGVPEDFRNHKDLYDF
Gpa2 .....E.....L.....E....K.....
Rx1 .....N.....Q.....SQ.....
Rx1 .....N.....Q.....SQ.....
SH-RGH1 .....H...W.....I.....IQQ...R..P...
SH-RGH5 .....YH.....DE.....
RH-RGH4 .....YH.....DE.....
SH-RGH6 ..Q.....N....T.....P.....E.....
RH-RGH5 ..Q.....N....T.....P.....E.....
SH-RGH7 .....DH.....L.....K.....
RH-RGH1 .....DH.....L.....K.....
RH-RGH2 .....E.....E....R.....
RH-RGH3 .....N.....Q.....
RH-RGH6 ..Q.....W.....

```
                 00000000000000000000000000000000000000000000000000000000000000000000
                 11111111111111111112222222222222222222222222222222222222222222222222
                 88888888899999999990000000000111111111122222222223333333334
                 12345678901234567890123456789012345678901234567890123456789012345678890
                              *   *
                            * **   *                                  * *  *
                            * ******   *  *  *               * * ***** **
CONSENSUS        RYLYQLEKLAFRIYYPSGQVAACFLKNTAPSGSTPQDPLRFQTEILHKETDFGGTAPPTD
Gpa2             ..........ST..S.---S.........L...........M.T..L..HSRA...P..
Rx1              .......E.T..L...Y---.............-.............I........--
Rx2              .......E.T..L...Y---.................................I........--
SH-RGH1          .........E.S.WR.AGQV...I...........E........F.RY.F.WE...---
SH-RGH5          ..............TP---D.....I..............I.T.Y.A...R..----
RH-RGH4          ..............TP---D.....I..............I.T.Y.A...R..----
RH-RGH5          .........V.SS..-------------------------------------------...
SH-RGH6          .........V.SS..-------------------------------------------...
SH-RGH7          .C.........H.H.S.---....VES..................R--K..WA..S...
RH-RGH1          .C.........H.H.S.---....VES..................R--K..WA..S...
RH-RGH2          ..........GF.H.Y---D..........Y...........T..............
RH-RGH3          ............V...G---...V.ES..............KL.Y.K.Q..KAV..AD
RH-RGH6          .........V.-------------------------------------------------
```

```
                 00000000000000000000000000000000000000000000000000000000000000000000
                 22222222222222222222222222222222222222222222222222222222222222222223
                 44444444455555555556666666666777777777788888888899999999990
                 12345678901234567890123456789012345678901234567890123456789012345678890
                                                          *
                                          *               *
                           *          * * *            *  *
CONSENSUS        VPTLLLPPPPAFPQNLKSLTFRGEFSLAWKDLSIVGKLPKLEVLKLSYNAFKGEEWEVVE
Gpa2             ...F.....DC..........S.D.F..................Q..H............
Rx1              --.........................V..................I..W...I.K......
Rx2              --.........................V..................I..W...I.K......
SH-RGH1          --.......................F.........................P.........A
SH-RGH5          L.....................S.Y.............................K......
RH-RGH4          L.....................S.Y.............................K......
SH-RGH6          ...................A.K.N.L..................Q..G...I........
RH-RGH5          ...................A.K.N.L..................Q..G...I........
SH-RGH7          ..I....SL.D.....................................CY...........
RH-RGH1          ..I....SL.D.....................................CY...........
RH-RGH2          ............................R.................G.....K......
RH-RGH3          .....................S...F..S..........E.......D..VI........
RH-RGH6          ---------------------------------------------------------------
```


CONSENSUS EGFPHLKFLFLDDVYIRYWRASSDHFPYLERLFLRDCYNLDSIPRDFADITTLALIDISY
Gpa2 .....SI.....S..FY.....FR
Rx1 .....VI....R.....D.
Rx2 .....N.....I....R.....D.
SH-RGH1 .....K.....A..F.....TR
SH-RGH5 .....S...YK.....R.....GG..S.....
RH-RGH4 .....S...YK.....R.....GG..S.....
SH-RGH6 .....NLN.Q.....G.N.....D
RH-RGH5 .....NLN.Q.....G.N.....D
SH-RGH7 .....L..E.D.....KN..Y.....L.....EH
RH-RGH1 .....L..E.D.....KN..Y.....L.....EH
RH-RGH2 .....S.....Q.....K...S.....K
RH-RGH3 .....NI.....N..D.....P.....S
RH-RGH6 ----------


CONSENSUS CQQSVGNSAKQIQQDIQDNYGSSIEVHTRHLFIPKRCTTVEDDDDSVTTDEDDDDDDFEK
Gpa2 .....Y.YRNGAFLV.----------
Rx1 .....V.....SV.....
Rx2 .....V.....SV.....
SH-RGH1 .....N.LVF----------.....
SH-RGH5 .R.....F----------.....-G...
RH-RGH4 .R.....F----------.....-G...
SH-RGH6 .....Y..D..QKL..S----------..--------
RH-RGH5 .....Y..D..QKL..S----------..--------
SH-RGH7 .....V...-..R.....GY..KHF..--------...I.....-I....
RH-RGH1 .....V.....E.....----------....Y....-V....
RH-RGH2 .....E.....----------....Y....-V....
RH-RGH3 .R....K.....FDIT.----------.....
RH-RGH6 ----------.T.....

```
            0000000000000000000
            4444444444444444444
            2222222233333333334
            1234567890123456789
                      *
                      ♦
CONSENSUS   EVASCRNNVEVLIPACSFLL
Gpa2        --------------------
Rx1         ..........----------
Rx2         ..........----------
SH-RGH1     .....H...Q..........
SH-RGH5     ..........----------
RH-RGH4     ..........----------
SH-RGH6     --------------------
RH-RGH5     --------------------
SH-RGH7     ..........----------
RH-RGH1     ..........----------
RH-RGH2     .........K----------
RH-RGH3     ..........----------
RH-RGH6     .........DK----------
```

Figure 2 Multiple amino acid alignment of the LRRs of the RGHs. Dots indicate nucleotides identical to the consensus sequence and dashes indicate sequence gaps. Asterisks indicate the variability of the amino acids at the different positions. No asterisk = 1 or 2 different amino acids, 1, 2, and 3 asterisks indicate positions with 3, 4, and 5 or more different amino acids respectively. The places of the introns, if applicable, are marked with diamonds. The acidic tail consisting of two direct sequence repeats is framed. Single acidic repeats are also framed. The putative β-strand/β-turn motifs are indicated in the consensus sequence in bold.

The number of LRRs could affect the specificity of resistance genes (Dixon *et al.* 1998). Due to the position of the cluster specific primers used to amplify the *Gpa2/Rx* homologues, only fifteen out of sixteen leucine rich repeats of the resistance genes *Gpa2, Rx1,* and *Rx2* are under investigation in this study. The majority of the other homologues also have fifteen leucine rich repeats. *SH-RGH6* and *RH-RGH5* have fourteen repeats as a result of a 126 bp deletion in the LRR domain and *RH-RGH6* has only nine repeats, due to a 834 bp deletion in the LRR domain. The VLDL motif, which is present in the third β-strand/β-turn motif of many NBS-LRR proteins (Axtell *et al.* 2001) and is supposed to be a negative regulatory domain (Bendahmane *et al.* 2002), is present in all thirteen RGHs that could be translated into a putative amino acid sequence.

Within the LRR domain, the solvent exposed amino acids are the most likely candidates to be involved in protein-protein interaction and therefore, these amino acids could be involved in *R* gene specificity. It is supposed that amino acids involved in *R* gene specificities are under selective pressure and these amino acid

positions will be hyper-variable in a multiple amino acid alignment. According to Parniske *et al.* (1997), who compared eleven *Cf* homologues, sites with four different amino acids were considered as variable and sites with five or more different amino acids hyper-variable. Comparing the amino acids of the different RGHs at the same position (Fig. 2) revealed that at the majority the positions (371 positions or 84.3%), a single or two different amino acids were present, whereas at 35 positions (8.0%) three different amino acid residues were detected, at 20 positions (4.5%) four amino acids were observed (variable sites) and at 14 positions (3.2%) five or more different amino acids were found (hyper-variable sites). Figure 2 shows that the variable and hyper-variable sites are not randomly distributed over the sequence. They can be roughly divided into fourteen clusters of (hyper-)variable amino acid positions often located in the LRR β-strand/β-turn motifs. The variability in the LRR β-strand/β-turn motifs points at a role in specific protein-protein interactions.

The ratio of nucleotide changes that result in asynonymous amino acids (K_a) and those that result in synonymous amino acids (K_s) can be a measure to identify the evolutionary pressure on a stretch of coding sequences (Parniske *et al.* 1997). When the evolutionary pressure is neutral, the amount of changes resulting in the same amino acid should equal the amount of changes resulting in different amino acid and thus result in $K_a/K_s = 1$. When $K_a/K_s < 1$, changes are not favoured and the evolutionary pressure is towards conservation. The third possibility is a $K_a/K_s > 1$. In that case, changes in the amino acid composition are favoured and the evolutionary pressure is toward diversification. The K_a/K_s ratios were calculated pair wise for the complete LRR domains (Table 1a) and the LRR β-strand/β-turn motifs (Table 1b) of all RGHs, including the pseudogene *SH-RGH3*. The mean K_a/K_s ratio for the complete sequence under inspection is 2.69, but the K_a/K_s ratio of the LRR β-strand/β-turn motifs is even 3.33. And although the two mean K_a/K_s ratios are not statistically significantly different from each other, this is an indication that the LRR β-strand/β-turn motifs are under stronger selective evolution than the rest of the sequence under inspection. Interestingly, consistent with the role of the LRR in pathogen recognition and *R* gene specificity, the K_a/K_s ratio of the LRRs of *Rx1* and *Rx2* is 0.5. Although there are not enough substitutions to make a statistically significant calculation, it is striking that the K_a/K_s ratio of the C-terminal half of *Rx1* and *Rx2* is the only ratio that suggests conserved selection pressure.

Table 1 The K_a/K_s ratios of the complete 3' sequences (a) and of the LRR -strand/-turn motifs of the RGHs (b). The G test of independence was used to calculate if a value is significantly different from 1 (*= 0.05, **= 0.025, ***= 0.01, ****= 0.001, no asterisk means no significance, X = K_a or K_s = 0, *SH* = *SH-RGH*, and *RH* = *RH-RGH*).

a

	RH6	*RH5*	*RH4*	*RH3*	*RH2*	*RH1*	*SH7*	*SH6*	*SH5*	*SH3*	*SH1*	*Rx2*	*Rx1*
Gpa2	3.23	2.68**	2.39****	4.42****	2.55***	3.46****	3.61****	2.62**	2.39****	3.15***	2.83****	2.72***	2.24***
Rx1	1.93	2.32***	2.11****	2.80****	1.97**	3.40****	3.07****	2.38***	2.11****	2.91****	2.07***	0.51	
Rx2	2.00	2.36***	2.52****	3.46****	2.43***	4.47****	3.82****	2.42***	2.52****	3.79****	2.43***		
SH1	1.23	1.49	2.26****	2.57***	2.04***	2.58****	2.22***	1.73	2.26****	1.97***			
SH3	1.75	2.96***	2.19****	3.96****	2.71****	5.69****	5.24****	2.89***	2.19****				
SH5	1.56	1.89**	X	2.02****	2.35****	2.64****	2.51****	1.58***					
SH6	2.01	X	1.85**	3.22***	3.31**	3.07****	2.88****						
SH7	2.38*	2.93****	2.51****	4.97****	3.58****	2.17							
RH1	2.01	3.12****	2.64****	5.65****	3.96****								
RH2	1.94	3.31***	2.35****	3.74****									
RH3	2.44	3.24***	2.02****										
RH4	1.56	1.89**											
RH5	2.01												
RH6													

b

	RH6	*RH5*	*RH4*	*RH3*	*RH2*	*RH1*	*SH7*	*SH6*	*SH5*	*SH3*	*SH1*	*Rx2*	*Rx1*
Gpa2	2.87	4.01**	2.58****	4.49****	4.67***	3.58****	3.67****	4.01**	2.58****	4.10****	3.43****	3.34***	3.27***
Rx1	2.00	3.58***	2.37***	3.80****	3.84***	3.16***	3.26***	3.54***	2.37***	4.00***	2.62***	X	
Rx2	2.14	3.66***	2.41***	3.87****	3.93***	3.23***	3.33***	3.66***	2.41***	4.08***	2.66***		
SH1	1.25	1.84	2.75***	3.05***	2.30**	2.04**	2.10**	1.84	2.75***	1.80**			
SH3	1.87	4.85***	1.98***	4.68****	4.67****	4.19****	4.31****	4.85***	1.98***				
SH5	1.10	1.34	X	1.90***	2.42***	2.16**	2.22**	1.34					
SH6	1.86	X	1.34*	5.81***	4.41**	3.64****	3.54****						
SH7	2.03	3.54****	2.22**	5.25****	5.59****	X							
RH1	1.93	3.64****	2.16**	5.11****	5.42****								
RH2	2.75	4.41***	2.42***	5.81****									
RH3	3.36	5.81***	1.92***										
RH4	1.10	1.34											
RH5	1.86												

Using Modeltest for selecting the appropriate substitution model and parameter values for constructing a similarity dendrogram with the nucleotide alignment of the fourteen *Gpa2/Rx1* homologues (Posada and Crandall 1998), resulted in the HKY model with invariable sites en heterogeneous substitution rates. The similarity dendrogram was subsequently produced using this model in a neighbour joining analysis. Figure 3 shows that, as expected, the two sets of (virtually) identical homologues *SH-RGH5/RH-RGH4* and *SH-RGH6/RH-RGH5* branch together separated from the main group. Also the couples *Rx1/Rx2* and *SH-RGH7/RH-RGH1,* for which the similarity is higher than average, branch together separated from the main group. It is surprising that *Gpa2* and the pseudogene *SH-RGH3* form a separated branch, although the bootstrap value is only 76. The other homologues do not form separated branches, or show low (<65) bootstrap values.

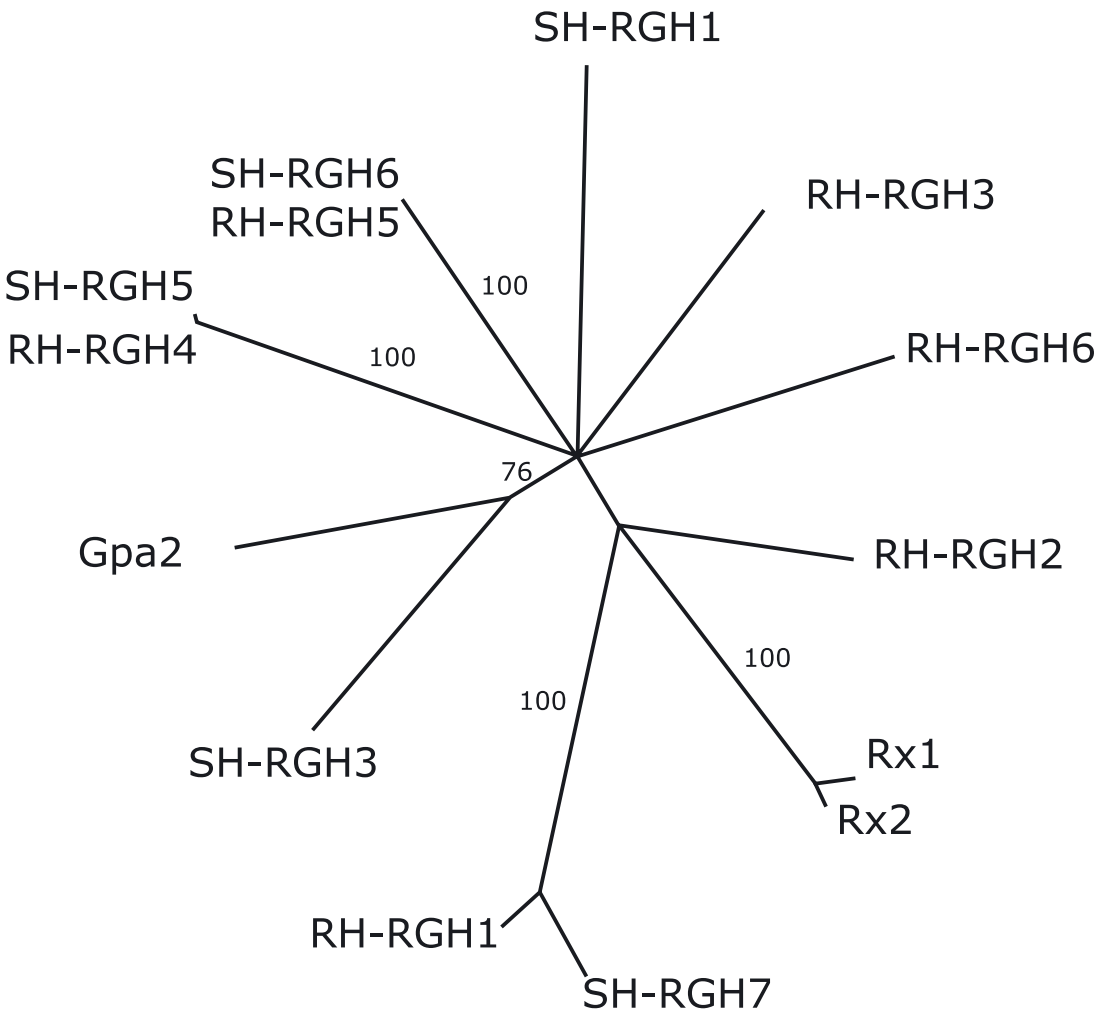
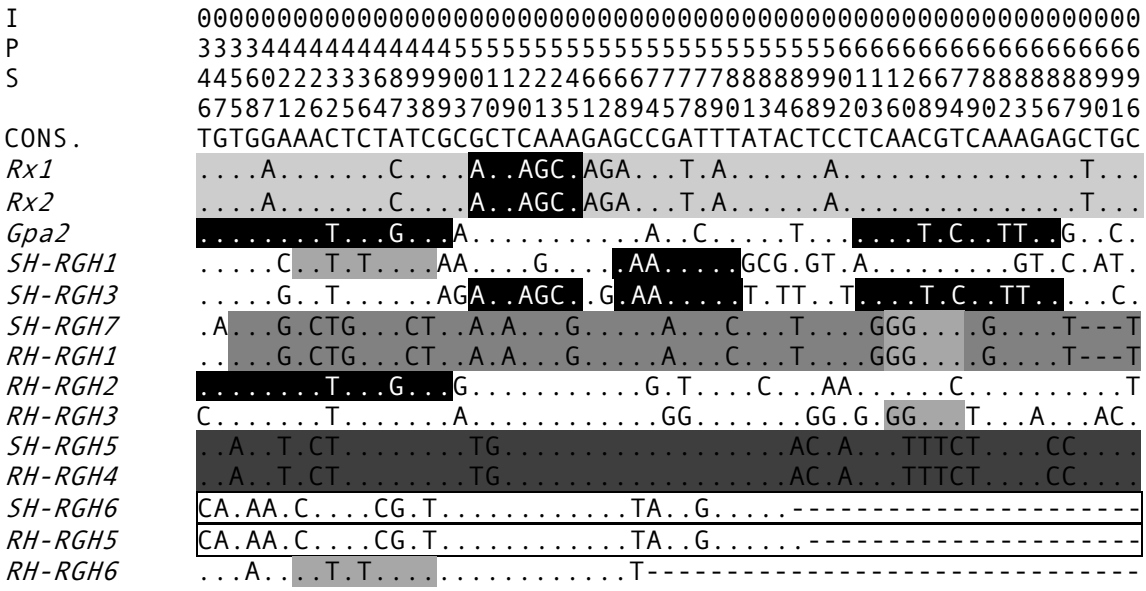
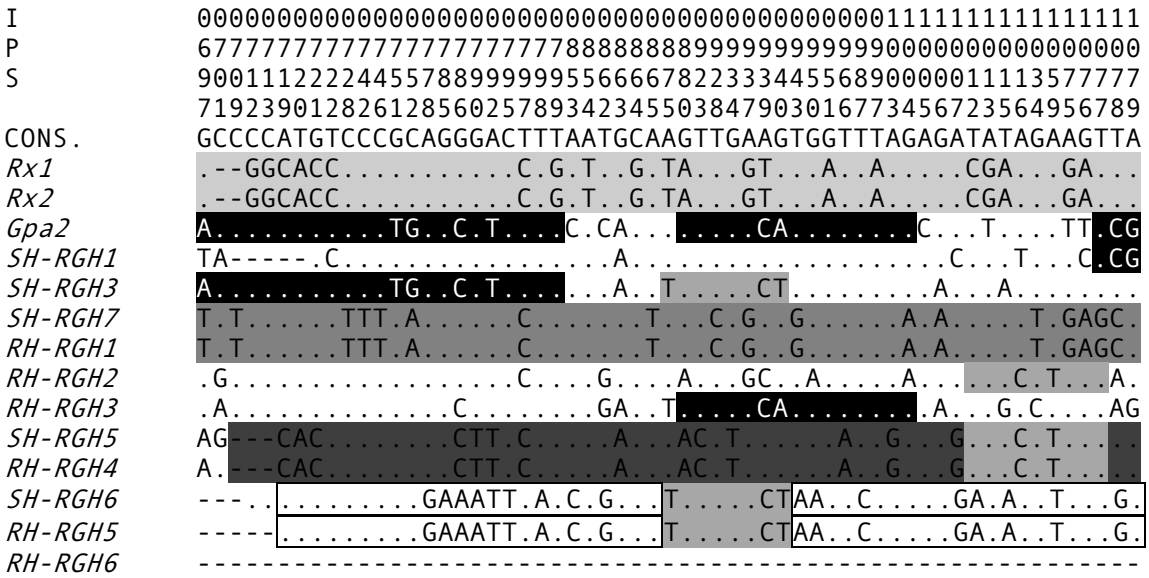


Figure 3 Similarity dendrogram of the C-terminal part of the *Gpa2/Rx1* homologues. Branch lengths are indicative for sequence similarity and bootstrap values >65 are indicated.

In addition, a nucleotide alignment was made to determine the Informative Polymorphic Sites (IPS; Parniske *et al.* 1997). At an IPS in a nucleotide alignment of fourteen *Gpa2/Rx1* homologues, two or more of the RGHs share the same nucleotide, while they are different from the nucleotides at the other RGHs. Positions that carry thirteen or fourteen identical nucleotides are therefore considered not to be informative. The IPS analysis showed that at certain positions, different RGHs share the same sequence resulting in a mosaic pattern of identical sequences (Fig. 4). For example, the *Gpa2* sequence is identical to *SH-RGH1, SH-RGH5* and *RH-RGH4* from position +9 to +17, but from position +98 to +170 *Gpa2* is identical to *Rx1* and *Rx2*. Than, from +346 to +499, *Gpa2* is identical to *RH-RGH2 and* from +610 to +686 and from +697 to +799, *Gpa2* is identical to *SH-RGH3.* From +880 to +1004 *Gpa2* is identical to *RH-RGH3 and* finally, *Gpa2* is identical again to *SH-RGH1* from +1083 to +1159. A similar patchwork pattern was observed for the sequences of the other homologues.


I
P
S
CONS.
Rx1
Rx2
Gpa2
SH-RGH1
SH-RGH3
SH-RGH7
RH-RGH1
RH-RGH2
RH-RGH3
SH-RGH5
RH-RGH4
SH-RGH6
RH-RGH5
RH-RGH6


I
P
S
CONS.
Rx1
Rx2
Gpa2
SH-RGH1
SH-RGH3
SH-RGH7
RH-RGH1
RH-RGH2
RH-RGH3
SH-RGH5
RH-RGH4
SH-RGH6
RH-RGH5
RH-RGH6

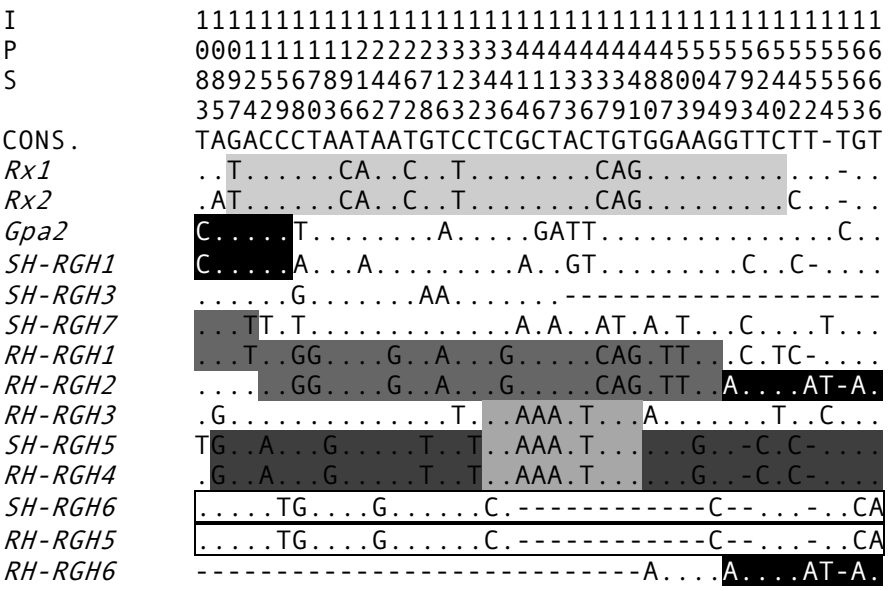


Figure 4 This alignment shows Informative Polymorphic Sites (IPS), which are the nucleotide positions were 2 or more nucleotides are different from the other nucleotides at the same positions. At the top of the alignment, the IPS position and the consensus sequence is shown. Dots indicate a nucleotide identical to the consensus sequence and dashes indicate sequence gaps. Stretches of sequence that are identical have the same shade or are framed, resulting in a patchwork of sequence similarities.

In the multiple nucleotide alignment of the LRR domains of the RGHs, 13 regions with InDels could be identified. De distribution of InDels displayed a similar patchwork of common InDels as was obtained from the IPS analysis (Fig. 5). Note that the 71, 126 and 158 bp deletions in *SH-RGH6* and *RH-RGH5* and the 834 bp deletion in *RH-RGH6* are not included. Remarkably, InDels present in exons are mostly multiples of three and thus maintaining the reading frame, while InDels in introns do not. The only exception is the region just after intron I where an 11 bp insertion in the *Gpa2* sequence and a 4 bp insertion in the *SH-RGH7* sequence resulted in a frame shift, followed by an InDel that is not a multiple of three for all RGHs except for *SH-RGH7.* As a result, all RGHs continue in the same frame, except *Gpa2*.


Rx1
Rx2
Gpa2
SH-RGH1
SH-RGH3
SH-RGH7
RH-RGH1
RH-RGH2
RH-RGH3
SH-RGH5
RH-RGH4
SH-RGH6
RH-RGH5
RH-RGH6

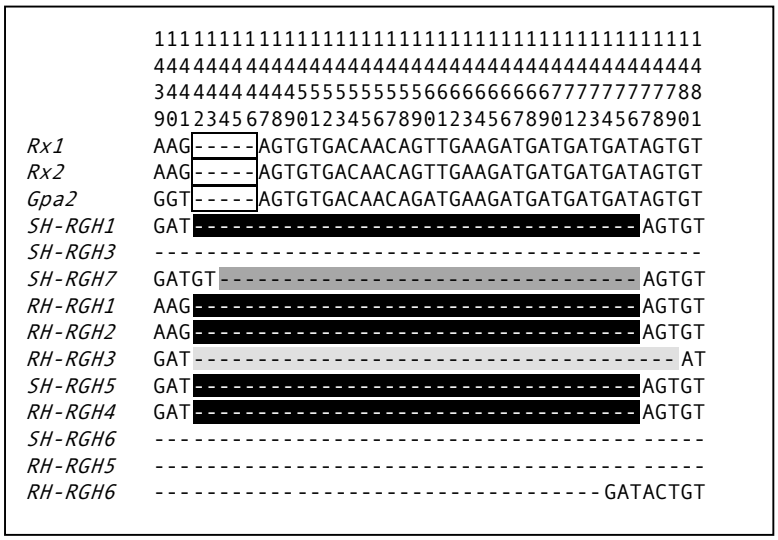
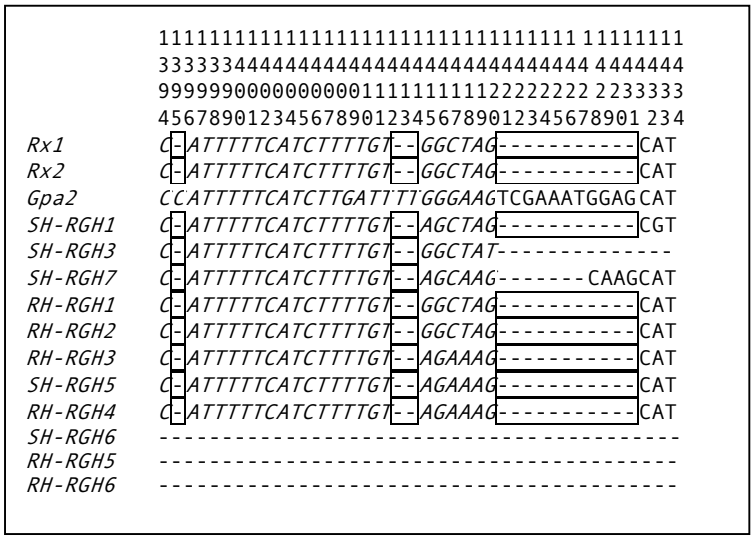
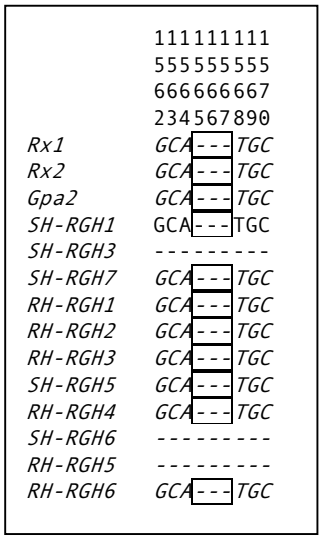
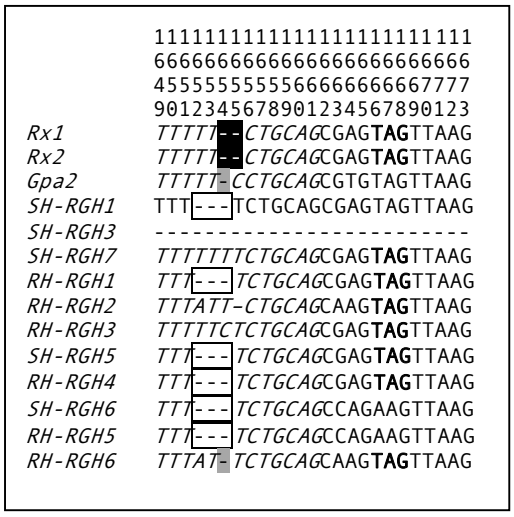
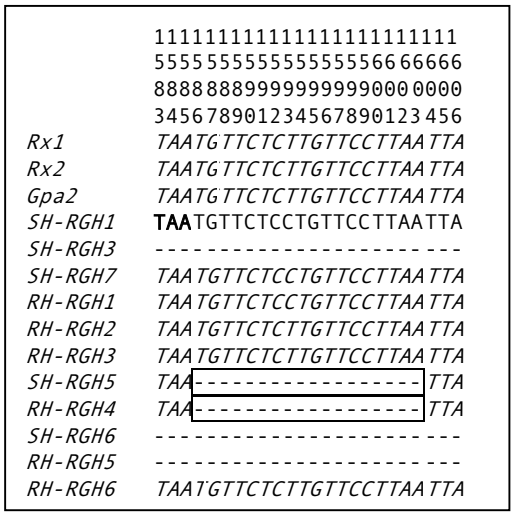


Figure 5 Overview of InDels present in the multiple nucleotide alignment of the 5' half *Gpa2/Rx1* homologues. Parts of the alignment containing InDels are displayed in boxes. Deletions are represented by dashes and common deletions are indicated with a frame or a pattern, showing a patchwork of identical InDel positions. Intron sequences are marked in italics and stop codons in bold.









Discussion

The structural relationship of the C-terminal part of fourteen *Gpa2/Rx1* homologues has been investigated to get more insight in the role of the LRR domain and the acidic tail in *R* gene specificity and signalling. Comparison of the nucleotide and amino acid sequences revealed patterns that might reflect the evolutionary mechanisms of *R* genes.

Gene structure
From a multiple nucleotide alignment produced for all fourteen *Gpa2/Rx1* homologues, it was clear that intron splice-site positions are very conserved among the members of the *Gpa2/Rx1* family, while the introns themselves are variable in size. This is consistent with the findings of Graham *et al.* (2002). They isolated nine genes that appear to be full-length resistance genes with homology to the TIR-NBS-LRR family. The genes have three introns at identical positions but with considerable size differences (Intron I: 95-196, Intron II: 155-356, Intron III: 85-219). *In silico* splicing of the introns resulted in the identification of the putative open reading frames of all *Gpa2/Rx1* homologues, except *SH-RGH3* and *8* (see Bakker *et al.* 2003). The presence of a putative open reading frame indicates that these homologues may encode for R proteins with unknown specificity. For pseudogenes it is expected that mutations result in a disrupted reading frame with multiple stop codons as was found for *SH-RGH3.* To ensure that the putative open reading frames of *Gpa2/Rx1* homologues *SH-RGH1* and *5* and *RH-RGH1-6* are correct and expressed we are currently isolating the 5′ genomic sequences and the cDNA sequences of all RGHs. *Taq*I digestion patterns of *Gpa2/Rx1* cluster specific PCR products derived from an RH cDNA pool revealed the patterns that match the *in silico Taq*I digestion pattern of *RH-RGH1* and *RH-RGH2, RH-RGH4* and/or *RH-RGH5* and *RH-RGH6* (unpublished results). This indicates that all six *Gpa2/Rx1* homologues of RH, including *RH-RGH6* that has a 834 bp deletion in the LRR domain, are transcribed.

The number of *R* gene homologues containing full-length ORFs differs substantionaly per *R* gene cluster. In both the tomato *Cf4* and *Cf9* clusters, all five RGHs encode full-length open reading frames (Parniske and Jones 1999). In the *RPP5* cluster of *Arabidopsis*, however, only *RPP5* out of ten RGHs is predicted to have a full-length open reading frame (Noel *et al.* 1999) and the *Xa21* cluster in rice contains seven *R* gene homologues (including *Xa21*), of which five are truncated (Song *et al.* 1997). The *Rp3* cluster in maize consists of five paralogues. Four of the *Rp3* homologues show a full-length open reading frame and one appears to be a pseudogene due to a reading frame disruption by a retrotransposon (Webb *et al.* 2002). Interestingly, removal of the retrotransposon restores the open reading frame. We have amplified a cDNA fragment for *SH-RGH6*, a *Gpa2/Rx1* homologue that has a 71 bp, a 126bp and a 158 bp deletion in the LRR domain and may be considered to be

a truncated gene. Other truncated genes have been reported to be expressed. For instance in an *R* gene cluster in soybean cDNA has been identified for three truncated homologues (Graham *et al.* 2000; Graham *et al.* 2002). One homologue consists of a complete TIR domain, a truncated NB domain and a single LRR, the other two truncated homologues contain a complete TIR domain followed by sequences with homology to a Ca^{2+} binding domain (Graham *et al.* 2002). In addition, all nine genes of the *Rp1* cluster in maize, including one homologue that is truncated, are transcribed (Sun *et al.* 2001).

Acidic tail

A major difference between *Rx1/Rx2* and *Gpa2* is the C-terminal acidic tail consisting of two direct sequence repeats (Bendahmane *et al.* 1999) that is present in both *Rx1* and *Rx2*, but absent in *Gpa2*. The acidic tail is unique for *Rx1* and *Rx2* and has not been identified in other *R* genes (Bendahmane *et al.* 1999; Bendahmane *et al.* 2000). It is striking that among the thirteen RGHs that could be translated into undisrupted putative open reading frames none have both repeats of the acidic tail. The role of acidic tails is not clear, but between *Gpa2* and the *Rx* genes a difference in resistance reaction has been observed; *Gpa2* triggers a very mild and slow response, while *Rx1* and *Rx2* trigger a very rapid extreme resistance response (Ritter *et al.* 1991; Bakker *et al.* unpublished data). The presence or absence of the acidic tail could play a role in this difference (Bendahmane *et al.* 2000; Bendahmane *et al.* 1999) and it is interesting to know what the function of a single repeat is as observed for a subset of homologues of which at least one is expressed (*SH-RGH7*).

The LRR domain

Comparing the thirteen homologues that could be translated into a putative amino acid sequence revealed 143 changes compared to the 440 aa consensus sequence that are most likely due to nucleotide point mutations accumulated over time. Major deletions have occurred in *SH-RGH6/RH-RGH5* and *RH-RGH6* that resulted in the loss of one and six LRRs, respectively. This altered number of repeats did not change the reading frames and could result in a different specificity, in a complete loss of function or even in a gain of specificity. Dixon *et al.* (1998) identified a gene that is very homologues to *Cf-5*, but has two additional LRRs and was not functional against *Cladosporium fulvum*.

Analysis of the complete nucleotide fragments of all fourteen *Gpa2/Rx1* homologues revealed a mean K_a/K_s ratio of 2.69. This ratio is well above one and therefore the C-terminal region of the *Gpa2/Rx1* homologues that harbours the LRR domain is under positive selection pressure. Analyses of the LRR β-strand/β-turn motifs resulted in a mean K_a/K_s ratio of even 3.33. These findings support the idea that the LRR domain and more precisely the solvent exposed residues of the LRR β-strand/β-turn motifs are involved in *R* gene specificity. The strong positive selection pressure that acts on these regions is an indication that diversifying selection is one of the mechanisms

underlying the development of new *R* gene specificities. For the regions outside the LRR β-strand/β-turn motifs the K_a/K_s ratios usually are smaller than one (Noel *et al.* 1999; Parniske *et al.* 1997; Graham *et al.* 2002).

In this study, however, in several cases, the K_a/K_s ratios of the LRR domain sequences outside the LRR β-strand/β-turn motif of the SH homologues were significantly larger than one (data not shown). It was expected that these ratios did not significantly differ from one, which would be consistent with the findings of Van der Vossen *et al.* (2000) for *Gpa2* and *Rx1*. The poor LRR structure of the *Gpa2/Rx1* homologues could be an explanation for these large K_a/K_s ratios. Palomino *et al.* (2002) have employed a maximum likelihood-based analysis to identify single amino acid positions that are subjected to diversifying selection among NBS-LRR genes in *Arabisopsis*. In two groups of sequences they identified such amino acid positions in regions between the LRR β-strand/β-turn motifs. They suggest that, based on the assumption that these amino acids are involved in protein-protein interaction, these findings may help to further elucidate the three-dimensional structure of these LRR domains. In addition, for three alleles of the *RPP13* locus, these regions also seem to be under positive selection pressure(Bittner-Eddy *et al.* 2000). They suggested that this region may be as important for generating new specificities in the *RPP13* cluster as the LRR β-strand/β-turn motifs themselves.

The pseudogene *SH-RGH3* is also under positive selection pressure. This was also found for the pseudogenes in an *R* gene cluster in soybean (Graham *et al.* 2002). *SH-RGH3* is the only sequenced pseudogene in the *Gpa2/Rx1* cluster and therefore, the K_a/K_s ratio could not be determined for the comparison of two pseudogenes. If a K_a/K_s ratio for two pseudogenes is still greater than one, it might be an indication that at least one of these pseudogenes had until recently been a functional gene. This could also be the case for *SH-RGH3,* but the high K_a/K_s ratio could also be explained by the diversification of the comparing partner.

Diversifying selection has been shown for resistance genes that reside in simple loci (Caicedo *et al.* 1999; Bittner-Eddy *et al.* 2000) as well as for genes that reside in complex loci (Parniske *et al.* 1997; Noel *et al.* 1999; McDowell *et al.* 1998). The K_a/K_s ratio of the LRR β-strand/β-turn motifs is generally higher than the K_a/K_s ratio found for the complete LRR domain (Ellis *et al.* 1999; Parniske *et al.* 1997). Positive selection pressure resulted in differences in amino acid composition. As was expected from the K_a/K_s ratio analysis, most of the variable amino acid positions are present in the LRR β-strand/β-turn motifs. Some amino acid positions harbour even five or more different amino acids. These hyper-variable amino acid positions might harbour the residues directly involved in protein-protein interaction and therefore responsible for *R* gene specificity. If this is indeed the case, we do not expect differences in these amino acids among *R* genes that have the same specificity.

The two *R* genes *Rx1* and *Rx2* have the same specificity. They both confer extreme resistance to the same strain of the potato virus X and respond to the same elicitor, the coat protein (Bendahmane *et al.* 2000; Bendahmane *et al.* 1999). In contrast, the *R* gene *Gpa2* confers resistance to the potato cyst nematode *G. pallida* (Van der Vossen *et al.* 2000). Additionally, the extreme resistance triggered by *Rx1* and *Rx2* is very rapid and without symptoms (Bendahmane *et al.* 2000; Bendahmane *et al.* 1999), while the response of the *Gpa2* gene is very mild and slow (unpublished results). When the amino acids are compared between the three *R* genes at the variable and hyper-variable amino acid positions, we find that *Rx1* and *Rx2* are always identical at those positions, but that forty-six differences have been found between *Rx1/Rx2* and *Gpa2*. Apparently, (hyper)-variable amino acid positions are conserved among *R* genes with the same specificity. Based on this assumption, we expect the specificities of the (nearly) identical couples *SH-RGH5/RH-RGH4* and *SH-RGH6/RH-RGH5* to be the same, but also *SH-RGH7* and *RH-RGH1* have the same amino acids at the hyper-variable amino acid positions in the LRR β-strand/β-turn motifs and may therefore have the same specificity. The K_a/K_s ratio could not be calculated for *SH-RGH5* and *RH-RGH4* (no substitutions) and for *SH-RGH6* and *RH-RGH5* (one substitution (synonymous) and for the LRR β-strand/β-turn motifs of *SH-RGH7* and *RH-RGH1* (one substitution (asynonymous)). For the complete sequences of *SH-RGH7* and *RH-RGH1*, however, the K_a/K_s ratio does not significantly differ from one. This supports the hypothesis that they could have the same specificity, because the K_a/K_s ratio for the LRR β-strand/β-turn motifs of *Rx1* and *Rx2* could also not be calculated (no substitutions) and the ratio for the complete LRR does not differ significantly from 1.

A patchwork of sequence similarities is an indication of mosaic evolution

A physical map of the *Gpa2/Rx1* cluster (Bakker *et al.* 2003) combined with the sequence identities lead to the assumption of several 'orthologues' groups. The pairs *Gpa2/SH-RGH5, SH-RGH1/SH-RGH6,* and *Rx1/SH-RGH7* are assumed to be allelic based on their positions on the physical map. Although we only have the C-terminal parts, the couples *SH-RGH5/RH-RGH4, SH-RGH6/RH-RGH5* and *SH-RGH7/RH-RGH1* are assumed to be 'allelic' based on their sequence identities (99.9%, 98.5% and 100% respectively). This resulted in the groups *Gpa2/SH-RGH5/RH-RGH4, SH-RGH1/SH-RGH6/RH-RGH5* and *Rx1/SH-RGH7/RH-RGH1* that could be orthologues. The first homologue in an orthologues group is always derived from the SH *S. tuberosum* ssp. *andigena* haplotype, the second from the SH *S. tuberosum* ssp *tuberosum* haplotype and the third from the RH *S. tuberosum* ssp. *tuberosum* haplotype.

A similarity dendrogram has been produced for the *Gpa2/Rx1* homologues in order to determine which homologues are most related. This resulted, however, in a tree where most homologues could not be separated from the main group. A possible explanation for these phenomena is the exchange of sequences among the homologues during evolution, so called mosaic evolution (Parniske *et al.* 1997). When comparing the IPS of the *Gpa2/Rx1* homologues significant stretches of sequence

identity were found between members of different haplotypes as well as between members of the same haplotype. This results in a patchwork of sequence identities as was observed for *Rx1, Rx2* and *Gpa2* (Bendahmane *et al.* 2000) and other *R* gene clusters (Parniske *et al.* 1997; Noel *et al.* 1999; Ellis *et al.* 1999; McDowell *et al.* 1998). Like the IPS, the InDels were also organised in a patchwork. This in contrast to the InDels in the *Mla* cluster in barley, were the presence of shared InDels among *Mla* family members were an indication of common ancestry (Wei *et al.* 2002).

The exchange events between the *Gpa2/Rx1* homologues could be the result of unequal crossing-over or gene conversion. Interestingly, break points of the exchange events often coincide with hyper-variable amino acid positions, suggesting a role of unequal crossing-over and gene conversion for the generation of new *R* gene specificities. In addition, the number of paralogues on one of the haplotypes of RH is two, whereas the other three haplotypes carry four paralogues (Bakker *et al.* 2003). This is another indication of unequal crossing-over. Three pairs of homologues (*SH-RGH5/RH-RGH4, SH-RGH6/RH-RGH5* and *SH-RGH7/RH-RGH1*), however, have a very high similarity (99.9, 100.0 and 98.5% respectively). Therefore, these three couples may be orthologues that did not undergo a sufficient amount of unequal crossing-over and gene conversion to cause concerted evolution. As for the *P* locus in flax (Dodds *et al.* 2001) the *Gpa2/Rx1* homologues all show a high degree of nucleotide sequence identity (apart from the severely truncated *SH-RGH3* all >90% in the most variable part of the genes) and therefore we think that on the *Gpa2/Rx1* locus, the birth and death model proposed by Michelmore and Meyers (1998) can still be applicable, but unequal crossing-over and gene conversion may be more important in the evolution of new *R* gene specificities at this locus than at the *Dm3* locus.

Chapter 6

Summary and concluding remarks

Parts of this chapter have been accepted for publication in a modified form in:

Erin Bakker, Jeroen Rouppe van der Voort, Herman van Eck, Ronald Hutten, Jaap Bakker and Aska Goverse
Molecular markers for resistance to cyst nematodes in potato.
in: Proceedings of the Fourth International Congress of Nematology

Resistance to the potato cyst nematodes *G. pallida* and *G. rostochiensis* is an important aspect in potato breeding. To gain insight in the structure of the genes underlying nematode resistance, it is necessary to first map nematode resistance loci, as described for the resistance loci *Gpa5* and *Gpa6* (Chapter 2). Subsequently, both for map-based cloning and for marker-assisted selection, a high-resolution map of the resistance locus is required. In Chapter 3, such a high-resolution map is produced for the *H1* locus. To gain more insight in the genomic organisation of a complex disease resistance locus, *Gpa2/Rx1* homologues derived from different haplotypes have been characterised (Chapter 4). Finally, the *Gpa2/Rx1* homologues derived from five haplotypes have been subjected to extensive sequence analysis to elucidate their structural diversity and evolutionary relationships (Chapter 5).

Marker assisted selection

A useful tool for pre-selection of potentially interesting genotypes for PCN resistance breeding in potato is marker-assisted selection (MAS). MAS considerably accelerates the selection process because segregating populations can be screened with molecular markers instead of the labour-intensive and time consuming PCN resistance assays. Moreover, selection for desired traits could already take place in the seedling stage. In addition, QTLs may be difficult to assess phenotypically, but the use of molecular markers eliminates environmental effects resulting in a more effective selection. For common bean, for instance, it has been reported that MAS for drought tolerance improved yield by 11% under stress conditions and 8% under non-stress conditions compared to conventional selection (Schneider *et al.* 1997). Finally, pyramiding genes in one cultivar that cannot be phenotypically distinguished is feasible with MAS.

For map-based cloning and MAS, a marker closely linked to the resistance gene is indispensable with the *R* gene itself as the ultimate marker. For this, a dense genetic map is required of the region where the resistance gene is located. Bulked segregant analysis is often employed for the generation of markers linked to the locus of interest. For instance for the high-resolution map of *Gro1,* 144 AFLP primer combinations resulted in 1 marker/cM in the *Gro1* region (Ballvora *et al.* 1995) and for the high-resolution map of the *Gpa2/Rx1* locus, even 728 AFLP primer combinations were used, which resulted in 2 markers/cM in the *Gpa2/Rx1* region (Bendahmane *et al.* 1997). Markers linked to a resistance gene locus can also be found with the aid of the AFLP catalogue consisting of chromosome-specific AFLP markers (Rouppe van der Voort *et al.* 1998a). This AFLP catalogue facilitated the localisation of the resistance loci *Gpa2*, *Grp1*, *Gpa3*, *Gpa5* and *Gpa6* (Rouppe van der Voort *et al.* 1997a; Rouppe van der Voort *et al.* 1998b; Chapter 2). Additionally, an Ultra High Density (UHD) genetic map of potato has been produced of the potato population SH×RH that has global marker saturation. AFLP analysis with 387 primer combinations resulted in 10,365 segregating AFLP loci in 1,118 BINs (groups of

markers that cannot be separated within the population used to construct the map). The results are available on the URL http://www.dpw.wageningen-ur.nl/uhd/.

To illustrate the benefit from the UHD genetic map, the high-resolution map of the *H1* locus can be used as an example. In a bulked segregant analysis, 704 AFLP primer combinations resulted in thirteen candidate markers linked to the *H1* locus (Chapter 3). When these candidate markers were implemented in the high-resolution map of the *H1* locus, however, only one marker was located at a distance of less than 1 cM. Mapping the *H1* locus in the UHD genetic map, which is produced with 387 AFLP primer combinations, the local saturation is 4 AFLP markers per cM. In potato, many more resistance loci will need local marker saturation, as a first step towards the isolation of the underlying genes and the UHD genetic map will be a useful tool to accomplish this. In the same way that a UHD genetic map can be beneficial for future efforts in local marker saturation, a physical map of the potato genome can be beneficial for future BAC contig construction and cloning the resistance gene of interest.

Lateral cloning
Many interesting PCN resistance loci are present in wild relatives of potato for which no UHD genetic map and BAC library are available. Nevertheless, the genetic and physical map of SH can be used as a standard in potato genetics. Previous research already indicated that the genomes of different potato species are highly similar and that co-migrating AFLP markers usually represent similar genomic regions (Rouppe van der Voort *et al.* 1997b). Even in the related genera *Lycopersicon* (tomato) and *Capsicum* (pepper), *R* gene clusters are found in homologous regions (Grube *et al.* 2000).

This allows lateral cloning of disease resistance genes by using the sequence information of a region harbouring RGHs in a potato genotype of one species for the cloning of resistance genes in other (related) species. Figure 1 illustrates the principle of lateral cloning of *Gpa3* from *S. tarijense* using the genetic map and BAC library of SH. *Gpa3* has been located at the short arm of chromosome XI in *S. tarijense.* In the first step, AFLP markers will be selected from the corresponding region in SH. From the SH BAC library, BAC clones will be selected that harbour one or more of the selected AFLP markers. Subsequently, a physical map will be constructed from these BAC clones. Using degenerated primers that anneal to conserved *R* gene regions, the BAC clones will be screened for *R* gene homologues (as proposed by Leister *et al.* 1996). These homologues will then be sequenced and with the sequence information of the *R* gene homologues, cluster specific primers will be designed. Finally, the cluster specific primers will be used to screen genomic DNA of *S. tarijense* and the resulting *R* gene homologues will be sequenced. Homologues encoding full-length open reading frames will be transformed to a susceptible potato clone for complementation assays to select the gene of interest in a nematode resistance test.

The *Gpa2/Rx1* cluster was used in an effort to proof that cluster specific primers could be designed for a complex resistance gene locus (Chapter 4). It was shown that one primer pair that anneals to the LRR was sufficient to amplify all the homologues in the *Gpa2/Rx1* cluster on chromosome XII in both SH and RH. This primer pair did not amplify *R* gene homologues elsewhere in the genome. Currently, a cluster specific primer pair is used to amplify *Gpa2/Rx1* homologues from other *Solanum* species (Butterbach *et al.*, unpublished data).

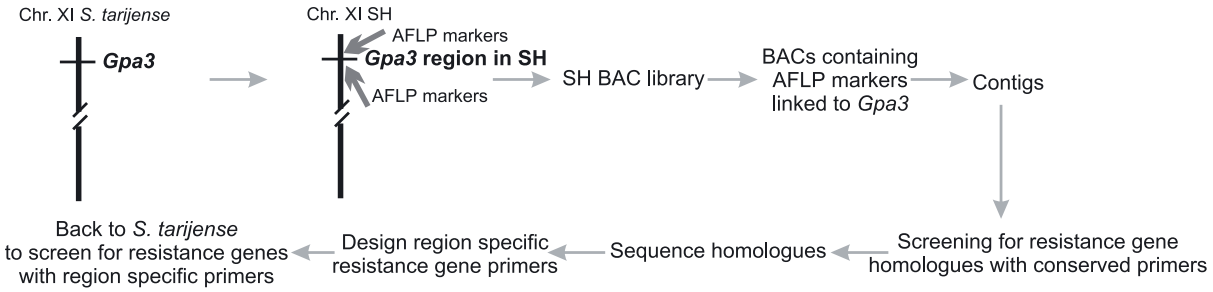


Figure 1 Lateral cloning strategy to isolate *Gpa3* from *Solanum tarijense* using the UHD genetic map and BAC library that are available for SH.

Hot spots of resistance

Various resistance loci against fungi, viruses, insects and nematodes have been mapped on the potato genome (reviewed by Gebhardt and Valkonen 2001; Fig. 2). Interestingly, many PCN resistance loci map to a region where at least one other single dominantly inherited resistance gene is present. These so called hot spots of resistance harbour genes that often confer resistance to unrelated pathogens. The most striking example is the nematode resistance gene *Gpa2* that is located in the same cluster as the virus resistance gene *Rx1*. The two genes are present at the same haplotype separated by 115 kb (Van der Vossen *et al.*, 2000). But also the *Gpa5* locus is located on a chromosomal region where two nematode (*Gpa* and *Grp1*), two virus (*Rx2* and *Nb*) and one fungus resistance locus (*R1*) have been mapped, respectively (Chapter 2). The *Gpa3* locus is located in the vicinity of the root-knot nematode resistance locus *Rmc1*, the fungus resistance locus *Sen1* and the two viral resistance loci N*a* and *Ry* (Gebhardt and Valkonen, 2001; Van Eck *et al.*, unpublished results).

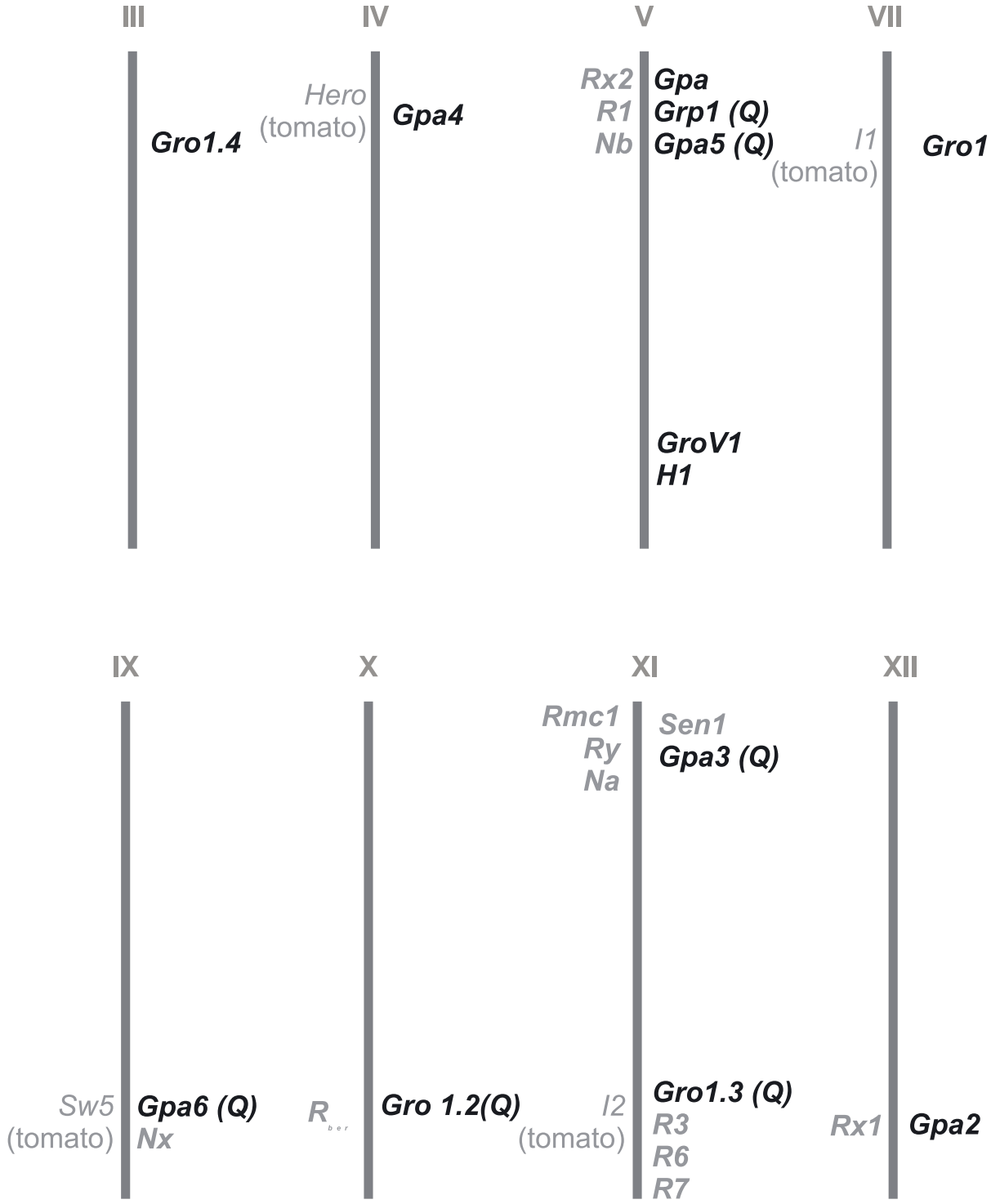


Figure 2 (After Gebhardt and Valkonen 2001) Schematic drawing of hot spots for resistance on potato chromosomes harbouring PCN resistance genes. Vertical lines depict the chromosomes with the chromosome number at the top. PCN resistance genes are typed in black. Other resistance genes are in grey.

Many quantitative trait loci that confer resistance to PCN are located in hot spots for resistance. Theoretically, the quantitative behaviour of QTLs for PCN resistance can be caused by partial resistance of the host plant. Another option is that quantitative resistance is mediated by single *R* genes and that the quantitative effect is caused by variation in the PCN population used in the resistance assay. For instance, using the Pa3-Rookmaker population, the quantitative resistance locus *Gpa6* is additive to the *Gpa5* locus (Chapter 2). PCN are out-crossing species and hence, resistance tests are performed with a genetic mixture of virulent and avirulent individuals. The latter explanation complies with the locations of most QTLs, which are often linked to single dominantly inherited *R* genes. In contrast to PCN, root-knot nematodes such as *Meloidogyne incognita*, *M. javanica* and *M. arenaria* reproduce via mitotic parthenogenesis. Therefore, resistance assays for root-knot nematode resistance can be performed with genetically identical populations. The fact that all root-knot nematode resistances are conferred by single dominant genes supports the idea that the quantitative behaviour of the potato cyst nematode *R* genes is caused by the mixture of virulent and avirulent genotypes in the potato cyst nematode populations. To proof this hypothesis, further characterisation of the genes underlying the resistance loci *Grp1, Gpa3, Gpa5* and *Gpa6* is needed.

Identification of R genes for PCN resistance

Nematode resistance genes have been isolated from different crop species, like for instance the beet cyst nematode *R* gene *Hs1^{pro}* from beet (Cai *et al.* 1997), the root knot nematode *R* gene *Mi1* from tomato (Milligan *et al.* 1998; Vos *et al.* 1998) and the potato cyst nematode *R* gene *Hero* from tomato (Ernst *et al.* 2002). So far, two nematode *R* genes have been cloned from potato: the *Gpa2* gene that confers resistance to *G. pallida* (Van der Vossen *et al.* 2000) and the *Gro1* gene that confers resistance to *G. rostochiensis* (Gebhardt *et al.*, personal communication).

All nematode *R* genes, except *Hs1^{pro}*, share structural motifs with *R* genes that confer resistance to other plant pathogens like bacteria, fungi, oomycetes and viruses. This shows that nematode *R* genes are part of the plant survey system that results in the activation of a defence response upon infection. Therefore, it is assumed that the same signal transduction pathways are involved in nematode resistance as in resistance to other pathogens. The high-resolution map of the *H1* locus (Chapter 3) allows the identification of the *H1* resistance gene. This will help us to further elucidate the mechanism underlying the *H1* resistance. The *H1* resistance response, which triggers an HR, is rapid. Furthermore, an avirulent nematode line (Ro_1-19) and a virulent nematode line (Ro_1-22) are available (Janssen *et al.* 1990). Therefore, the *H1-AvrH1* interaction would be a suitable model system to study the mechanisms underlying potato cyst nematode resistance.

Complex R gene loci

Most *R* genes reside in complex loci (reviewed by Hulbert *et al.*, 2001). It is often reported that members of the same complex *R* gene cluster confer resistance to different pathotypes of the same pathogen species. For instance the *P* locus in flax carries at least six different rust specificities at different haplotypes (Dodds *et al.* 2001) and the *Dm* locus harbours at least ten genes conferring resistance to *Bremia lactucae* (Meyers *et al.* 1998). The *Gpa2/Rx1* cluster in potato, however, harbours two genes at the same haplotype that confer resistance to completely different organisms, a virus and a nematode (Van der Vossen *et al.*, 2000). In Chapter 4, nine *Gpa2/Rx1* homologues have been identified from three additional haplotypes and cDNA analysis revealed that at least six of them are expressing putative *R* genes with unknown specificity (Chapter 5).

Sequence analysis of *R* gene clusters will increase our understanding of the evolutionary history of *R* genes. The analysis of the *Gpa2/Rx1* homologues derived from 5 different haplotypes (Chapter 5) revealed that new *R* gene specificities may have evolved through positive selection in combination with crossing-over and gene conversion. This complies with the findings of Parniske *et al.* (1997) at the *Cf-4/9* cluster and Meyers *et al.* (1998) at the *Dm3* cluster. They both found evidence for positive selection acting on the solvent exposed regions of the LRR domains and a patchwork of sequence similarities, which points at unequal crossing-over and gene conversion. Comparing members of the same *R* gene cluster, especially in the case of *Gpa2* and *Rx1,* where two highly similar *R* genes confer resistance to completely different organisms, could also lead to a better understanding of their role in the activation of a resistance response and how the different structural domains are involved. Structure-function analyses, however, are required to test the hypotheses that result from such *in silico* comparisons of *R* gene homologues.

Literature cited

Axtell MJ, McNellis TW, Mudgett MB, Hsu CS, Staskawicz BJ (2001) Mutational analysis of the Arabidopsis *RPS2* disease resistance gene and the corresponding *Pseudomonas syringae avrRpt2* avirulence gene. Molecular Plant-Microbe Interactions 14: 181-188

Baker B, Zambryski P, Staskawicz B, Dinesh-Kumar SP (1997) Signaling in plant-microbe interactions. Science 276: 726-733

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. Theoretical and Applied Genetics 106: 1524-1531

Bakker J, Bouwman-Smits L, Gommers FJ (1992) Genetic relationships between *Globodera pallida* pathotypes in Europe assessed by using 2-dimensional gel-electrophoresis of proteins. Fundamental and Applied Nematology 15: 481-490

Bakker J, Folkertsma RT, Rouppe van der Voort JNAM, De Boer JM, Gommers FJ (1993) Changing concepts and molecular approaches in the management of virulence genes in potato cyst nematodes. Annual Review of Phytopathology 31: 169-190

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 Journal 30: 361-371

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*. Molecular & General Genetics 249: 82-90

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*. Molecular & General Genetics 224: 177-182

Bateson W (1899) Hybridisation and cross-breeding as a method of scientific investigation. Journal of the Royal Horticultural Society 24: 59-66

Bendahmane A, Farnham G, Moffett P, Baulcombe DC (2002) Constitutive gain-of-function mutants in a nucleotide binding site-leucine rich repeat protein encoded at the Rx locus of potato. Plant Journal 32: 195-204

Bendahmane A, Kanyuka K, Baulcombe DC (1997) High-resolution genetical and physical mapping of the *Rx* gene for extreme resistance to potato virus X in tetraploid potato. Theoretical and Applied Genetics 95: 153-162

Bendahmane A, Kanyuka K, Baulcombe DC (1999) The *Rx* gene from potato controls separate virus resistance and cell death responses. The Plant Cell 11: 781-791

Bendahmane A, Querci M, Kanyuka K, Baulcombe DC (2000) Agrobacterium transient expression system as a tool for the isolation of disease resistance genes: application to the *Rx2* locus in potato. The Plant Journal 21: 73-81

Bittner-Eddy PD, Crute IR, Holub EB, Beynon JL (2000) *RPP13* is a simple locus in *Arabidopsis thaliana* for alleles that specify downy mildew resistance to different avirulence determinants in *Peronospora parasitica*. The Plant Journal 21: 177-188

Blok VC, Phillips MS, Harrower BE (1997) Comparison of British populations of potato cyst nematodes with populations from continental Europe and South America using RAPDs. Genome 40: 286-293

Bonierbale MW, Plaisted RL, Pineda O, Tanksley SD (1994) QTL analysis of trichome-mediated insect resistance in potato. Theoretical and Applied Genetics 87: 973-987

Botella MA, Coleman MJ, Hughes DE, Nishimura MT, Jones JDG, Somerville SC (1997) Map positions of 47 Arabidopsis sequences with sequence similarity to disease resistance genes. The Plant Journal 12: 1197-1211

Bradshaw JE, Meyer RC, Milbourne D, McNicol JW, Phillips MS, Waugh R (1998) Identification of AFLP and SSR markers associated with quantitative resistance to *Globodera pallida* (Stone) in tetraploid potato (*Solanum tuberosum* subsp. *tuberosum*) with a view to marker-assisted selection. Theoretical and Applied Genetics 97: 202-210

Brommonschenkel SH, Frary A, Frary A, Tanksley SD (2000) The broad-spectrum tospovirus resistance gene *Sw-5* of tomato is a homolog of the root-knot nematode resistance gene *Mi*. Molecular Plant-Microbe Interactions 13: 1130-1138

Brommonschenkel SH, Tanksley SD (1997) Map-based cloning of the tomato genomic region that spans the Sw-5 tospovirus resistance gene in tomato. Molecular & General Genetics 256: 121-126

Bubeck DM, Goodman MM, Beavis WD, Grant D (1993) Quantitative Trait Loci Controlling Resistance to Gray Leaf- Spot in Maize. Crop Science 33: 838-847

Cai D, Kleine M, Kifle S, Harloff H-J, Sandal NN, Marcker KA, Klein-Lankhorst RM, Salentijn EMJ, Lange W, Stiekema WJ, Wyss U, Grundler FMW, Jung C (1997) Positional cloning of a gene for nematode resistance in sugar beet. Science 275: 832-834

Caicedo AL, Schaal BA, Kunkel BN (1999) Diversity and molecular evolution of the *RPS2* resistance gene in *Arabidopsis thaliana*. Proceedings of the National Academy of Science 96: 302-306

DeJong W, Forsyth A, Leister D, Gebhardt C, Baulcombe DC (1997) A potato hypersensitive resistance gene against potato virus X maps to a resistance gene cluster on chromosome 5. Theoretical and Applied Genetics 95: 246-252

Dixon MS, Hatzixanthis K, Jones DA, Harrison K, Jones JDG (1998) The tomato *Cf-5* disease resistance gene and six homologs show pronounced allelic variation in leucine-rich repeat copy number. The Plant Cell 10: 1915-1925

Dodds PN, Lawrence GJ, Ellis JG (2001) Six amino acid changes confined to the leucine-rich repeat b-strand/b-turn motif determine the difference between the *P* and *P2* rust resistance specificities in flax. The Plant Cell 13: 163-178

Dunnet JM (1961) Inheritance of resistance to potato root-eelworm in breeding lines stemming from *Solanum multidissectum* Hawkes. Annual Report Scottisch Plant Breeding Station: 39-46

Ellenby C (1952) Resistance to the potato root eelworm, *Heterodera rostochiensis* Wollenweber. Nature 170: 1016

Ellis J, Dodds P, Pryor T (2000) Structure, function and evolution of plant disease resistance genes. Current Opinion in Plant Biology 3: 278-284

Ellis JG, Lawrence GJ, Luck JE, Dodds PN (1999) Identification of regions in alleles of the flax rust resistance gene *L* that determine differences in gene-for-gene specificity. The Plant Cell 11: 495-506

Ernst K, Kumar A, Kriseleit D, Kloos D-U, Phillips MS, Ganal MW (2002) The broad-spectrum potato cyst nematode resistance gene (*Hero*) from tomato is the only member of a large gene family of NBS-LRR genes with an unusual amino acid repeat in the LRR region. Plant Journal in press.

Evans K (1993) New Approaches for Potato Cyst-Nematode Management. Nematropica 23: 221-231

Evans K, Stone AR (1977) A review of the distribution and biology of the potato cyst nematodes *Globodera rostochiensis* and *G. pallida*. PANS 23: 178-189

Flor HH (1942) Inheritance of pathogenicity in *Melampsora lini*. Phytopathology 32: 653-669

Flor HH (1956) The complementary genetic systems in flax and flax rust. Advances in Genetics 8: 29-54

Flor HH (1971) Current status of the gene-for-gene concept. Annual Review of Phytopathology 9: 275-296

Folkertsma RT, De Groot KE, Van Koert PHG, Van Gent-Pelzer MPE, Rouppe van der Voort JNAM, Schots A, Bakker J, Gommers FJ, Helder J (1996a) Cluster analysis of 36 *Globodera pallida* field populations using two sets of molecular markers. European Journal of Plant Pathology 102: 577-584

Folkertsma RT, Spassova MI, Prins M, Stevens MR, Hille J, Goldbach RW (1999) Construction of a bacterial artificial chromosome (BAC) library of *Lycopersicon esculentum* cv. Stevens and its application to physically map the *Sw-5* locus. Molecular Breeding 5: 197-207

Folkertsma RT, Van derVoort J, De Groot KE, Van Zandvoort PM, Schots A, Gommers FJ, Helder J, Bakker J (1996b) Gene pool similarities of potato cyst nematode populations assessed by AFLP analysis. Molecular Plant-Microbe Interactions 9: 47-54

Ganal MW, Czihal R, Hannappel U, Kloos DU, Polley A, Ling HQ (1998) Sequencing of cDNA clones from the genetic map of tomato (*Lycopersicon esculentum*). Genome Research 8: 842-847

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. Theoretical and Applied Genetics 85: 541-544

Gebhardt C, Valkonen JPT (2001) Organization of genes controlling disease resistance in the potato genome. Annual Review of Phytopathology 39: 79-102

Geiger HH, Heun M (1989) Genetics of quantitative resistance to fungal diseases. Annual Review of Phytopathology 27: 317-341

Graham MA, Marek LF, Lohnes D, Cregan P, Shoemaker RC (2000) Expression and genome organization of resistance gene analogs in soybean. Genome 43: 86-93

Graham MA, Marek LF, Shoemaker RC (2002) Organization, expression and evolution of a disease resistance gene cluster in soybean. Genetics 162: 1961-1977

Grube RC, Radwanski ER, Jahn M (2000) Comparative genetics of disease resistance within the *Solanaceae*. Genetics 155: 873-887

Hammond-Kosack KE, Jones JDG (1997) Plant disease resistance genes. Annual review of Plant Physiology and Molecular Biology 48: 575-607

Hawkes JG (1990) The potato. Evolution, biodiversity and genetic resources. Belhaven, London

Hermsen JGT, Verdenius J (1973) Selection from *Solanum tuberosum* Group Phureja of genotypes combining high-frequency haploid induction with homozygosy for embryo-spot. Euphytica 22: 244-259

Hulbert SH (1997) Structure and evolution of the *rp1 c*omplex conferring rust resistance in maize. Annual Review of Phytopathology 35: 293-310

Hulbert SH, Webb CA, Smith SM, Sun Q (2001) Resistance gene complexes: Evolution and utilization. Annual Review of Phytopathology 39: 285-312

Hutten RCB, Scholberg E, Huigen DJ, Hermsen JGT, Jacobsen E (1993) Analysis of Dihaploid Induction and Production Ability and Seed Parent X Pollinator Interaction in Potato. Euphytica 72: 61-64

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*. Molecular Breeding 2: 51-60

Jansen RC (1993) Interval Mapping of Multiple Quantitative Trait Loci. Genetics 135: 205-211

Jansen RC, Stam P (1994) High-Resolution of Quantitative Traits into Multiple Loci Via Interval Mapping. Genetics 136: 1447-1455

Janssen R, Bakker J, Gommers FJ (1990) Selection of lines of *Globodera rostochiensis* virulent and avirulent to the *H1* resistance gene in *Solanum tuberosum* ssp. *andigena* CPC 1673. Revue de Nematologie 13: 11-15

Janssen R, Bakker J, Gommers FJ (1991) Mendelian proof for a gene-for-gene relationship between virulence of Globodera rostochiensis and the H-1 resistance gene in Solanum tuberosum ssp. andigena CPC 1673. Revue de Nematologie 14: 207-212

Jones DA, Jones JDG (1997) The role of leucine-rich repeat proteins in plant defences. Advances in Botanical Research 24: 90-167

Jones FGW, Jones MGK (1974) Pests of field crops. 2nd ed. Edward Arnold, London, UK

Jones FGW, Parrott DM, Perry JN (1981) The gene-for-gene relationship and its significance for potato cyst nematodes and their Solanaceous hosts. In: Zuckerman BM, Rohde RA (eds) Plant Parasitic Nematodes. Academic Press, New York vol 3, pp 23-36

Jones MGK, Northcote DH (1972) Nematode-induced syncytium: A multinucleate transfer cell. Journal of Cell Science 10: 789-809

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. Theoretical and Applied Genetics 98: 679-689

Konieczny A, Ausubel FM (1993) A procedure for mapping Arabidopsis mutations using codominant ecotype-specific PCR-based markers. Plant Journal 4: 403-410

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*. Theoretical and Applied Genetics 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. Theoretical and Applied Genetics 92: 463-470

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121: 185-199

Lawrence GJ, Finnegan EJ, Ayliffe MA, Ellis JG (1995) The *L6* gene for flax rust resistance is related to the Arabidopsis bacterial resistance gene *Rps2* and the tobacco viral resistance gene *N*. Plant Cell 7: 1195-1206

Leister D, Ballvora A, Salamini F, Gebhardt C (1996) A PCR-based approach for isolating pathogen resistance genes from potato with potential for wide application in plants. Nature Genetics 14: 421-429

Leonards-Schippers C, Gieffers W, Salamini F, Gebhardt C (1992) The *R1* gene conferring race-specific resistance to *Phytophthora infestans* in potato is located on potato chromosome V. Molecular & General Genetics 233: 278-283

Leonards-Schippers C, Gieffers W, Schaefer PR, Ritter E, Knapp SJ, Salamini F, Gebhardt C (1994) Quantitative resistance to *Phytophthora infestans* in potato - a

case-study for QTL mapping in an allogamous plant-species. Genetics 137: 67-77

Li ZK, Luo LJ, Mei HW, Paterson AH, Zhao XZ, Zhong DB, Wang YP, Yu XQ, Zhu L, Tabien R, Stansel JW, Ying CS (1999) A "defeated" rice resistance gene acts as a QTL against a virulent strain of *Xanthomonas oryzae* pv. *oryzae*. Molecular & General Genetics 261: 58-63

McDowell JM, Dhandaydham M, Long TA, Aarts MGM, Goff S, Holub EB, Dangl JL (1998) Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the *RPP8* locus of Arabidopsis. The Plant Cell 10: 1861-1874

Meksem K, Leister D, Peleman J, Zabeau M, Salamini F, Gebhardt C (1995) A high-resolution map of the vicinity of the *R1* locus on chromosome V of potato based on RFLP and AFLP markers. Molecular & General Genetics 249: 74-81

Mesbah LA, Kneppers TJA, Takken FLW, Laurent P, Hille J, Nijkamp HJJ (1999) Genetic and physical analysis of a YAC contig spanning the fungal disease resistance locus *Asc* of tomato (*Lycopersicon esculentum*). Molecular & General Genetics 261: 50-57

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

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

Milligan SB, Bodeau J, Yaghoobi J, Kaloshian I, Zabel P, Williamson V (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. The Plant Cell 10: 1307-1319

Mindrinos M, Katagiri F, Yu GL, Ausubel FM (1994) The *A. thaliana* disease resistance gene *Rps2* encodes a protein containing a nucleotide-binding site and leucine-rich repeats. Cell 78: 1089-1099

Niewohner J, Salamini F, Gebhardt C (1995) Development of PCR assays diagnostic for RFLP marker alleles closely linked to alleles *Gro1* and *H1*, conferring resistance to the root cyst nematode *Globodera rostochiensis* in potato. Molecular Breeding 1: 65-78

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, Schonbeck F, Weber A (1994) Crop production and crop protection; Estimated losses in major food and cash crops. Elsevier Science B.V., Amsterdam, The Netherlands

Palomino MM, Meyers BC, Michelmore RW, Gaut BS (2002) Patterns of positive selection in the complete NBS-LRR gene family of *Arabidopsis thaliana*. Genome Research 12: 1305-1315

Parniske M, Hammond-Kosack KE, Golstein C, Thomas CM, Jones DA, Harrison K, Wulff BBH, Jones JDG (1997) Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the *Cf-4/9* locus. Cell 91: 821-832

Parniske M, Jones JDG (1999) Recombination between diverged clusters of the tomato *Cf-9* plant disease resistance gene family. Proceedings of the National Academy of Science 96: 5850-5855

Paul E, Goto M, Tanksley SD (1994) Solgenes - a Solanaceae database. Euphytica 79: 181-186

Payne RW, Lande PW, Ainsley AE, K.E. B, Digby PGN (1987) Genstat 5 Reference Manual. Clarendon Press, Oxford, UK

Phillips MS, Forrest JMS, Wilson LA (1980) Screening for resistance to potato cyst nematode using closed containers. Annals of Applied Biology 96: 317-322

Pineda O, Bonierbale MW, Plaisted RL (1993) Identification of RFLP markers linked to the *H1* gene conferring resistance to the potato cyst nematode *Globodera rostochiensis*. Genome 36: 152-156

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14: 817-818

Pryor T, Ellis J (1993) The genetic complexity of fungal resistance in plants. Advances in plant pathology 10: 282-305

Qi X, Niks RE, Stam P, Lindhout P (1998) Identification of QTLs for partial resistance to leaf rust (*Puccinia hordei*) in barley. Theoretical and Applied Genetics 96: 1205-1215

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

Richter TE, Pryor TJ, Bennetzen JL, Hulbert SH (1995) New rust resistance specificities associated with recombination in the *Rp1* complex in maize. Genetics 141: 373-381

Ritter E, Debener T, Barone A, Salamini F, Gebhardt C (1991) RFLP mapping on potato chromosomes of 2 genes controlling extreme resistance to potato virus X (PVX). Molecular & General Genetics 227: 81-85

Robinson MP, Atkison HJ, Perry RN (1988) The association and partial characterisation of a fluorescent hypersensitive response of potato roots to the potato cyst nematodes *Globodera rostochiensis* and *G. pallida*. Revue de Nematologie 11: 99-107

Ross H (1986) Potato breeding -problems and perspectives. Journal of Plant Breeding Supplement 13

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. Molecular Plant-Microbe Interactions 12: 197-206

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

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 (1997a) Mapping of the cyst nematode resistance locus *Gpa2* in potato using a strategy based on comigrating AFLP markers. Theoretical and Applied Genetics 95: 874-880

Rouppe van der Voort JNAM, Van Eck HJ, Draaistra J, Van Zandvoort PM, Jacobsen E, Bakker J (1998a) An online catalogue of AFLP markers covering the potato genome. Molecular Breeding 4: 73-77

Rouppe van der Voort JNAM, Van Zandvoort P, Van Eck HJ, Folkertsma RT, Hutten RCB, Draaistra J, Gommers FJ, Jacobsen E, Helder J, Bakker J (1997b) Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. Molecular & General Genetics 255: 438-447

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

Sambrook J, Fritsch EF, Maniatis T (1989) Moleculal cloning: a loboratory manual. Cold Spring Harbor Laboratory Press, New York

Schneider KA, Brothers ME, Kelly JD (1997) Marker-assisted selection to improve drought resistance in common bean. Crop Science 37: 51-60

Schnick D, Rumpenhorst HJ, Burgermeister W (1990) Differentiation of closely related *Globodera pallida* (Stone) populations by peans of DNA Restriction-Fragment-Length- Polymorphisms (RFLPs). Journal of Phytopathology 130: 127-136

Segal G, Sarfatti M, Schaffer MA, Ori N, Zamir D, Fluhr R (1992) Correlation of genetic and physical structure in the region surrounding the *I2 Fusarium oxysporum* resistance locus in tomato. Molecular & General Genetics 231: 179-185

Shen KA, Meyers BC, Islam-Faridi MN, Chin DB, Stelly DM, Michelmore RW (1998) Resistance gene candidates identified by PCR with degenerate oligonucleotide primers map to clusters of resistance genes in lettuce. Molecular Plant-Microbe Interactions 11: 815-823

Snedecor GW, Cochran WG (1967) Statistical methods, 6th edition. The Iowa State University Press, Ames, Iowa, USA

Sokal RR, Rohlf RF (1995) Biometry. W.H. Freeman and company, New York, USA

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

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package - Joinmap. Plant Journal 3: 739-744

Steffenson BJ, Hayes PM, Kleinhofs A (1996) Genetics of seedling and adult plant resistance to net blotch (*Pyrenophora teres* f. *teres*) and spot blotch (*Cochliobolus sativus*) in barley. Theoretical and Applied Genetics 92: 552-558

Stewart CN, Via LE (1993) A rapid CTAB DNA isolation technique useful for rapid fingerprinting and other PCR applications. Biotechniques 14: 748-750

Sun Q, Collins NC, Ayliffe M, Smith SM, Drake J, Pryor T, Hulbert SH (2001) Recombination between paralogues at the *Rp1* rust resistance locus in maize. Genetics 158: 423-438

Swofford DL (1998) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts

Takken FLW, Joosten M (2000) Plant resistance genes: their structure, function and evolution. European Journal of Plant Pathology 106: 699-713

Tanksley SD, Ganal MW, Prince JP, De Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Roder MS, Wing RA, Wu W, Young ND (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132: 1141-1160

Thompson JD, Higgins DG, Gibson TJ (1994) Clustal-W - Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position- specific gap penalties and weight matrix choice. Nucleic Acids Research 22: 4673-4680

Tommiska TJ, Hamalainen JH, Watanabe KN, Valkonen JPT (1998) Mapping of the gene *Nx(phu)* that controls hypersensitive resistance to potato virus X in *Solanum phureja* lvP35. Theoretical and Applied Genetics 96: 840-843

Trudgill DL (1967) The effect of environment on sex determination in *Heterodera rostochiensis*. Nematologica 13: 263-272

Van der Beek JG, Verkerk R, Zabel P, Lindhout P (1992) Mapping strategy for resistance genes in tomato based on RFLPs between cultivars: *Cf9* (resistance to *Cladosporium fulvum*) on chromosome 1. Theoretical and Applied Genetics 84: 106-112

Van der Plank JE (1982) Disease resistance in plants 2nd ed. Academic Press, New York, USA

Van der Vossen E, Rouppe van der Voort J, Kanyuka K, Bendahmane A, Sandbrink H, Baulcombe D, Bakker J, Stiekema W, Klein-Lankhorst R (2000) Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: a virus and a nematode. The Plant Journal 23: 567-576

Van Eck HJ, Rouppe van der Voort J, Draaistra J, Van Zandvoort P, Van Enckevort E, Segers B, Peleman J, Jacobsen E, Helder J, Bakker J (1995a) The Inheritance and Chromosomal Localization of Aflp Markers in a Noninbred Potato Offspring. Molecular Breeding 1: 397-410

Van Eck HJ, Rouppe van der Voort J, Draaistra J, Van Zandvoort P, Van Enckevort E, Segers B, Peleman J, Jacobsen E, Helder J, Bakker J (1995b) The inheritance and chromosomal localization of AFLP markers in a non-inbred potato offspring. Molecular Breeding 1: 397-410

Van Ooijen JW (1994) Drawmap - a computer-program for drawing genetic-linkage maps. Journal of Heredity 85: 66-66

Van Ooijen JW, Maliepaard C (1996a) MapQTL version 3: software for the calculation of QTL positions on genetic maps. CPRO-DLO, Wageningen, The Netherlands

Van Ooijen JW, Maliepaard C (1996b) MapQTL version 3: software for the calculation of QTL positions on genetic maps. Plant Genome IV

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

Vos P, Simons G, Jesse T, Wijbrandi J, Heinen L, Hogers R, Frijters A, Groenendijk J, Diergaarde P, Reijans M, Fierens-Osterenk 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. Nature Biotechnology 16: 1365-1369

Waetzig GH, Sobczak M, Grundler FMW (1999) Localization of hydrogen peroxide during the defence response of *Arabidopsis thaliana* against the plant-parasitic nematode *Heterodera glycines*. Nematology 1: 681-686

Wang GL, Mackill DJ, Bonman JM, McCouch SR, Champoux MC, Nelson RJ (1994) RFLP mapping of genes conferring complete and partial resistance to blast in a durably resistant rice cultivar. Genetics 136: 1421-1434

Webb CA, Richter TE, Collins NC, Nicolas M, Trick HN, Pryor T, Hulbert SH (2002) Genetic and molecular characterization of the maize *Rp3* rust resistance locus. Genetics 162: 381-394

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

Whitham S, Dineshkumar SP, Choi D, Hehl R, Corr C, Baker B (1994) The product of the tobacco mosaic-virus resistance gene-*N* - similarity to Toll and the Interleukin-1 receptor. Cell 78: 1101-1115

Young ND (1996) QTL mapping and quantitative disease resistance in plants. Annual Review of Phytopathology 34: 479-501

Samenvatting

In deze samenvatting is getracht de inhoud van het proefschrift en de achtergronden zo weer te geven dat de hoofdlijnen van het onderzoek ook begrijpelijk zijn voor niet-ingewijden. Voor een vaktechnische weergave wordt verwezen naar de Engelse samenvatting.

Introductie
Aardappel (*Solanum tuberosum* ssp. *tuberosum* L.) is een belangrijk voedselgewas in gematigde klimaatzones. De productie van aardappel is ongeveer 300 miljoen ton per jaar en daarmee is aardappel het vierde voedselgewas in de wereld, na tarwe, maïs en rijst. In de zestiende eeuw is de aardappel door de Spaanse "conquistadores" vanuit Zuid-Amerika, waar hij waarschijnlijk al zo'n 8000 jaar geleden was gedomesticeerd naar Europa gebracht. Onze cultuuraardappel heeft twaalf chromosomen en is tetraploïd. Een aardappel heeft dus in totaal achtenveertig chromosomen waarvan steeds vier chromosomen homoloog zijn. Ter vergelijking, een mens heeft drieëntwintig chromosomen en is diploïd. Genetica met tetraploïde organismen is erg ingewikkeld, maar gelukkig is er in de jaren zeventig van de twintigste eeuw een methode ontwikkeld waarbij met behulp van schijnbevruchting een diploïde aardappelplant verkregen kan worden. Sindsdien wordt veel genetisch onderzoek met diploïde aardappels gedaan.

Zoals elk gewas, wordt ook aardappel belaagd door ziekten en plagen. Aardappelmoeheid is een groot probleem in de aardappelteelt en wordt veroorzaakt door de twee soorten aardappelcystenaaltjes *Globodera rostochiensis* en *G. pallida,* die parasiteren op de wortels van aardappel. De schade die ze veroorzaken bestaat uit het verstoren van de knolontwikkeling, het verwelken en uiteindelijk het afsterven van de plant. Plekken op een akker waar een besmetting is zijn vaak te herkennen aan de verwelkte en/of compleet afgestorven planten en worden ook wel valplekken genoemd. De aardappelcystenaaltjes zijn obligaat parasitair, dat wil zeggen dat ze zich niet kunnen voortplanten zonder de aanwezigheid van een geschikte gastheer. Ondanks de betrekkelijk nauwe waardplantenreeks van deze nematoden (ze kunnen zich alleen voortplanten op enkele genera van de nachtschadefamilie) is de toepassing van vruchtwisseling economisch gezien niet een voldoende bevredigende tactiek om een aardappelmoeheidsplaag onder controle te houden. Dit komt omdat, nadat de eieren in het lichaam van het vrouwtje gevormd zijn, het vrouwtje verhardt tot een beschermend omhulsel. De nieuwe generatie is op deze manier in staat jarenlang zonder waardplant in de bodem te overleven. Wanneer opnieuw een geschikte waardplant in de nabijheid van een cyste aanwezig is, zorgen stoffen uit de wortels ervoor dat de juveniele nematoden uit het ei kruipen en de nieuwe waardplant parasiteren. Na de tweede wereldoorlog kwam daarom de chemische grondontsmetting in zwang. Door het gebruik van toxische en aspecifieke middelen

werd de aardappeloogst beschermd tegen aardappelmoeheid en lange tijd werd deze methode als dé oplossing voor het aardappelmoeheidsprobleem gezien. Pas in 1989 is in het zogenaamde Meerjarenplan Gewasbescherming besloten dat het gebruik van deze chemicaliën sterk aan banden moest worden gelegd. Daarom is men naarstig op zoek naar nieuwe economisch rendabele methoden die een aardappelmoeheidsplaag tegen kunnen gaan.

Merkergestuurde veredeling
Resistentie tegen aardappelmoeheid wordt gezien als een veelbelovend alternatief voor chemische bestrijding en een goede aanvulling op vruchtwisseling. Resistenties worden vaak ontdekt in zogenaamde wilde aardappelsoorten dat wil zeggen, leden van het geslacht *Solanum* die van nature in Zuid-Amerika voorkomen. In verschillende wilde aardappelsoorten (zie Tabel 1 in hoofdstuk 1) zijn diverse resistentie-eigenschappen gevonden, zowel tegen *G. pallida* (b.v. *Gpa, Gpa2, Gpa3, Gpa5* en *Gpa6*) als tegen *G. rostochiensis* (b.v. *Gro1, Gro1.2, Gro1.3, Gro1.4, GroV1* en *H1*). Sommige van deze resistentie-eigenschappen geven kwantitatieve resistentie en zorgen voor een gereduceerde vermeerdering van de nematoden (bv *Gpa5* en *Gpa6* in hoofdstuk 2), terwijl andere voor kwalitatieve resistentie zorgen en er op de resistente plant geen vermeerdering van de nematoden plaatsvindt (b.v. *H1* in hoofdstuk 3).

Nu is het probleem met wilde aardappelsoorten dat ze vaak niet eetbaar zijn of niet aan de hoge kwaliteitseisen voldoen die wij in Nederland aan onze aardappelen stellen. Daarom wordt een wilde aardappel met een gewenste resistentie-eigenschap gekruist met een aardappel die aan alle andere kwaliteitseisen voldoet. De nakomelingen van deze kruising worden vervolgens geselecteerd op zowel de resistentie-eigenschap afkomstig van de wilde ouder als de kwaliteitseigenschappen van de andere ouder. Dit zogenaamde terugkruisen wordt meerdere malen herhaald. Uiteindelijk hoopt men een nieuwe aardappelvariëteit te hebben verkregen die aan alle kwaliteitseisen voldoet en bovendien de gewenste resistentie-eigenschap bezit.

De klassieke manier om de individuen van een nakomelingschap te testen op het bezit van de gewenste aardappelmoeheidsresistentie-eigenschap is een zogenaamde resistentietoets, die wordt uitgevoerd met aardappelknollen. Het duurt ongeveer 10 à 12 weken voordat gezien kan worden of een plant resistent is. Wanneer de cysten volgroeid zijn worden de wortels opgespoeld met behulp van een “Kort trechter”. Na het opspoelen kunnen de cysten geteld worden en aan de hand van de hoeveelheid cysten per plant wordt bepaald of een plant wel of niet resistent is. Deze toets kan pas een jaar nadat de zaadjes zijn ontkiemd worden uitgevoerd: eerst moet de plant volwassen worden en aardappels produceren en daarna ondergaan de aardappels nog een periode van kiemrust voordat ze in staat zijn weer uit te lopen.

Om de selectie van resistente aardappelplanten te versnellen en minder arbeidsintensief te maken wordt gezocht naar een methode om snel en gemakkelijk te kunnen testen of een plant een bepaalde resistentie-eigenschap bevat. Dit testen wordt gedaan met "moleculaire merkers". Merkers zijn plaatsen op het aardappel-DNA die eigenschappen representeren. Merkers kunnen "morfologische merkers" zijn en "echte" eigenschappen zoals resistentie, schil-, of bloemkleur vertegenwoordigen, maar "moleculaire merkers" representeren "nietszeggende" stukjes DNA. De voordelen van moleculaire merkers zijn dat 1) de hoeveelheid merkers in een organisme nagenoeg eindeloos is en met behulp van één kruising voldoende merkers kunnen worden verkregen, 2) de merkers niet beïnvloed worden door de externe omstandigheden waarin de plant zich bevindt.

Om te beginnen wordt met de moleculaire merkers een "framework" opgezet, een zogenaamde genetische kaart. Een genetische kaart is een representatie van de relatieve volgorde van de moleculaire merkers. Er wordt geen onderscheid gemaakt tussen de homologe chromosomen van de aardappelplant, dus een genetische kaart van aardappel bestaat uit twaalf groepen merkers die binnen elke groep op volgorde zijn gelegd. Daarna kan met behulp van de genetische kaart worden bepaald waar een eigenschap (bijv. resistentie) zich op het aardappel-DNA bevindt. Dit wordt gedaan door naar het verband te kijken tussen het voorkomen van merkers en het voorkomen van de resistentie-eigenschap in nakomelingen van een kruising. Als bijvoorbeeld in een nakomelingschap van 100 planten een bepaalde merker 25 keer samen met de resistentie overerft en 75 keer niet, dan is er tussen die merker en de resistentie geen verband. Als nu een andere merker juist 99 keer samen met de resistentie overerft en 1 keer niet, dan kan dat geen toeval meer zijn en is er wel verband. Hoe vaker een merker en een eigenschap samen overerven, hoe groter het verband en hoe dichter ze bij elkaar op het aardappel-DNA liggen. Omdat de volgorde en plaats van de moleculaire merkers bekend is (genetische kaart) wordt zodoende ook de plek van de resistentie vastgesteld. In hoofdstuk 2 is op deze manier de locatie van twee resistentie-eigenschappen, *Gpa5* en *Gpa6,* bepaald. Door nu het aardappel-DNA te testen op de aanwezigheid van de dichtstbijzijnde merker kan al in een vroeg stadium (zaailing) bepaald worden of een bepaalde resistentie wel of niet aanwezig is in een nakomeling.

R genen

Als de locatie van een resistentie-eigenschap eenmaal bepaald is, is het interessant om te proberen het gen dat verantwoordelijk is voor die resistentie te isoleren. De plaatsbepaling zoals beschreven in de vorige paragraaf is vaak te grof om het gen makkelijk te kunnen isoleren. Daarom moet er worden ingezoomd op de regio waar zich de eigenschap ongeveer bevindt. Dat inzoomen wordt gedaan door de hoeveelheid nakomelingen en de hoeveelheid merkers te vergroten, zodat er een beter verband gelegd kan worden tussen de resistentie en de merkers. Op deze manier is in hoofdstuk 3 een genetische kaart gemaakt met een hoge merker-

dichtheid van het stukje van chromosoom V waar zich de *H1* resistentie bevindt. Als het gen geïsoleerd is, kan de structuur van dat gen worden bepaald.

Het is gebleken dat de meeste *R* genen niet alleen op het aardappel-DNA zitten, maar dat ze in groepjes voorkomen. Bovendien lijken de genen die in zo'n groepje liggen vaak erg op elkaar, terwijl ze niet dezelfde resistentie-eigenschappen bezitten. Het *Gpa2* gen, bijvoorbeeld, bevindt zich op chromosoom XII in een groepje waar zich ook een *R* gen tegen een virus bevindt (*Rx1*). In hoofdstuk 4 zijn uit twee aardappelplanten alle genen van het *Gpa2/Rx1* cluster geïsoleerd. In die twee planten zijn in totaal veertien genen gevonden die allemaal in de buurt van het *Gpa2* gen liggen en die inderdaad allemaal erg op elkaar lijken. Bovendien is gebleken dat tenminste zes van deze genen tot expressie komen. Mogelijk liggen deze genen ten grondslag aan een onbekende resistentie-eigenschap.

Sequentie-analyse van *R* genclusters zal ons begrip vergroten met betrekking tot hun evolutie. Het vergelijken van de structuur van de *Gpa2/Rx1* homologen (hoofdstuk 5) liet zien dat nieuwe *R* genen zouden kunnen ontstaan door middel van positieve selectie in combinatie met overkruising en genconversie. Het vergelijken van leden van hetzelfde *R* gencluster, vooral in het geval van *Gpa2* en *Rx1,* waar twee genen die veel op elkaar lijken resistentie geven tegen twee compleet verschillende organismen, kan ons ook helpen meer inzicht te verkrijgen in de rol die *R* genen hebben bij de activering van een resistentiereactie. Daarnaast kan het ons helpen begrijpen op welke manier de verschillende domeinen van *R* genen hierbij betrokken zijn. Het is echter noodzakelijk om hypothesen die voortkomen uit de *in silico* analyses van *R* genen te toetsen met *in vivo* structuur-functie analyses.

List of publications

Bakker, E., Achenbach, U., Bakker, J., Van Vliet, J., Peleman, J., Segers., Van der Heijden, S., Van der Linde, P., Van Eck, H., 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*. *submitted*

Bakker, E., Rouppe van der Voort, J., Van Eck, H., Hutten, R., Bakker, J. Goverse, A. (2004) Molecular markers for resistance to cyst nematodes in potato. In: Proceedings of the Fourth International Congress of Nematology. *in press*

Bakker, E., Van Vliet, J., Overmars, H., Smant, G., Sandbrink, J., Van der Vossen, E., Bakker, J., Goverse, A. (2004) *R* gene homologues in potato confer resistance to distinct pathogens: a virus and a nematode. In: Proceedings of the Fourth International Congress of Nematology. *in press*

Bakker, E., Dees, R., Bakker, J., Goverse, A. (2003) Mechanisms involved in plant resistance to nematodes. In: Multigenic and induced systemic resistance in plants. Edited by Sadik Tuzun and Elizabeth Bent. Kluwer Acedemic Press, New York *in press*

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. Theoretical and Applied Genetics 106: 1524-1531

Jones, J.T.; Cleber, F.; Bakker, E.; Banks, B.; Blok, V.; Chen, Q.; Phillips, M., Proir, A. (2003) Characterisation of a chorismate mutase from the potato cyst nematode *Globodera pallida* Molecular Plant Pathology 4 (1): 43-50

Bakker, E.; Vliet, J. van; Overmars, H.; Smant, G.; Sandbrink, H.; Vossen, E. van der; Bakker, J.; Goverse, A. (2002) *R* gene homologues in potato confer resistance to distinct pathogens: a virus and a nematode. Nematology 4 (2): 146

Bakker, E.; Rouppe Voort, J. van der; Eck, H. van; Bakker, J.; Goverse, A. (2002) Molecular markers for resistance to cyst nematodes in potato. Nematology 4 (2): 138-139

Goverse, A.; Bakker, E.; Smant, G.; Sandbrink, H.; Vossen, E. van der; Bakker, J.(2001) *R* gene homologues in potato confer resistance to distinct pathogens: Nematodes and viruses. Phytopathology 91 (6): 147

Bakker, E.; Vermaas, L.; Vossen, E. van der; Eck, H.J. van; Klein-Lankhorst, R.; Bakker, J.; Rouppe van der Voort, J. (2000) A novel, pcr based method to identify *Gpa2* homologues among Solanaceae. Durable disease resistance, key to sustainable agriculture: 72

Bakker, E.; Vermaas, L.; Vossen, E. van der; Eck, H. van; Klein-Lankhorst, R.; Bakker, J.; Rouppe van der Voort, J.(2000) A novel pcr based method to identify *Gpa2* homologues among *Solanaceae* Abstract book: Interaction between plants and attacking organisms, mechanisms, genetics, ecology and evolution: 20

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

Popeijus H, Blok VC, Cardle L, Bakker E, Phillips MS, Helder J, Smant G, Jones JT (2000) Analysis of genes expressed in second stage juveniles of the potato cyst nematodes *Globodera rostochiensis* and *G. pallida* using the expressed sequence tag approach Nematology 2: 567-574

Bakker, E.H.; Vermaas, L.C.; Vossen, E.A.G. van der; Eck, H.J. van; Bakker, J.; Rouppe van der Voort, J.N.A.M. (1999) Mapping of resistance gene homologues on a high density map of potato. Book of abstracts : 9th international congress on Molecular Plant-Microbe interactions, Amsterdam, 25-30 juli 1999: 152

Bakker, E.H.; Blok, V.C.; Stokkermans, J.P.W.G.; Burrows, P.R.; Phillips, M.S.; Jones, J.T.(1998) Characterisation of abundantly expressed genes from the two species of potato cyst nematode (PCN) Globodera rostochiensis and G. pallida. British Society for Parasitology Spring Meeting, Exeterm, UK: 51

Bakker, E.H.; Blok, V.C.; Stokkermans, J.P.W.G.; Burrows, P.R.; Phillips, M.S.; Jones, J.T. (1998) EST analysis of genes expressed in the two species of potato cyst nematode (PCN) Globodera rostochiensis and G. pallida. 24th International Nematology Symposium, Dundee, UK: 6

Bakker, E.H.; Blok, V.C.; Stokkermans, J.P.W.G.; Burrows, P.R.; Phillips, M.S.; Jones, J.T. (1998) EST analysis of genes expressed in the two species of potato cyst nematode (PCN) Globodera rostochiensis and G. pallida. Nematologica 44: 456-457

Nawoord

Met veel plezier heb ik de afgelopen jaren gewerkt aan mijn promotie-onderzoek bij het Laboratorium voor Nematologie. Nadat ik mijn afstudeervak bijna had afgerond ging ik, met de belofte het verslag na terugkomst af te ronden, op vakantie naar Amerika. Ik was net terug van Schiphol en zat nog wat na te suffen op de bank toen de telefoon ging. Het was Jeroen, de begeleider van mijn afstudeervak, met de vraag of ik even op de vakgroep langs wilde komen. “Moet ik vandaag al beginnen met het afronden van mijn verslag?” dacht ik en ik vertelde aan de telefoon dat ik een beetje moe was vanwege de jet lag. Dat ik moe was gaf helemaal niet en of ik toch maar even wilde komen. Enfin, zo gezegd zo gedaan en ik zat diezelfde middag nog met een hoofd vol watten op de vakgroep. Ze hadden goed nieuws. Er was een aio-plaats beschikbaar gekomen en Jaap Bakker en Jeroen Rouppe van der Voort vroegen zich af of ik het leuk zou vinden aio te worden op hun vakgroep. “Het zal wel voornamelijk moleculaire biologie worden” waarschuwden ze. Maar ik was na mijn stage in Schotland en het afstudeervak bij Jeroen helemaal om, dus moleculaire biologie was geen punt. Ik hoefde er dan ook niet lang over na te denken en in oktober 1998 kon ik beginnen. Uiteraard heb ook ik te maken gehad met de beroemde aio-dip. Jeroen had inmiddels een andere baan en mijn onderzoek zat zo vast als een huis. Gelukkig ben ik daar -met name met behulp van Aska Goverse en de geweldige assistentie van Joke van Vliet- weer helemaal bovenop gekomen en heb ik mijn promotie-onderzoek tot een goed einde kunnen brengen. Bedankt! Uiteraard heeft ook Jaap Bakker met zijn stroom goede ideeën een grote invloed op mijn werk gehad en daarvoor wil ik hem dan ook hartelijk bedanken. Herman van Eck en Edwin van der Vossen wil ik bedanken voor alle goede adviezen en discussies. En natuurlijk wil ik ook alle andere collega’s hartelijk bedanken voor de gezellige sfeer die er altijd op de vakgroep heerst. Zo gezellig, dat ik zelfs noodgedwongen mijn werkplek vanuit “de koffiekamer” in de “nieuwbouw” heb moeten verplaatsen naar de zolder van de “oudbouw”, waar het een stuk rustiger was. Verder wil ik nog kwijt dat niet alle eer naar de vakgroep gaat. Mijn ouders, die het studeren altijd al hebben gestimuleerd en me zelfs de kans hebben gegeven een jaar langer te studeren zodat ik me op de valreep nog kon verdiepen in de moleculaire biologie, en uiteraard ook Joost en Mirte wil ik hartelijk bedanken voor hun interesse in mijn werk en het wel en wee eromheen. Mijn AD&D vrienden wil ik bedanken voor de vele uren waarin ik me compleet kon verliezen in een andere wereld. En tenslotte Theo. Jij was en bent er altijd om al mijn verhalen (van gefrustreerd tot euforisch) aan te horen en zo nodig van commentaar te voorzien. Bedankt voor alle begrip, discussies, tips en ideeën!

Curriculum vitae

Erin Harriët Bakker werd op 13 februari 1973 geboren te Amsterdam. Ze behaalde in 1992 haar VWO-diploma aan het Bonhoeffer College te Castricum. Hierna is zij Plantenveredeling en Gewasbescherming gaan studeren aan de Wageningen Universiteit (toen nog Landbouwuniversiteit) met als specialisatie Ecologische Gewasbescherming. Haar eerste afstudeervak werd uitgevoerd bij het Laboratorium voor Fytopathologie onder begeleiding van dr. ir. W. Blok. Binnen dit afstudeervak is onderzoek gedaan naar het effect van semi-anaërobe condities op de overleving van sporen van de bodempathogenen *Verticillium dahliae* en *Fusarium oxysporum* f. sp. *asparagi*. Haar daaropvolgende afstudeervak werd uitgevoerd op het DLO Instituut voor Plantenziektekundig Onderzoek (IPO) onder begeleiding van dr. ir. F. Zoon (IPO) en ing. J. van Bezooijen (WU). Hier heeft zij onderzoek gedaan naar de status van verschillende gewassen als gastheer voor de nematoden *Trichodorus primitivus* en *Paratrichodorus teres.* Ook is gekeken in hoeverre deze gewassen als bron kunnen dienen voor het tabaksratelvirus (TRV), dat deze nematoden als vector gebruikt. Haar stage was bedoeld om kennis te maken met de moleculaire biologie. Op het Scottish Crop Research Institute (SCRI) in Invergowrie, Schotland heeft ze onder begeleiding van dr. J. Jones (SCRI) en dr. J. Helder (WU) onderzoek gedaan naar een set genen die tot expressie komen in preparasitaire juvenielen van het aardappelcystenaaltje *Globodera pallida* met behulp van de techniek "Expressed Sequence Tag (EST) analyses". Enthousiast geraakt voor de moleculaire biologie heeft ze een laatste afstudeervak uitgevoerd bij het Laboratorium voor Nematologie onder begeleiding van dr. J. Rouppe van der Voort. Binnen dit afstudeervak is met behulp van AFLP merkers een genetische kaart van aardappel geproduceerd die vervolgens is vergeleken met bestaande genetische kaarten van andere aardappel-populaties. In 1998 heeft ze deze studie afgerond. Aansluitend is zij aangesteld als assistent in opleiding (aio) bij het Laboratorium voor Nematologie. In eerste instantie binnen het EU-project FAIR5-PL97-3565, "Construction and application of a multifunctional ultra-dense genetic map of potato", en later binnen het STW-project 805.18.760, "Ontwikkeling van een economisch rendabele strategie voor de isolatie en karakterisering van een set duurzame resistentiegenen tegen aardappel-cystenaaltjes". Sinds december 2002 is zij bij het Laboratorium voor Nematologie werkzaam als post-doc onderzoeker.

The research presented in this thesis has been carried out at the Laboratory of Nematology, Wageningen University.

Financial support for this thesis was kindly provided by the European Union grant FAIR5-PL97-3565 and the Dutch Technology Foundation (STW) grant 805.18.760.

Bibliographic abstract
This thesis describes the genomic localisation and organisation of loci that harbour resistance to the potato cyst nematode species *Globodera pallida* and *G. rostochiensis*. Resistance to the potato cyst nematodes *G. pallida* and *G. rostochiensis* is an important aspect in potato breeding. To gain insight in the structure of the genes underlying nematode resistance, it is necessary to first map nematode resistance loci, as described for the resistance loci *Gpa5* and *Gpa6*. Subsequently, both for map-based cloning and for marker-assisted selection, a high-resolution map of the resistance locus is required. A high-resolution map is produced for the *H1* locus. To gain more insight in the genomic organisation of a complex disease resistance locus, *Gpa2/Rx1* homologues derived from different haplotypes have been characterised. Finally, the *Gpa2/Rx1* homologues derived from five haplotypes have been subjected to extensive sequence analysis to elucidate their structural diversity and evolutionary relationships.

Cover:
Vincent van Gogh (1853-1890)
De aardappeleters, 1885
Olieverf op doek, 82 x 114 cm
Van Gogh Museum, Amsterdam (Vincent van Gogh Stichting)

Druk:
Ponsen & Looijen BV, Wageningen