Theor Appl Genet (2010) 120:679–689
DOI 10.1007/s00122-009-1185-0

ORIGINAL PAPER

Development of diagnostic markers for use in breeding potatoes resistant to *Globodera pallida* pathotype Pa2/3 using germplasm derived from *Solanum tuberosum* ssp. *andigena* CPC 2802

Claire Moloney · Denis Griffin · Peter W. Jones · Glenn J. Bryan · Karen McLean · John E. Bradshaw · Dan Milbourne

Received: 20 February 2009 / Accepted: 9 October 2009 / Published online: 1 November 2009

Abstract Quantitative resistance to *Globodera pallida* pathotype Pa2/3, originally derived from *Solanum tuberosum* ssp. *andigena* Commonwealth Potato Collection (CPC) accession 2802, is present in several potato cultivars and advanced breeding lines. One genetic component of this resistance, a large effect quantitative trait locus (QTL) on linkage group IV (which we have renamed $GpaIV^{s}_{adg}$) has previously been mapped in the tetraploid breeding line 12601ab1. In this study, we show that $GpaIV^{s}_{adg}$ is also present in a breeding line called C1992/31 via genetic mapping in an F_1 population produced by crossing C1992/31 with the *G. pallida* susceptible cultivar Record. C1992/31 is relatively divergent from 12601ab1, confirming that $GpaIV^{s}_{adg}$ is an ideal target for marker-assisted selection in currently available germplasm. To generate markers exhibiting diagnostic potential for $GpaIV^{s}_{adg}$, three bacterial artificial chromosome clones were isolated from the QTL region, sequenced, and used to develop 15 primer sets generating single-copy amplicons, which were examined for polymorphisms exhibiting linkage to $GpaIV^{s}_{adg}$ in C1992/31. Eight such polymorphisms were found. Subsequently, one insertion/deletion polymorphism, three single nucleotide polymorphisms and a specific allele of the microsatellite marker STM3016 were shown to exhibit diagnostic potential for the QTL in a panel of 37 potato genotypes, 12 with and 25 without accession CPC2082 in their pedigrees. STM3016 and one of the SNP polymorphisms, C237(119), were assayed in 178 potato genotypes, arising from crosses between C1992/31 and 16 *G. pallida* susceptible genotypes, undergoing selection in a commercial breeding programme. The results suggest that the diagnostic markers would most effectively be employed in MAS-based approaches to pyramid different resistance loci to develop cultivars exhibiting strong, durable resistance to *G. pallida* pathotype Pa2/3.

Introduction

The potato cyst nematode *Globodera pallida* is one of the most significant soilborne pests of potatoes worldwide. In Northern Europe, field populations of *G. pallida* comprising mixtures of pathotypes Pa2 and Pa3 are prevalent, mixed populations sometimes being referred to as pathotype Pa2/3. Developing cultivars expressing high levels of resistance to these heterogeneous populations is a major goal of potato breeders. The availability of DNA-based markers, which are easy to score, cost-effective and diagnostic for resistance to *G. pallida* Pa2/3 would greatly speed up the process of new variety development by allowing a marker-assisted selection (MAS)-based approach to develop lines in which multiple resistance sources were combined, or pyramided, preferably with each individual resistance locus present in high dosage states in the tetraploid potato genome.

Communicated by C. Gebhardt.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-009-1185-0) contains supplementary material, which is available to authorized users.

C. Moloney · D. Griffin · D. Milbourne (✉)
Crops Research Centre, Teagasc, Oak Park, Carlow, Ireland
e-mail: dan.milbourne@teagasc.ie

P. W. Jones
Department of Zoology, Ecology and Plant Science,
University College Cork, Butler Building,
Distillery Field, North Mall, Cork, Ireland

G. J. Bryan · K. McLean · J. E. Bradshaw
Genetics Programme, Scottish Crop Research Institute,
Invergowrie, Dundee DD2 5DA, UK

A prerequisite for this MAS-based approach to resistance breeding is the genetic mapping of loci conferring resistance to *G. pallida* pathotypes Pa2 and Pa3 in potato. Several such loci have been mapped, including *Gpa* (Kreike et al. 1994), *Gpa2* (Rouppe van der Voort et al. 1997), *Gpa3* and *Gpa4* (Wolters et al. 1999), *Gpa5* and *Gpa6* (Rouppe van der Voort et al. 2000), *Grp1* (Rouppe van der Voort et al. 1998), $GpaV^{s}_{spl}$ and $GpaXI^{s}_{spl}$ (Caromel et al. 2005), $GpaXI^{l}_{tar}$ (Tan 2008) and several unnamed loci (Bryan et al. 2002; Caromel et al. 2003). These loci mainly originate from diploid wild relatives of cultivated potato, including *Solanum vernei*, *S. sparsipilum*, *S. spegazinii* and *S. tarijense*, or, in the case of *Gpa2*, *S. tuberosum* ssp. *andigena* CPC1673. The loci differ in their effects, some conferring complete or partial resistance to populations of one or other of the pathotypes, and others conferring varying levels of resistance to populations of both pathotypes. However, no single locus confers complete resistance to all populations of both pathotypes of *G. pallida*.

Quantitative resistance to *G. pallida* pathotype Pa2/3, originally derived from *S. tuberosum* ssp. *andigena* Commonwealth Potato Collection (CPC) accession 2802, is present in some cultivars and advanced breeding lines. This resistance source is particularly useful from a breeding perspective due to the fact that *S. tuberosum* ssp. *andigena* is a cultivated tetraploid form of potato closely related to modern *S. tuberosum* ssp. *tuberosum*, obviating the extensive backcrossing to the Tuberosum genepool required to eliminate undesirable donor traits when the resistance is introgressed from wild species. A major genetic component of this resistance has previously been genetically mapped in a segregating population resulting from a cross between the *G. pallida*-resistant advanced breeding line 12601ab1 and the cultivar Stirling (Bradshaw et al. 1998; Bryan et al. 2004). A large effect QTL in the duplex configuration was mapped to the distal arm of chromosome IV of 12601ab1. The QTL maps to a position similar to that of the reasonably well characterised R gene locus to which the blight resistance genes *R2*, *R2-like*, *Rpi-blb3* and *Rpi-abpt* have previously been mapped (Park et al. 2005a, b), although insufficient common markers exist between the map of 12601ab1 and the other studies to make accurate positional comparisons. A second, smaller QTL was also detected on LG XI of 12601ab1. The major LGIV QTL was not named in the studies of Bradshaw et al. 1998 or Bryan et al. 2004, and was referred to as Gpa4 in the review by Gebhardt and Valkonen (2001). Given the increasing number of PCN resistance loci being isolated, we propose that, for clarity, this QTL be referred to in future as $GpaIV^{s}_{adg}$, according to the nomenclature described first by Caromel et al (2005) using the scheme [pathogen species, potato linkage group (roman), long/short arm (superscript), source species (subscript)].

Unfortunately, markers found to be linked to QTL in mapping experiments frequently do not satisfy the requirements for use in marker-assisted breeding. Markers such as AFLPs and RFLPs generally require conversion to some form of sequence-characterised marker. Even when sequence-characterised markers are available, they may not be fully diagnostic for the trait in question because the particular allele of the marker found linked to the QTL in the original mapping population is not consistently found associated with the QTL in a broader set of germplasm, due to recombination, or because that particular marker allele is present at a high frequency in the genepool of material used for breeding. For example, Niewohner et al. (1995) developed a PCR assay for detecting RFLP alleles at a marker locus, CP56, which was closely linked in coupling to the nematode resistance gene *Gro*1 on potato chromosome VII in a segregating population. However, when tested in 136 diverse tetraploid potato varieties, the marker allele originally found linked to *Gro*1 was not correlated with PCN resistance. Conversely, from an AFLP marker linked to a *S. vernei*-derived QTL conferring resistance to *G. pallida* pathotype Pa2/3 resistance located in the major disease resistance locus on potato chromosome V, Bryan et al. (2002) developed a dominant single-locus PCR marker that was consistently associated with PCN resistance in genotypes with *S. vernei* in their pedigree. In a slightly different approach to develop a diagnostic marker to the same quantitative resistance locus, Sattarzadeh et al (2006) sequenced PCR amplicons derived from end sequences of BACs spanning the region containing the *R1* blight resistance gene region to identify SNPs linked to resistance in a segregating population, identified a subset of these SNPs diagnostic for resistance in a broad germplasm set of genotypes with and without *S. vernei* in their background and subsequently used this information to develop a highly diagnostic PCR-based marker for *S. vernei*-derived resistance to *G. pallida* pathotype Pa2/3.

Our objectives in this study were to confirm that $GpaIV^{s}_{adg}$ is an appropriate target for MAS by showing that it is present in another breeding line called C1992/31, which is divergent from 12601ab1 (Fig. 1), to identify DNA polymorphisms linked to $GpaIV^{s}_{adg}$ that exhibit diagnostic potential in a broad set of breeding germplasm, and to validate the utility of these markers for MAS in a commercial breeding programme.

Materials and methods

Plant material and DNA isolation

Three sets of germplasm were used in the experiments described in this study.

As part of the potato breeding programme at Oak Park Research Centre, an F_1 population comprising 94 progeny individuals was produced by crossing the advanced tetraploid breeding line C1992/31 (male parent) with the *G. pallida* Pa2/3 susceptible potato cultivar Record (female parent). This is referred to as the T3683 population in this study.

Secondly, a panel of 37 clones comprising breeding material from both Oak Park and the Scottish Crop Research Institute (SCRI) was established. This 'validation panel' was composed of 12 breeding lines and cultivars that are known to have *S. tuberosum* ssp *andigena* accession CPC2802 in their pedigrees, and 25 which do not. The genotypes are: with CPC2802—C1992/31, C1992/42, H1H3-6, H1H3-140, 10920ad9, 12601ab1, 12636a2, 12674ab1, Vales Everest, Eden, 97MT200B61, Scarborough.

Without CPC2802—10300-13, 15119ac5, 657/d3*, 8204a4, Cara*, Desiree*, Druid*, Famosa, Habibi*, Lady Balfour, Maris Peer*, Morene*, Navan*, Orla, P55/7, Panda*, Pentland Ivory*, Record, Rooster, Sandra, Saturna*, T2270/50*, T4069/22, Z89/6 (*also in breeding panel, see below).

A third set of germplasm was assembled that comprised almost all of the available progeny with C1992/31 as one parent, which were undergoing selection in the fourth year of the Oak Park Potato Breeding Programme. A total of 178 clones originating from crosses between C1992/31 and 16 other genotypes were used for this study. Twelve parental genotypes are represented in the validation panel, the others were Ausonia, Buchan, Estima and Ulster Glade. None are known to possess significant resistance to *G. pallida*. This germplasm set is referred to as the 'breeding panel' in this study.

DNA was isolated from all of the above using either a Qiagen DNAeasy kit or a standard CTAB-based DNA extraction.

Assessment of resistance to *G. pallida* Pa2/3

PCN testing was carried out on the 94 individuals of the T3683 population (including C1992/31) and the 178 individuals in the breeding panel. Tubers from each genotype were used for resistance testing to *G. pallida* pathotype Pa2/3 (Lindley) in a closed container test with four replicates (one tuber per replicate) in a completely randomised design experiment as described by Phillips et al. (1980). Data were collected in the form of the numbers of cysts visible on roots through the transparent walls of the containers 8 weeks after inoculation. The mean cyst count over the four replicates was established for each genotype. Data from the T3683 population were subjected to the variance stabilizing square root transformation for all subsequent calculations. Data from the breeding panel were not transformed.

Broad-sense heritability of PCN resistance in the T3683 population was calculated in Minitab (Minitab Inc 2006) using the formula $h^2 = V_g/(V_g + V_e/r)$, Components of variance were calculated from the mean squares in ANOVA using Minitab.

Development of sequence-characterised markers in the $GpaIV^s_{adg}$ region

To develop sequence-characterised markers specific for the region surrounding the $GpaIV^s_{adg}$ locus, a BAC library of the diploid potato RH, the male parent of the ultra high density mapping population of potato (van Os et al. 2006) was screened with the sequenced AFLP marker DPGAM-CAG_155.0 and the microsatellite marker STM3016, which map to the $GpaIV^s_{adg}$ region in 12601ab1. This resulted in the isolation of three BACs (GB003E01, GB005C13—isolated using DPGAMCAG_155.0, and GB080G22—isolated using STM3016). One of these (GB005C13) was submitted to a preliminary round of low-coverage shotgun sequencing, which resulted in the availability of a number of sequence contigs. Thirteen of these contigs exhibited homology to Solanaceous ESTs when subjected to BLASTn analysis (results not shown). PCR primers were designed to these sequence contigs, amplified in Record, C1992/31 and ten progeny from the mapping population. To identify polymorphisms, the resulting amplicons were subjected to restriction digest with a range of 4-cutter restriction endonucleases. One primer set, called Contig237, developed by this process, was subsequently used for mapping in the T3683 population.

To develop SNP markers, all three BACs were subsequently sequenced to sixfold coverage (GATC, Germany) using a standard Sanger shotgun approach, and assembled (using the SeqMan module of Lasergene) to phase 1 (unordered, unoriented contigs). Preliminary annotation of the BAC contigs using ab initio gene prediction and homology approaches revealed that none of the BACs contain R-genes (data not shown). BACs GB003E01 and GB080G22 were submitted to GenBank as AC236732 and AC236733, respectively. GB005C13 could not be submitted due to sequence quality issues, but largely overlaps GB003E01. Following the assembly, 11 contigs exceeding 6 kb in size from the three BACs were searched using BLASTn against the entire plant EST database at NCBI. Fourteen primer pairs capable of generating single-copy amplicons in potato, when resolved on a 2% agarose gel, were designed to EST hits within the BAC contig sequences. The 14 primer sets, and the Contig237 primer set, were used to generate amplicons in C1992/31 (resistant parent), Record (susceptible parent) and five of the most resistant

and susceptible progeny from the T3683 population. Direct PCR product sequencing (Agowa, Berlin) was performed on the amplicons using the forward primer. Following the process outlined by Sattarzadeh et al. (2006), SNPs/Indels in the amplicons were identified by manual inspection of the sequence chromatograms, and their probable dosage state estimated by virtue of the relative height/area of overlapping peaks in the sequence chromatogram using the software package DAx version 8.1 (Van Mierloo Software Consultancy) Simplex SNPs/InDels that were consistently present in the resistant parent and resistant samples and absent in the susceptible parent and susceptible samples were considered to be highly likely to be in coupling phase linkage with the $GpaIV^{s}_{adg}$ QTL. The dosage state of diagnostic SNPs in C1992/31 and 12601ab1 was reconfirmed by sequencing the appropriate PCR product using the reverse primer and re-examining the polymorphic nucleotide.

Molecular marker assays

SSR assays were carried out using either radioactively or fluorescently labelled primers. The procedures followed were as described in Febrer et al. (2007). The primer sequences for STM3016 and STM5140 are listed in Bryan et al. 2004.

AFLP assays were performed using the modified protocol of Vos et al. (1995) as described in Bryan et al. (2002, 2004). The rare cutter restriction enzyme *Pst*I (P) and the frequent cutter restriction enzyme *Mse*I (M) were used for all assays, with two and three selective nucleotides, respectively. For bulked segregant analysis (BSA) experiments, 10 µl of the AFLP pre-amp templates for the eight most resistant and eight most susceptible individuals were mixed to make resistant and susceptible bulks. AFLP assays were then performed using 5 µl of each bulk in the selective amplification.

For CAPS marker assays using the Contig237 marker, PCR reactions were performed using the same procedure employed for the radioactively labelled SSR markers, but omitting the labelling step. An annealing temperature of 54°C was used. PCR products were subsequently digested with 5U of *Taq*I restriction enzyme at 65°C for 2 h, before visualisation on a 2% agarose gel. The sequences of the Contig237 primers are

contig237_fwd	GCAGTCCTAATTGCACGTAACA
contig237_rev	CTTACTTGGGCAACCCAGAAT

As described in the previous section, 14 primer sets were designed to sequence contigs derived from BACs known to be from the $GpaIV^{s}_{adg}$ region. The sequences of the two primer sets subsequently used on the validation panel are listed below

From BAC GB005C13	Bac8_15_fwd_TCCTCTCTGGTCAGTGTTGG
	Bac8_15_rev_GAGAATCTCACGGTGGTGGT
From BAC GB080G22	BAC9_3_fwd_CGCCAACTCTGCTTCATACA
	BAC9_3_rev_TGCAGAGGTCAGTTGTTTGG

PCRs using these primers were performed as for the Contig237 primer set.

Linkage and QTL analysis in the T3683 population

Marker segregation data were analysed using Joinmap 3.0 (Van Ooijen and Voorrips 2001) and TetraploidMap (Hackett et al. 2007). A map was calculated with segregating simplex markers present only in C1992/31 using the default settings of Joinmap 3.0 (LOD score larger than 1.0, pairwise recombination estimates smaller than 0.4, a jump threshold of 5.0, a ripple value of 1.0, Kosambi mapping function). An additional microsatellite marker (STM5140), not conforming to the segregation type above was subsequently added using TetraploidMap (Hackett et al. 2007), which is designed specifically to calculate genetic maps in autotetraploids.

Single marker QTL analysis was performed using a *t* test for two samples in Minitab (2003). The coefficient of determination (r^2) of markers exhibiting significant associations with resistance to *G. pallida* Pa2/3 was calculated by performing a regression analysis on the mean square root transformed cyst scores and the corresponding marker allele scores in Minitab. The marker and trait segregation data were analysed by interval mapping using MapQTL 4.0 (Van Ooijen et al. 2002). A LOD threshold of three was chosen to declare significant QTL effects. The resulting QTL data were represented as an LOD plot against the linkage map.

Analysis of association between diagnostic markers and resistance in the breeding panel

Potentially diagnostic markers were analysed for their ability to select for resistance in the 178 members of the breeding panel by comparing the median (untransformed) cyst counts for genotypes with the appropriate diagnostic allele to those without the allele using a Mann–Whitney *U* test, in Minitab.

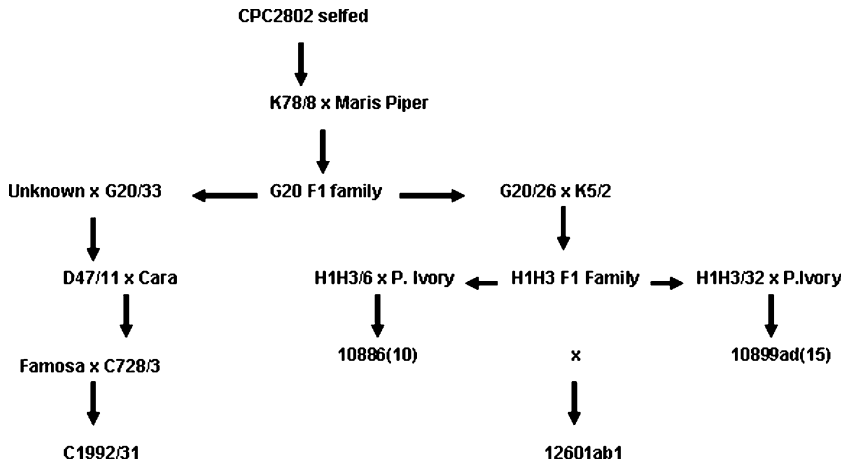

Fig. 1 Pedigree of the breeding lines C1992/31 (Oak Park) and 12601ab1 (Scottish Crop Research Institute). Clone K78/8, the G20 F1 family, the clone D47/11 and the H1H3 F1 family were produced at the Plant Breeding Institute, Cambridge, UK in the late 1960s/early 1970s. Subsequent crosses were carried out at Oak Park (*left hand side of pedigree*) and SCRI (right hand side of figure) to produce C1992/31 and 12601ab1

Results

Segregation of resistance to *G. pallida* Pa2/3 in the T3683 population

The mean cyst count over the four replicates was established for each genotype in the mapping population. This ranged from 0.8 to 44 cysts, while the resistant parent, C1992/31, had a mean cyst count of 14.3. The susceptible control, Maris Piper, had a cyst count of 58. Record was not included in the assay, although historical data indicate that it is susceptible to *G. pallida* Pa2/3. The distribution of square root transformed cyst counts over the population is illustrated in Fig. 2. The broad-sense heritability of the trait was 0.78.

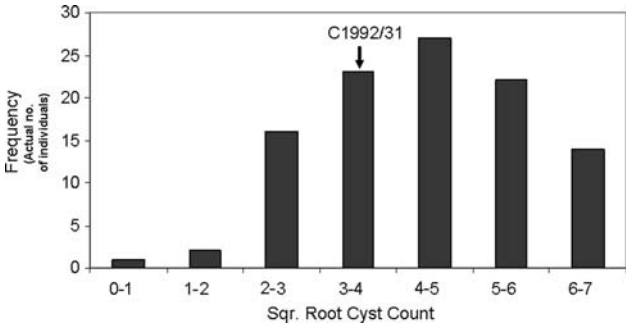

Fig. 2 Distribution of *G. pallida* Pa2/3 cyst counts (square root transformed) over the T3683 population. The resistance value of the resistant parent is indicated approximately, no data were available for the susceptible parent

Single marker QTL analysis reveals $GpaIV^{s}_{adg}$ is present in C1992/31

To investigate whether $GpaIV^{s}_{adg}$ was present in C1992/31, we first tested the ability of the SSR marker STM3016 to detect a QTL in the T3683 population. STM3016 produced three alleles in C1992/31 exhibiting segregation in the 94 progeny individuals. Using a single marker QTL analysis, only one allele (subsequently called STM3016-122/177), present in C1992/31, but absent in Record, was found to be significantly associated with resistance ($P \leq 0.001$, $t = 8.13$). On the chromatogram output of an ABI3130, STM3016-122/177 is characterised by a main peak at 122 nucleotides that is consistently accompanied by a secondary peak at 177 nucleotides. The allele segregates in a 1:1 ratio in the population (as confirmed by non-significant χ^2 for departures from this ratio: $\chi^2 = 2.084$, $P > 0.05$), indicating that it is present in single-dose in C1992/31. No other allele exhibited any association with resistance. STM3016-122/177 accounts for 45% of the variance in PCN scores in the population. These results indicate that the $GpaIV^{s}_{adg}$ QTL is present in C1992/31 in the simplex dosage state (present on only 1 of the 4 homologous chromosomes), and is linked in coupling phase to allele STM3016-122/177.

Development of a linkage map of $GpaIV^{s}_{adg}$ in C1992/31

To further characterise the $GpaIV^{s}_{adg}$ QTL in C1992/31, we generated a linkage map specific to the homologous chromosome carrying both the QTL and STM3016-122/177. Using a BSA-based approach, simplex markers linked in coupling phase to STM3016-122/177 and the QTL were identified. The analysis was carried out on bulks comprising the eight most resistant (R) and susceptible (S) clones from the T3683 population. All of the plants in the resistant bulk possessed the allele STM3016-122/177, while all of the plants in the susceptible bulk lacked the allele. The average PCN scores for the R and S bulks were 3.85 and 40.7, respectively. The two bulks and both parents were subjected to AFLP analysis using 120 *Pst*I/*Mse*I AFLP primer combinations. Primer combinations yielding at least one AFLP band that was present in C1992/31 and the

resistant bulk, but absent in Record and the susceptible bulk were selected for application to the entire population. Five bands from four primer combinations were identified as being potentially linked to $GpaIV^{s}_{adg}$ using this approach (these were designated: PagMacc1, PagMacc2, PgtMagg1, PtgMctg1, PagMata1 and PtaMagg1). On analysis in the T3683 population, all of the markers segregated in a 1:1 ratio (indicating a simplex configuration). An additional marker was developed from the primer set called Contig237 (described in "Materials and methods"). Digestion of the amplicon produced by this primer set with the enzyme *Taq*I produced a segregating allele (referred to as Contig237-a) that exhibited a similar segregation pattern to that of AFLP and SSR markers mapping to the $GpaIV^{s}_{adg}$ region in the T3683 population. Preliminary analysis of STM3016-122/177, the five AFLP markers and the Contig237-a CAPS marker in the T3683 population suggested that they were all linked in coupling phase. In terms of linkage analysis, dominant markers linked in simplex coupling phase in tetraploids can be treated in the same manner as heterozygous coupling markers in diploids, allowing the data to be analysed in a standard mapping package designed for diploids (in this case, JoinMap 3.0). On analysis in Joinmap 3.0, a genetic map of that region on chromosome IV in C1992/31 was created using the seven linked markers (Fig. 3b).

The $GpaIV^{s}_{adg}$ QTL has been extensively characterised by Bryan et al. (2004) in the genotype 12601ab1, using an interval mapping approach developed specifically for autotetraploids (Hackett et al. 2007). This yielded a marker dense map of the QTL interval comprising mainly AFLP markers, and the potato SSR markers STM3016 and STM5140, which allow the orientation of the LGIV map relative to reference maps (Fig. 3a). In order to orient the C1992/31 map relative to that of 12601ab1, the SSR STM5140 was also analysed on the T3683 population. This marker did not yield any alleles in the configuration that could be included by analysis in JoinMap. However, a single C1992/31-derived allele of STM5140, segregating in a 5:1 ratio, indicating its presence in the duplex dosage state in that genotype, was placed on the map by re-analysing the data in TetraploidMap, which handles both codominant and dominant markers, in all possible configurations in autotetraploids. On the basis of this analysis, the position of STM5140 was approximated on the Joinmap-derived map of C1992/31 (Fig. 3b). Interval mapping of the linkage and phenotypic data identified a QTL in C1992/31 exhibiting a peak at marker Contig237-a (LOD = 15, percentage of variation = 52%), flanked by markers STM3016-122/177 and PgtMagg1. Figure 3 shows a comparison of the LG IV linkage/QTL maps in 12601ab1 and C1992/31, aligned by virtue of three common markers. Although the C1992/31 map is less marker dense than that of 12601ab1, it is apparent that a large effect QTL is present in approximately the same location in the two breeding lines. Given the shared pedigree of the two breeding lines, we propose that the $GpaIV^{s}_{adg}$ QTL is present in both lines, in the simplex dosage state in C1992/31, and the duplex dosage state in 12601ab1.

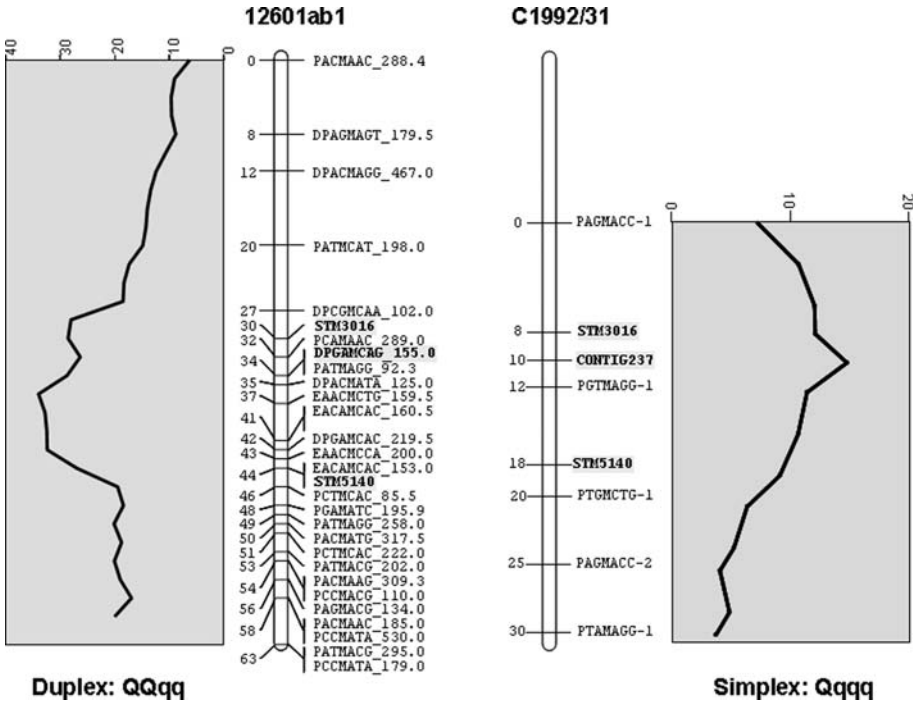

Fig. 3 Interval maps of the $GpaIV^{s}_{adg}$ *l*ocus on linkage group IV in 12601ab1 (**a** *left* from Bryan et al. 2004) and C1992/31 (**b** *right* this study). Locus-specific markers common to both studies are highlighted in *bold* on both maps. Marker Contig 237 on the C1992/31 map is equivalent to marker DPGAMCAG_155.0 (the *D* prefix indicates a duplex marker) on the 12601ab1 map. The position of a duplex allele of SSR marker STM5140 was estimated on C1992/31 as described in the "Results", the marker was not used in the interval mapping

Identifying and testing markers for diagnostic potential in the validation panel

A total of 15 primer sets were designed to BAC-derived sequence from the $GpaIV^{s}_{adg}$ region and tested on the T3683 population. Four of the primer sets (including the Contig237 primers) revealed a total of seven simplex SNPs and one single nucleotide In/Del in the simplex state in apparent coupling phase linkage with $GpaIV^{s}_{adg}$ in C1992/31.

Table 1 Simplex polymorphisms associated with resistance evident from the examination of the sequence of amplicons produced by three primer sets in C1992/31 and Record

Primer	Amplicon size (bp)	Configuration-C1992/31	Configuration-Record	Base pos C1992/31
BAC8-15	635	*TTT/A (In/Del)*	*TTTT*	241
BAC9-3	650	*CCC/G*	*CCCC*	45
BAC9-3	650	*TTT/A*	*TTTT*	87
Contig 237	423	*AAA/T*	*AAAA*	119

Total amplicon size, base position of the polymorphism in the sequence and the configurations in the sequence and their dosage

SNPs/In/Del nomenclature is according to their primer name and base position in the amplicon (see Electronic Supplementary Material for the sequences of the amplicons)

The four primer sets BAC8_15, BAC9_3, BAC7_16 and Contig 237 were applied to the entire validation panel. Even though they were associated with CPC2802 resistance in C1992/31, four of the SNPs examined apparently represented alleles present at a high frequency in the *S. tuberosum* genepool on the basis of the fact that they were also found in some of the non-CPC2802-derived clones. In contrast, three SNPs, C237 (119) BAC9-3 (45) and BAC 9-3 (87) and the In/Del, BAC8-15 (241), seem to be diagnostic. Table 1 summarises the configuration and position of the polymorphisms in the sequenced amplicons of Record and C1992/31. The actual sequences of the amplicons, with the relevant polymorphic positions highlighted, are provided as electronic supplementary material. The SSR marker STM3016 was also applied to the validation panel, and allele STM3016-122/177, which was linked to the QTL in the T3686 population, was also diagnostic for resistance in this germplasm set. For all of these markers, the target allele is only present in genotypes known to have CPC2802 in their pedigree and is never present in genotypes which do not (Table 2).

On examination of the results of the SNP/InDel markers and STM3016 on the validation panel, it is apparent that 3 out of the 12 CPC2802 derived lines have lost some or all of the potential diagnostic marker alleles described above,

Table 2 Thirty-seven genotypes comprising the validation panel and their scores for marker alleles, exhibiting diagnostic potential for the $GpaIV^{s}_{adg}$ QTL developed during this study

	GB005C13		GB080G22	
	Bac8-15 (241)	C237 (119)	Bac9-3 (45) and (87)	STM3016-122/177
Marker validation panel for CPC2802 resistance				
CPC 2802-derived genotypes				
C1992/31	+	+	+	+
C1992/42	+	+	+	+
H1H3-6	+	+	+	+
H1H3-140	+	+	+	+
10920ad9	+	+	+	+
12601ab1[a]	+	+	+	+
12636a2	+	+	+	+
12674ab1	+	+	+	+
Vales Everest	+	+	+	+
Eden	+	+	−	−
97MT200B61	−	−	−	−
Scarborough[b]	−	−	−	−
All non-CPC 2802 genotypes[c]	−	−	−	−

+ presence of diagnostic allele, − absence of diagnostic allele

[a] The only CPC2802-derived genotype for which the dosage state of *the* $GpaIV^{s}_{adg}$ QTL is definitively duplex

[b] Scarborough is susceptible to *G. pallida* Pa2/3

[c] Twenty-five cultivars and breeding lines without CPC2802 in their pedigree are present in the validation panel, these are: 10300-13, 15119ac5 (*S. vernei*-based resistance to *G. pallida* Pa2/3), 657/d3, 8204a4, Cara, Desiree, Druid, Famosa, Habibi, Lady Balfour, Maris Peer, Morene, Navan, Orla, P55/7, Panda, Pentland Ivory, Record, Rooster, Sandra, Saturna, T2270/50, T4069/22, Z89/6

through recombination or loss of the QTL-containing chromosome during meiosis. It is interesting to examine these individuals in terms of their allele score for these SNP/InDel markers and STM3016 to assess which have the greatest relative diagnostic power. For the purposes of this discussion, it is important to remember that markers BAC8-15(241) and C237(119) are derived from BAC GB005C13, while markers STM3016 and BAC9-3(45) and (87) are derived from BAC GB080G22, and indeed, marker alleles from these two clusters had the exact same scores over the entire validation panel. Thus, we are effectively examining only two genetic loci, defined on the map in Fig. 3 by STM3016 and Contig237/DPGAMCAG_155.0. The latter is closest to the peak of the QTL effect on the maps of C1992/31 and 12601ab1, and it might be assumed that it is more closely linked to the gene underlying the $GpaIV^{s}_{adg}$ QTL. Closer examination of the recombinant individuals in the T3683 population suggests that the probable arrangement of these markers relative to the underlying gene/s is STM3016—Contig237—Resistance locus. Given these observations, it would be reasonable to hypothesise that the "Contig237/ DPGAMCAG_155.0-derived" markers would be better predictors of resistance than the "STM3016-derived" markers. Of the genotypes that have lost diagnostic markers, 97MT200B61 and Eden exhibit resistance and Scarborough which is the F_1 progeny of a cross between Eden and the variety Brodick, is susceptible. All three have lost the STM3016-derived markers. Eden, which is resistant, possesses the Contig237/DPGAMCAG_155.0-derived markers, while its progeny Scarborough, which is susceptible, has lost them. However, 97MT200B61, which is three generations removed from 12601ab1 and 12636a2 and also resistant, has lost the Contig237/DPGAMCAG_155.0-derived markers. Thus, the markers representing locus Contig 237/DPGAMCAG_155.0 successfully diagnose the resistance status of the genotype on 11 out of 12 occasions in comparison to STM3016-derived markers, which successfully diagnose the resistance status on 10 out of 12 occasions.

Testing the utility of SSR marker STM3016 and SNP marker C237(119) for MAS in a breeding programme

To test the predictive abilities of the some of the markers developed in this study in a real commercial breeding programme, two markers, representing the STM3016 and Contig237/DPGAMCAG_155.0 loci (namely the STM3016 SSR marker and the C237(119) SNP marker) were tested in a panel of breeding material that comprised 178 progeny clones with C1992/31 as one parent, which were undergoing conventional selection in the fourth year of the Oak Park Potato Breeding Programme (Table 3).

Table 3 The composition of the breeding panel as described in the "Materials and methods" of the study

Cross	No. of genotypes
C1992/31 × Desiree	21
C1992/31 × 657d3	3
C1992/31 × Estima	1
C1992/31 × Habibi	1
C1992/31 × Ausonia	25
C1992/31 × Cara	6
C1992/31 × Panda	2
Buchan × C1992/31	2
Cara × C1992/31	5
Druid × C1992/31	11
Maris Peer × C1992/31	23
Morene × C1992/31	7
Navan × C1992/31	12
Pentland Ivory × C1992/31	21
Saturna × C1992/31	31
T2270/50 × C1992/31	6
Ulster Glade × C1992/31	1
Total	178

All genotypes are from the third field generation of the Oak Park breeding programme in 2004

In parallel to the genotyping experiment, the resistance of each clone in the breeding panel was assessed as previously described. Due to time constraints, and the large number of samples to be analysed, any clone with more than 100 cysts was given a nominal score of 100. The mean cyst counts ranged from 17 to 100+ cysts. Counting of representative samples from the susceptible clones revealed actual cyst counts of up to 160. The susceptible controls, Desiree and Maris Piper had mean cyst counts of 123 and 112, respectively. The "standard" resistant control Morag had a cyst count of 36. The SCRI cultivar, Vales Everest (which has $GpaIV^{s}_{adg}$) had a cyst count of 13.75, and the SCRI advanced breeding lines 12601ab1 and 12380 (*S. vernei* derived resistance) had counts of 14.5 and 1.25, respectively. The distributions of untransformed cyst counts for breeding panel members with and without markers STM3016-122/177 and C237(119) are shown in Fig. 4. The medians are 37.75 (with STM3016-122/177) and 75.75 (without), and 33.50 (with C237(119)) and 75.75 (without). Both differences are statistically significant ($P < 0.0001$, Mann–Whitney U test). Similar results were obtained when each of the specific crosses was analysed independently. Figure 4 shows the consequences of using the two markers for MAS in a breeding programme. The use of either marker results in the selection of a subpopulation of individuals which exhibits a median corresponding to the resistant end

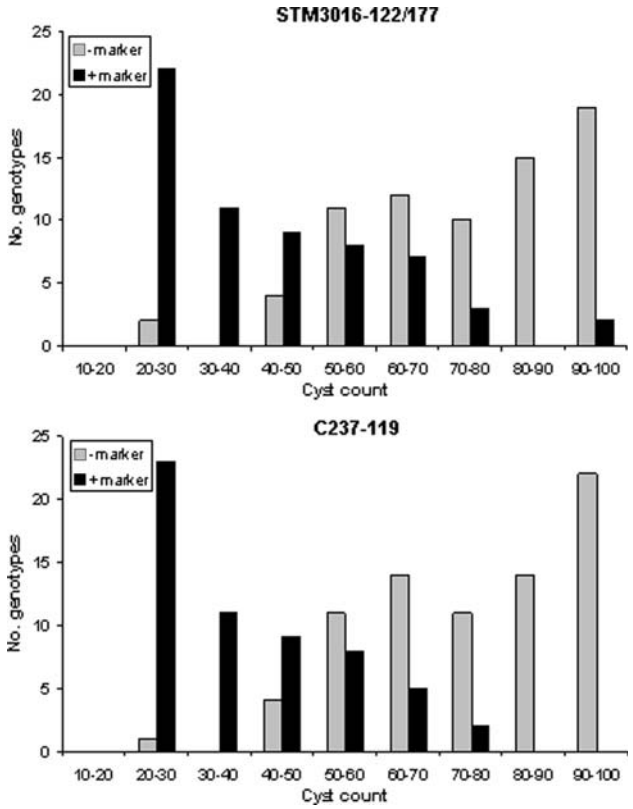

Fig. 4 Histograms of the distribution of cyst counts (untransformed) amongst the 178 individuals of the breeding panel with (*black*) and without (*grey*) the $GpaIV^{s}_{adg}$ diagnostic marker alleles STM3016-122/177 and C237(119)

of the spectrum. The practical implications of these results for MAS are discussed below.

Discussion

By demonstrating the presence of the $GpaIV^{s}_{adg}$ QTL in relatively divergent material from two independent breeding programmes, we have confirmed that it is a viable target for MAS-based strategies to develop potato varieties exhibiting resistance to *G. pallida* pathotype Pa2/3. Our initial comparisons between 12601ab1 and C1992/31 were based on the anchor marker STM3016, and one allele of this marker (STM3016-122/177) subsequently exhibited diagnostic potential for $GpaIV^{s}_{adg}$ across a wide range of germplasm. However, STM3016 is several centimorgans from the peak of the QTL in the maps of both 12601ab1 and C1992/31, and the results in the validation panel suggest that the diagnostic allele may occasionally be uncoupled from the resistance gene by recombination. In addition, unless they reveal four distinct alleles at a locus, microsatellite markers give no indication of the dosage state of the associated target gene, a factor which is of great interest in a tetraploid, since more highly multiplex parents transmit the target gene to the progeny at a greater frequency. The diagnostic SNP/InDel polymorphisms developed during the study will provide a basis for the development of dosage-specific markers for breeding. Examination of diagnostic SNP/InDel polymorphisms between C1992/31 and 12601ab1, the only genotype in the study known to be duplex for $GpaIV^{s}_{adg}$ generally reveals a shift in peak heights, indicating the expected duplex state in the latter genotype (data not shown). However, as outlined by Sattarzadeh et al. (2006), this method of dosage detection is subject to inconsistent results due to factors such as the influence of the surrounding nucleotides or events such as asymmetric amplification of different alleles. For this reason, and the high per assay cost of Sanger sequencing, dosage detection based on the methods such as pyrosequencing, allele-specific hybridisation or quantitative PCR would be more appropriate for routine application in breeding.

Our results suggest that the gene/s responsible for the QTL effect in C1992/31 might lie between markers Contig237 and PGTMAGG-1. In the longer term, it would be useful to develop another SNP-based marker, or a series of diagnostic markers flanking the QTL, allowing a highly informative marker haplotype, rather than an individual marker genotype to be used to test for the presence of the introgression segment, eliminating the problem of recombination between the marker and the gene. Although it might be possible to convert the aforementioned AFLP marker for this purpose, further BAC sequencing in this region, followed by SNP screening in the same manner as performed in this study, is more likely to achieve the development of such markers.

The experiments on the validation panel offer a retrospective analysis of the diagnostic ability of $GpaIV^{s}_{adg}$ -linked markers in advanced breeding material. However, our goal was to develop markers for MAS, we felt it was also necessary to assess their utility as a selection tool in a commercial potato breeding programme in a manner similar to the way that they could potentially be employed in that programme. To achieve this, STM3016 and the C237(119) SNP marker were tested in a panel of breeding material that comprised 178 progeny clones with C1992/31 as one parent, from year 4 of the Oak Park Potato Breeding Programme. At this point in the programme, there are generally 2,500 genotypes undergoing field evaluation in plots comprising 10 plants per genotype. Within the programme, PCN testing is generally not carried out until the subsequent year, when 300 genotypes are generally retained, and even then, testing is carried out on a relatively limited amount of material. Deployment of marker-based testing for CPC2802-derived resistance at the year 4 stage of the programme would allow it to be combined with a

reasonable level of traditional selection for generally commercially desirable phenotypes. This could then be followed by confirmation in the selected clones of PCN resistance and other desirable traits. Thus, testing the marker in this material gives a useful measure of the probable utility of the marker in a "real-world" breeding scenario.

As outlined in the results, the use of the diagnostic markers resulted in the selection of a subpopulation of individuals from the breeding panel that exhibits a median corresponding to the resistant end of the spectrum. However, there is still considerable variation around the medians, and this must be due to factors such as recombination breaking the linkage between the markers and $GpaIV^{s}_{adg}$, environmental variation in the PCN tests and genetic variation at other loci. The wild species *S. vernei*, which is another common source of resistance to *G. pallida*, is in the pedigree of several of the non-CPC2802 parents of the breeding panel. Although all of these parents are considered susceptible to varying degrees, we tested them with the HC and SPUD1636 markers (Sattarzadeh et al. 2006; Bryan et al. 2002) which are diagnostic for the large effect chromosome V resistance QTL from this source (results not shown). None were positive for either marker, largely eliminating any contributions from that locus, although it is possible that other, smaller effect loci from *S. vernei* and other sources are present.

The results in the breeding panel suggest that the breeder could confidently use the markers to practise negative selection at this stage in the breeding programme, discarding all genotypes that lack the marker, since there are very few instances in which the most resistant or susceptible genotypes are incorrectly predicted by the marker genotype for either marker. Thus, adopting a negative selection strategy would eliminate the most susceptible individuals from the population. Although useful for raising the median levels of resistance in a breeding programme, this approach may not be the best strategy for breeding cultivars expressing very high levels of resistance to *G. pallida* pathotype Pa2/3. It is clear from this and other studies that no single resistance source gives complete resistance to all populations of *Globodera pallida* pathotypes Pa2 and Pa3. The main reason for this likely to be the relatively complex structure of the PCN populations, leading to heterogeneity of avirulence factors in these populations. Using MAS to combine multiple resistance loci with differing specificities has the potential of broadening the resistance spectrum of varieties, resulting in more durable resistance. In addition, combining different resistance loci can result in apparent additive effects on resistance levels as Rouppe van der Voort et al. (2000) found for *Gpa5* and *Gpa6*. The effect can be quite pronounced, for instance, Caromel et al (2005) found that combining $GpaV^{s}_{spl}$ and $GpaXI^{s}_{spl}$ resulted in the expression of a strong hypersensitive response in the roots of genotypes containing both loci, which was not present when either locus was present individually. Conversely, combining different resistance loci does not always result in an additive effect. For example, Tan (2008) demonstrated that combining root knot nematode resistance QTL from *S. tarijense* and *S. chacoense* did not result in any reduction in infestation beyond that conferred by the more effective of the two loci.

MAS then, is ideal not only for the actual process of pyramiding *G. pallida* pathotype Pa2/3 resistance loci from various sources, but also in helping to identify the most effective combinations of these loci. The availability of the diagnostic markers described in this study will allow the combination of the $GpaIV^{s}_{adg}$ QTL with other resistance loci to develop potato varieties, with high levels of resistance to mixed populations of *G. pallida* pathotypes Pa2 and Pa3.

Acknowledgments CM was supported by a Teagasc Walsh Fellowship during the course of this study. DM is supported by the National Development Plan (NPD) of Ireland. The potato breeding programme at Oak Park (DG) is supported by the NDP and Irish Potato Marketing Limited. GJB and JEB acknowledge financial support of the Scottish Government Rural Environment Research and Analysis Directorate (RERAD). The authors wish to thank Dr Della Milbourne for help with the figures and proofreading the manuscript, Dr Helena Meally for help with fluorescent genotyping of STM3016 and Emmet Dalton for genotyping the HC and SPUD1636 markers.

References

Bradshaw JE, Hackett CA, 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. Theor Appl Genet 97:202–210

Bryan GJ, McLean K, Bradshaw JE, Phillips M, Castelli L, DeJong W, Waugh R (2002) Mapping QTLs for resistance to the cyst nematode *Globodera pallida* derived from the wild potato species *Solanum vernei*. Theor Appl Genet 105:68–77

Bryan GJ, McLean K, Pande B, Purvis A, Hackett CA, Bradshaw JE, Waugh R (2004) Genetical dissection of H3-mediated polygenic PCN resistance in a heterozygous autotetraploid potato population. Mol Breed 14:105–116

Caromel B, Mugniery D, Lefebvre V, Andrzejewski S, Ellisseche D, Kerlan MC, Rousselle P, Rousselle-Bourgeois F (2003) Mapping QTLs for resistance against *Globodera pallida* (Stone) Pa2/3 in a diploid potato progeny originating from *Solanum spegazzinii*. Theor Appl Genet 106:1517–1523

Caromel B, Mugniery D, Kerlan MC, Andrzejewski S, Palloix A, Ellisseche D, Rousselle-Bourgeois F, Lefebvre V (2005) Resistance quantitative trait loci originating from *Solanum sparsipilum* act independently on the sex ratio of *Globodera pallida* and together for developing a necrotic reaction. MPMI 18(11):1186–1194

Febrer M, Cheung F, Town CD, Cannon SB, Young ND, Abberton MT, Jenkins G, Milbourne D (2007) Construction, characterisation and preliminary BAC-end sequencing of a bacterial artificial

chromosome library of white clover (*Trifolium repens* L.). Genome 50:412–421

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

Hackett CA, Milne I, Bradshaw JE, Luo Z (2007) TetraploidMap for windows: linkage map construction and QTL mapping in autotetraploid species. J Hered 98:727–729

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

Minitab Inc (2006) Minitab Statistical Software, Release 15 for Windows, State College, Pennsylvania Minitab® is a registered trademark of Minitab Inc

Niewohner J, Salamani 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. Mol Breed 1:57–78

Park TH, Gros J, Sikkema A, Vleeshouwers VGAA, Muskens M, Allefs S, Jacobsen E, Visser RGF, van der Vossen EAG (2005a) The late blight resistance locus *Rpi-blb3* from *Solanum bulbocastanum* belongs to a major late blight R gene cluster on chromosome 4 of potato. MPMI 18:722–729

Park TH, Vleeshouwers VGAA, Hutten RCB, van Eck HJ, Van Der Vossen E, Jacobsen E, Visser RGF (2005b) High-resolution mapping and analysis of the resistance locus *Rp1-abpt* against *Phytophthora infestans* in potato. Mol Breed 16:33–43

Phillips M, Forrest J, Wilson L (1980) Screening for resistance to potato cyst nematode using closed containers. Ann Appl Biol 96:317–322

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

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

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

Sattarzadeh A, Achenbach U, Lubeck J, Strahwald J, Tacke E, Hofferbert HR, Rothsteyn T, Gebhardt C (2006) Single nucleotide polymorphism (SNP) genotyping as basis for developing a PCR-based marker highly diagnostic for potato varieties with high resistance to *Globodera pallida* pathotype Pa2/3. Mol Breed 18:301–312

Tan AMY (2008) Genetic mapping and pyramiding of resistance genes in potato. PhD Thesis, Wageningen University, The Netherlands, ISBN 978-90-8504-977-7

Van Ooijen JW, Voorrips RE (2001) Joinmap 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, The Netherlands

Van Ooijen JW, Boer MP, Jansen RC, Maliepaard C (2002) MapQTL4.0, Software for the calculation of QTL positions on genetic maps. Plant Research International, Wageningen, The Netherlands

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

Vos P, Hogers R, Bleeker M, Rerjans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucl Acid Res 23:4407–4444

Wolters P, Vinke H, Ilja Bontjer I, Rouppe van der Voort J, Colon L, Hoogendoorn C (1999) Mapping of major genes for resistance to *Globodera pallida* in wild *Solanum* species. Poster Abstract 362, Plant and Animal Genome Conference VII, Town & Country Hotel, San Diego, CA, 17–21 Jan 1999