Theoretical and Applied Genetics (2025) 138:52
https://doi.org/10.1007/s00122-024-04794-8

ORIGINAL ARTICLE

Identification by GWAS of marker haplotypes relevant to breed potato for *Globodera pallida* resistance

J. Leuenberger[1,2] · F. Esnault[1] · P. L. Lebas[1] · S. Fournet[3] · M. P. Cann[1] · S. Marhadour[1,4] · C. Prodhomme[3,4] · M. L. Pilet-Nayel[3] · M. C. Kerlan[1]

Received: 30 July 2024 / Accepted: 30 November 2024 / Published online: 24 February 2025

Abstract

Key message **Identified and validated QTL *GpaVa_MRQ* and *GpaIX_MRQ* provide robust tools for improving potato resistance to *Globodera pallida* via marker-assisted selection.**

Abstract Potato (*Solanum tuberosum* L.), a vital food crop globally, faces significant yield losses due to potato cyst nematodes (PCN). This study aimed to identify and validate genomic regions conferring resistance to *Globodera pallida*, to facilitate the development of resistant potato varieties through marker-assisted selection (MAS). We conducted Genome-Wide Association Studies (GWAS) on a pre-breeding panel genotyped using Genotyping by Sequencing (GBS) and the SolCAP DNA array. Significant resistance-associated SNP markers were identified on chromosomes III, IV, V, IX and XI. Quantitative trait loci (QTL), including the major-effect QTL *GpaVa_MRQ* on chromosome V and the QTL *GpaIX_MRQ* on chromosome IX, were validated and shown to account for substantial phenotypic variance in a validation potato panel. Haplotype-based marker sets were defined at four QTL regions, enabling the practical application of MAS. The successful conversion of SNPs to PACE markers at the two main QTL *GpaVa_MRQ* and *GpaIX_MRQ* further supports their use in breeding programs. This study provides valuable insights and robust tools for enhancing potato resistance to *G. pallida*, contributing to sustainable agricultural practices and global food security.

Introduction

Potato (*Solanum tuberosum* L.) ranks as the fourth most important food crop worldwide with a production of 17 Mega tons, supporting food security and economic stability across various regions. Its global cultivation and good nutritional value highlight its crucial role in meeting dietary needs and addressing challenges of global food security.

Potato cyst nematodes (PCN) *Globodera rostochiensis* and *Globodera pallida* are recognized as major quarantine pests due to their potential to cause up to 90% of crop loss (Pulavarty et al. 2022). Due to their conservation form as cyst, PCN can survive during twenty years or more (Wainer and Dinh 2021), which implies the implementation of control methods. The management of these pathogens involves different strategies including seed certification, crop rotation, chemical control and plant resistance (Price et al. 2021).

The use of nematicides is more and more banned due to their detrimental impacts on soil microbiota and the risk of engendering resistant nematode populations. Besides, most of these nematicides have been removed from the market and the development of sustainable management practices are needed (Palomares-Rius et al. 2021).

The vast genetic diversity present in wild potato species enabled the identification of genomic regions associated with resistance to PCN. For both nematodes, several genes have been identified and cloned (Rouppe van der Voort et al. 2000; Ernst et al. 2002; Paal et al. 2004). The majority of

Communicated by Benjamin Stich.

✉ J. Leuenberger
Julien.leuenberger@inrae.fr

1 Institut de Génétique, Environnement et Protection des Plantes (IGEPP), INRAE, Institut Agro, Univ Rennes, Ploudaniel, France

2 Association Des Créateurs de Variétés Nouvelles de Pomme de Terre (ACVNPT), Paris, France

3 Institut de Génétique, Environnement et Protection des Plantes (IGEPP), INRAE, Institut Agro, Univ Rennes, Le Rheu, France

4 Inov3PT, Recherche, Développement, Innovation des Producteurs de Plants de Pomme de Terre, Paris, France

the genetic loci involved in resistance are Quantitative Trait Loci (QTL). Regarding resistance to *G. rostochiensis*, the *H1* gene is the most important one and appears stable across years (Gartner et al. 2021). The genetic architecture of resistance to *G. pallida* is more complex. One major-effect QTL, originating from *S. vernei*, has been identified on the north of chromosome V (G*paV*) (Rouppe van der Voort et al. 2000). This major-effect QTL has been successfully introgressed by breeders into commercial varieties. Due to its intensive use, some nematode populations have evolved to overcome it. Indeed, some virulent populations have been observed in fields in Germany (Mwangi et al. 2019) and confirmed what has been shown in controlled conditions, which reinforced the need to develop new resistant varieties based on diversified resistance sources (Fournet et al. 2013). Other QTL have been identified in different wild species, including *S. vernei* (Bryan et al. 2002), *S. tuberosum ssp andigena* (Bradshaw et al. 1998), *S. spegazzinii* (Caromel et al. 2003; Gartner et al. 2024), *S. sparsipilum* (Caromel et al. 2005) and *S. tarijense* (Adillah Tan et al. 2009). These major resistance QTL have been less used in breeding programs due to their linkage to deleterious alleles for other traits or the absence of effective molecular markers tagging these resistance regions (Hasan et al. 2021).

The pyramiding of resistance genes and QTL with major or minor effect is a promising approach to improve potato resistance to PCN (Van der Plank 1968; Rouppe van der Voort et al. 2000; Caromel et al. 2005; Dalton et al. 2013; Pilet-Nayel et al. 2017). This strategy has the potential to enhance potato production resilience, for instance the use of marker-assisted selection led to breed "Buster" cultivar that combines resistance to *G. rostochiensis* and *G. pallida*. In order to facilitate this, the development of molecular markers is crucial. Indeed, breeding for some quantitative resistance traits is almost impossible based solely on phenotyping.

Several Genome-Wide Association Studies (GWAS) have been performed in potato breeding, facilitating the identification of genetic loci associated with desirable traits, such as resistance to diseases and pests (Marhadour and Prodhomme 2023), abiotic stress such as drought tolerance (Alvarez-Morezuelas et al. 2023), tuber quality (Pandey et al. 2022), agronomic traits (Vexler et al. 2024), or starch yield (Khlestkin et al. 2020). The implementation of GWAS has been facilitated by the development of models that consider kinship, which helps reducing the false discovery rate and promotes the identification of true genetic associations (Yu et al. 2006). The mixed linear model (MLM) is the most commonly used model in potato GWAS. Additionally, MLM has been adapted for use in tetraploid crops like potatoes with the development of the GWASpoly package (Rosyara et al. 2016). The Multi-Locus Mixed Model (MLMM) (Segura et al. 2012) has been utilized in other species to address complex genetic traits affected by multiple genes. Unlike traditional single-locus models, MLMM incorporates multiple loci simultaneously into the analysis, enhancing the understanding of genetic architecture of the trait (Tibbs Cortes et al. 2021).

Genotyping by sequencing (GBS) is increasingly employed in GWAS. This technology offers a cost-effective means to identify a high number of SNP markers. Unlike DNA arrays, GBS does not suffer from ascertainment bias. Bastien et al. (2018) demonstrated that GBS could significantly enhance the potential of GWAS for potatoes by providing high-quality SNPs. However, GBS is not entirely free from bias. Polymorphisms in restriction sites can influence which fragments are sequenced, potentially leading to underrepresentation of certain alleles, and the method can result in higher rates of missing data due to uneven sequencing coverage (Rasheed et al. 2017). Despite these limitations, GBS remains a valuable tool for capturing a broader range of genetic diversity. The prerequisite for sequence alignment in GBS is a well-assembled genome. The first potato genome was sequenced in 2011 by the Potato Genome Sequencing Consortium (The Potato Genome Sequencing Consortium 2011), and since then, long read sequencing allowed to improve the quality of this reference sequence (Pham et al. 2020).

The possibility to convert SNP markers coming from GBS or DNA arrays to technologies like KASP (Kompetitive Allele Specific PCR) or PACE (PCR Allele Competitive Extension) is crucial for carrying out marker-assisted selection programs (Semagn et al. 2014; von Maydell 2023). This process is more suitable to perform the genotyping of a high number of individuals with a few markers making possible the early selection of individuals with desirable traits. For resistance to *G. pallida*, Sattarzadeh et al. (2006) developed the PCR marker HC which allows a better detection of *GpaV* QTL than the use of the SPUD1636 marker (Bryan et al. 2002). Previous works showed that even if both markers allow the detection of G*paV,* they are not linked (Kerlan, personal communication). The HC marker is now routinely used by breeders (Milczarek et al. 2011; Asano et al. 2021).

Marker haplotypes proved to be of great interest in plant breeding, particularly in genetically complex crops like potatoes. An haplotype can be defined, as mentioned by Bhat et al. (2021), as a combination of alleles for different polymorphisms (such as SNPs, insertions/deletions and other markers or variants) present on the same chromosome, which are inherited together with minimum chance of contemporary recombination. In potato, haplotypes have been designed by van Eck et al. (2017) using graphical genotyping. Haplotype construction has also been used by Klaassen et al. (2019) to identify genomic regions associated with protein content. Marker haplotypes are more specific of the desired allele whereas a single marker may turn out to be non-specific in a larger panel.

This study aimed to (i) precisely confirm major or minor-effect QTL previously reported for resistance to *G. pallida* and potentially identify new ones, (ii) identify QTL combinations associated with increased resistance and (iii) develop useful marker haplotypes to target these QTL combinations in breeding. To achieve this objective, we conducted GWAS for resistance to *G. pallida* from a panel of potato pre-breeding clones that was previously genotyped using three genotyping technologies. We defined sets of closely linked SNPs in QTL regions significantly detected by GWAS for resistance to *G. pallida*. Then, we converted the most useful SNP markers from these sets into PACE markers and verified the significant association between PACE marker haplotypes in main QTL regions and resistance to *G. pallida* in a validation panel of genotypes.

Material and methods

Plant material and genotypic data

This study was carried out using 247 INRAE advanced breeding clones of the potato MRQ (Multi-Resistance and Quality) panel (Table S1). The breeding clones were created from various INRAE research programs. This panel is relevant for studying genetic resistance to various pathogens. The MRQ panel was genotyped using three different genotyping techniques (GBS, SolCAP DNA array (Felcher et al. 2012) and PCR markers) resulting in the identification of 27,148 high-quality markers, including 22,087 GBS SNPs, 5053 SolCAP SNPs and 8 PCR-based markers among which HC and SPUD1636. The GBS SNPs were identified by mapping the reads onto the DM reference genome v6.1 (Pham et al. 2020). Genotyping data are described and available in Leuenberger et al. (2024). A subset of 89 genotypes was chosen from that panel to validate the conversion of selected SNPs from GBS to the PACE genotyping method (Table S1).

A second panel of potato genotypes was established and used to check the association between haplotypes at the PACE markers and resistance to *G. pallida*. This panel included 61 breeding clones from INRAE research programs (Table S2). These genotypes present a diversity of resistance or susceptibility levels to *G. pallida.* The resistance sources in the panel were mainly derived from *S. vernei*. This panel was referred to as "Validation panel".

Genetic structuration of the MRQ panel

In order to analyze the genetic structure of the MRQ panel, a subset of 1800 SNP markers was carefully selected using the following criteria: (i) only markers with less than 5% missing data and at least 10 genotypes carrying the alternate allele were included in this subset, and (ii) 150 SNPs were then randomly chosen on each chromosome to ensure genome representation. To confirm reproducibility of the structure analysis, multiple random draws of 150 SNPs per chromosome were performed, yielding consistent results (results not shown).

Principal Component Analyses (PCA) were performed using the FactoMineR package (Lê et al. 2008) and multiple draw of 150 markers per chromosome. The FactoExtra package (Kassambara and Mundt 2020) was used to visualize the PCA results. Each genotype was qualified with the agronomic trait targeted in the research program it comes from (Leuenberger et al. 2024).

A Kinship matrix was calculated using the same set of markers and the VanRaden method (VanRaden 2008). The kinship heatmap was performed using the heatmap function from the Stats package.

Phenotyping for resistance to *G. pallida*

The 247 advanced breeding clones from the MRQ panel were evaluated for resistance to *G. pallida* pathotype Pa2/3 (Chavornay population), using the 'Foot test' as described by Foot (1977) (Table S1). The experiment was conducted by HLB (HLB BV, The Netherlands) and included four independent replicates for each genotype. For each replicate, a single tuber was planted in a 125-cm^3 transparent closed container filled with soil containing 5 juveniles of *G. pallida* per gram of soil. Two cultivars were added as controls, including 'Désirée' as the susceptible control and 'Seresta' as the resistant one. After 10 weeks, newly formed cysts were visually counted through the container. Two scoring variables were defined, including (i) Cyst Presence (CP) which is a qualitative variable with two scores, 0 if there is no cyst and 1 if there is at least one cyst, and (ii) Cyst Number 1 (CN1) which is a quantitative variable defined by the mean number of cysts counted on the genotypes presenting at least one cyst. For CN1, genotypes having no cyst have been excluded. Phenotyping data were analyzed using the Restricted Maximum Likelihood (REML) method in the BreedR package (Muñoz and Sanchez 2019) which allowed to estimate variance components. Broad-sense heritability (H^2) was estimated for each variable from variance components using the formula σ^2_G/σ^2_P where σ^2_G is the genotypic variance and σ^2_P the phenotypic variance.

Phenotypic data of the 61 breeding clones from the Validation panel were previously obtained for resistance to *G. pallida* (Table S2). Two phenotyping methods were used, including a "Foot test," as described for the MRQ panel, and a "pot-test". Pot-tests were carried out with four independent replicates for each genotype. For each replicate, a single tuber was planted in a 400-cm^3 plastic pot filled with a mixture of soil and sand (respectively, 2/3; 1/3) containing

10 cysts of *G. pallida* to achieve an infestation rate of 5 to 10 juveniles per gram of soil. After four months, newly formed cysts were counted for each plant individually after their extraction from the soil using a Kort elutriator. The genotypes of this panel were evaluated in various years. The use of the cultivar 'Désirée' as a susceptible control each year ensured the consistency of the data. For each genotype, the susceptibility rate to *G. pallida* was calculated using the formula: $(\text{Cyst\_Number}_{\text{genotype}}/\text{Cyst\_Number}_{\text{Désirée}}) \times 100$. Then the clones were classified into four groups according to their susceptibility rate, *i.e.,* > 10% as Susceptible (S); 5 to 10% as Slightly Susceptible (SS), 1 to 5% as Resistant (R), and < 1% as Very Resistant (VR).

Genome-wide association study

The Multi-Locus Mixed Model (MLMM) was employed to perform GWAS on the MRQ panel, using the mlmm.gwas R package (Bonnafous et al. 2018), as described by Segura et al. (2012). The MLMM is a modified mixed linear model that uses forward–backward stepwise linear regression to include significant SNP markers as fixed-effect cofactors. MLMM analyses were done using two models: the "additive model" and the "dominant model." In the additive model, the genotypes were coded according to the dosage of the alternate allele as followed: 0 for nulliplex, 0.25 for simplex, 0.5 for duplex, 0.75 for triplex and 1 for quadruplex. In the dominant model, the genotypes were coded as 0 if the alternate allele is not present and as 1 if the alternate allele is present, whatever its dosage. Both models were applied to the two variables CP and CN1 previously defined with a minor allele frequency (MAF) of 0.05. To determine the most appropriate number of steps in the MLMM analyses, we applied an adjusted Bonferroni correction with a p value of 0.05.

Definition of marker sets at GWAS peaks

At each GWAS peak, the most significant marker was extracted. Subsequently, additional markers in linkage disequilibrium (LD) with this peak marker were identified. Only the four most correlated markers, exhibiting a Pearson correlation coefficient with the peak marker higher than 0.9, were retained. When less than three markers fulfilled this criterion, the GWAS peak was not considered for subsequent analysis. For defining haplotypes, *i.e.,* combinations of alleles for the SNP markers composing the set, missing data in the case of a single SNP was imputed using the k-Nearest Neighbors (KNN) algorithm. At each locus, a genotype was defined as carrying the "Alt" or "Ref" haplotype when the alternate or reference allele, respectively, was present at all markers in the set. The genotypes which presented a recombination profile within the marker set were removed from subsequent analyses. The same approach was applied to identify a marker set in LD with marker SPUD1636. This marker is known to be associated with a QTL for resistance to *G. pallida* on chromosome V (Bryan et al. 2002). A haplotype–phenotype association analysis was then conducted in the MRQ panel at each marker set defined around a GWAS peak. A mixed model employing Restricted Maximum Likelihood (REML) was utilized, considering the marker set and the kinship matrix as random effects. The proportion of phenotypic variance explained by each locus was subsequently calculated for the variable that allowed the detection of the locus.

SNP conversion from GBS to PACE markers

For a use in marker-assisted selection, the development of PACE markers was performed for the SNPs of the marker sets associated with a higher level of resistance. The sequences surrounding each SNP, encompassing 50 bases both upstream and downstream, were extracted. Sequence polymorphisms were examined within these sequences, based on the aligned GBS data before variant selection. These sequences were then submitted to 3CR (CR Bioscience, Essex, UK) for primer definition. Primer synthesis was carried out by Integrated DNA Technologies Inc. (Munich, Germany). The PACE markers were then used to genotype the subset of the MRQ panel and the Validation panel.

Total genomic DNA was extracted from frozen leaves using a protocol derived from Doyle and Doyle (1990). The genotyping procedure was conducted in accordance with the protocol provided by 3CR Bioscience, utilizing their Master mix. This mix comprised a buffer solution optimized for KASP reactions, fluorescent tags (FAM and HEX), Taq polymerase, and a mix of free nucleotides. For the genotyping assay, a primer mix containing a shared antisense primer along with two allele-specific sense primers was prepared. The PCR reaction was performed using 5 µL of genomic DNA (at a concentration of 10 ng/µL), 5 µL of the KASP Master mix, and 0.138 µL of the primer mix. Control wells were set up with nuclease-free water to ensure assay integrity. After sealing the plates and centrifugation, the samples were subjected to PCR amplification in a Roche LightCycler 480 system. The amplification process consisted in a 15-min enzyme activation step at 94 °C, 10 cycles of denaturation at 94 °C for 20 s and annealing–extension for 1 min at 65 °C to 57 °C, with a decrement of 0.8 °C per cycle, and 30 cycles of denaturation at 94 °C for 20 s and annealing–extension for 1 min at 57 °C, followed by a final fluorescence read at 37 °C for 1 min. Endpoint analysis was carried out to determine the fluorescence intensities corresponding to the FAM and HEX fluorescence levels. To refine the quality of clustering, additional amplification steps were

employed when needed, consisting in 3 cycles of denaturation at 94 °C for 20 s and annealing–extension for 1 min at 57 °C, followed by a final fluorescence read at 37 °C for 1 min.

For each SNP, each individual was assigned to one of the four genotypic classes (AAAA, AAAB, AABB, ABBB, BBBB) using the KNN clustering algorithm, which was based on fluorescence intensities. Manual verification was conducted on the algorithmic outputs. Data points identified as outliers or those positioned ambiguously between clusters were coded as 'NA'. The concordance between the SNP genotyping data obtained from GBS and PACE in the MRQ subset panel was quantified using Cohen's Kappa index, with the degree of agreement interpreted according to the Landis and Koch (1977) scale. This scale provides a qualitative measure of agreement, ranging from poor to almost perfect. The PACE markers that demonstrated high concordance with GBS (Kappa index > 0.8) were subsequently used to genotype the Validation panel. The association between PACE marker haplotypes and resistance to *G. pallida* was evaluated using a Chi-2 test, with a *p* value threshold set at 0.05 corrected for multiple tests with a FDR method.

Results

Genetic structure of the MRQ panel

The PCA performed on MRQ panel from multiple draw of 1800 SNPs gave similar genetic structure (Fig. 1). They showed that the two first principal components captured only 7.7% of the total variance. The "*G. pallida*" and "*Pectobacterium*" breeding groups were more widely distributed across the PCA space. The "Multi-resistance" group appeared to be included within the "*G. pallida*" group. The other breeding groups were more aggregated and located toward the center of the PCA plot. All the groups overlapped. Therefore, the PCA analysis did not show a clear structuration of the MRQ panel. This lack of structuration was confirmed by the Kinship analysis (Fig. S1), where no distinct groups were visually discernible.

Phenotypic data for resistance to *G. pallida* in the MRQ panel

Some growth anomalies, such as the absence of root formation, occurred during the phenotyping experiment of the 247 advanced breeding clones, resulting in the exclusion of four genotypes from subsequent analyses. In the 243 remaining clones, the cyst number data showed a zero-inflated distribution (Fig. 2). Specifically, 13% of the genotypes did not present any cyst. Among the 211 genotypes presenting at least one cyst, 80% had between 5.5 and 41 cysts, with a mean of 24 cysts per genotype (Table S1). The cyst number for the control genotypes were, as expected, 0 for Seresta and 65 for Désirée. Broad-sense heritability for CP and CN1 was 0.70 and 0.48, respectively.

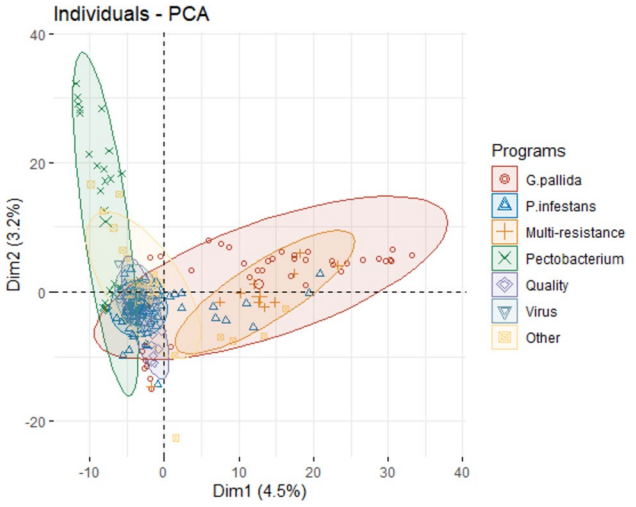


Fig. 1 Principal component analysis performed on the MRQ panel using a subset of 1800 SNPs. Each genotype belongs to a group named according to the agronomic trait targeted in the research program it comes from as followed: "*G. pallida*" for resistance to *G. pallida,* "*P. infestans*" for resistance to late blight, "Multi-Resistance" for the combination of resistance to late blight with resistance to *G. pallida* or *G. rostochiensis,* "*Pectobacterium*" for resistance to *Pectobacterium* ssp., "Quality" for cold-induced sweetening or after cooking blackening, "Virus" for resistance to PVX, PVY, or PLRV viruses, "Variety" for registered potato cultivars; "Other" for less represented agronomic traits, such as resistance to *Fusarium* spp. or *Meloidogyne incognita*

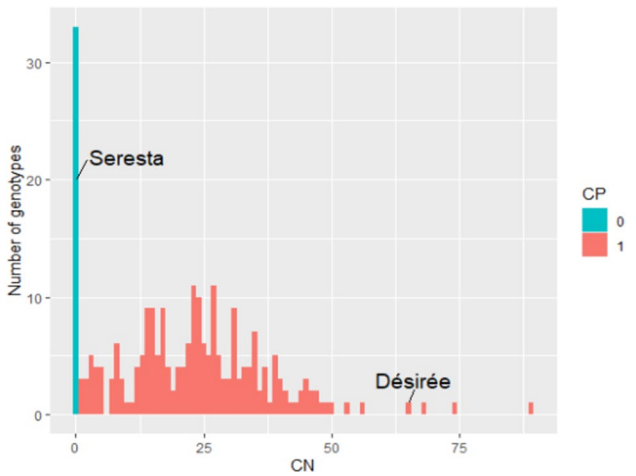


Fig. 2 Frequency distribution of *G. pallida* cyst number in the MRQ panel. Blue color represents the genotypes without any cyst (CP=O), red color represents the genotypes which have at least one cyst (CP = 1). The position of the susceptible control (cv Désirée) and the resistant control (cv Seresta) is shown

GWAS in the MRQ panel

The set of 27,148 markers from Leuenberger et al. (2024) was used to perform GWAS for the two phenotypic traits CP and CN1. For CP, five significant markers were detected using the dominance model. These significant SNPs were located on chromosomes III, IV, V and IX (Table 1, Fig. S2). On chromosome V, these markers were positioned on both arms of the chromosome and the marker located on the northern arm (chr05_5423448) was detected with the highest significance (p value = 7.6 e−16). For CN1, two significant markers were detected using the additive model. The markers were located on chromosomes III and XI. The significant markers identified on chromosome III for the two variables were different and 15 Mb apart. All identified SNPs were derived from GBS.

Table 1 GWAS results for CP (Cyst Presence) and CN1 (Cyst Number for the genotypes having at least one cyst) variables, using MLMM. The name of each marker indicates its chromosomal and physical position according to the reference genome DM v6.1

Variable	Model	Peak marker	p value
CP	Dominance	chr03_43430790	2.3 e−6
		chr04_8013962	1.6 e−7
		chr05_5423448	7.6 e−16
		chr05_46793576	3.8 e−9
		chr09_65310251	7.7 e−8
CN1	Additive	chr03_58425552	8.2 e−7
		chr11_26392764	2.6 e−6

Marker sets at GWAS peaks

Analysis of local LD at each significant GWAS peak identified more than three markers in high LD with the peak marker ($r^2 > 0.9$) for the regions detected on chromosomes V, IX and XI. Thus, marker sets were defined at these loci. Each locus or QTL was named based on the current nomenclature, using the chromosome number and the name of the panel (Table 2). As no SNP was found in high LD with the significant SNPs detected on chromosomes III and IV, no marker sets were defined for these regions. Genotyping data at each marker set identified the "Ref" or "Alt" haplotypes in more than 95% of the individuals in the panel, and few recombinant haplotypes were identified (Table S1).

The QTL *GpaVa_MRQ*, detected on the northern arm of chromosome V, was composed of four markers covering 1.3 Mb. Pairwise correlation values between each marker of the set and the peak marker (chr05_5423448) were higher than 0.97. Haplotype–phenotype association analysis showed that the locus accounted for a substantial 54% of the phenotypic variation for CP. This peak marker was linked to marker HC (not detected by GWAS, $r^2 = 0.55$). In the MRQ panel, we observed that all genotypes carrying the "Alt" haplotype at this locus also have marker HC (Fig. 3). The genotypes without the "Alt" haplotype and HC marker are mainly susceptible, but the ones having only HC marker are

Table 2 SNP marker sets defining the loci around each GWAS peak and variance explained by each locus for CP (Cyst Presence) and CN1 (Cyst Number for the genotypes having at least one cyst) variables. Markers in bold are the peak markers. Ref and Alt are the reference and alternate allele, respectively, at each marker

Locus name	Marker set	Ref	Alt	Pearson correlation to peak marker	Locus size (Mb)	Variance explained (%)	
		0	1			CP	CN1
GpaVa_MRQ	chr05_4636267	C	G	0.98	1.3	54	
	chr05_5423448	**G**	**C**	1			
	chr05_5433470	T	C	0.98			
	chr05_5995827	T	C	0.97			
GpaVb_MRQ	chr05_46729352	T	C	0.93	0.7	15	
	chr05_46793576	**T**	**C**	1			
	chr05_46811300	C	G	0.93			
	chr05_47508836	C	T	0.93			
GpaIX_MRQ	chr09_65292024	C	T	1	2.2	18	
	chr09_65310251	**G**	**A**	1			
	chr09_66676903	A	G	0.96			
	chr09_67244173	C	T	0.97			
	chr09_67484357	A	G	1			
GpaXI_MRQ	chr11_19611129	A	G	1	7.4		18
	chr11_23823590	C	A	0.96			
	chr11_23823612	A	G	0.94			
	chr11_26392764	**C**	**T**	1			
	chr11_27087899	G	A	0.95			

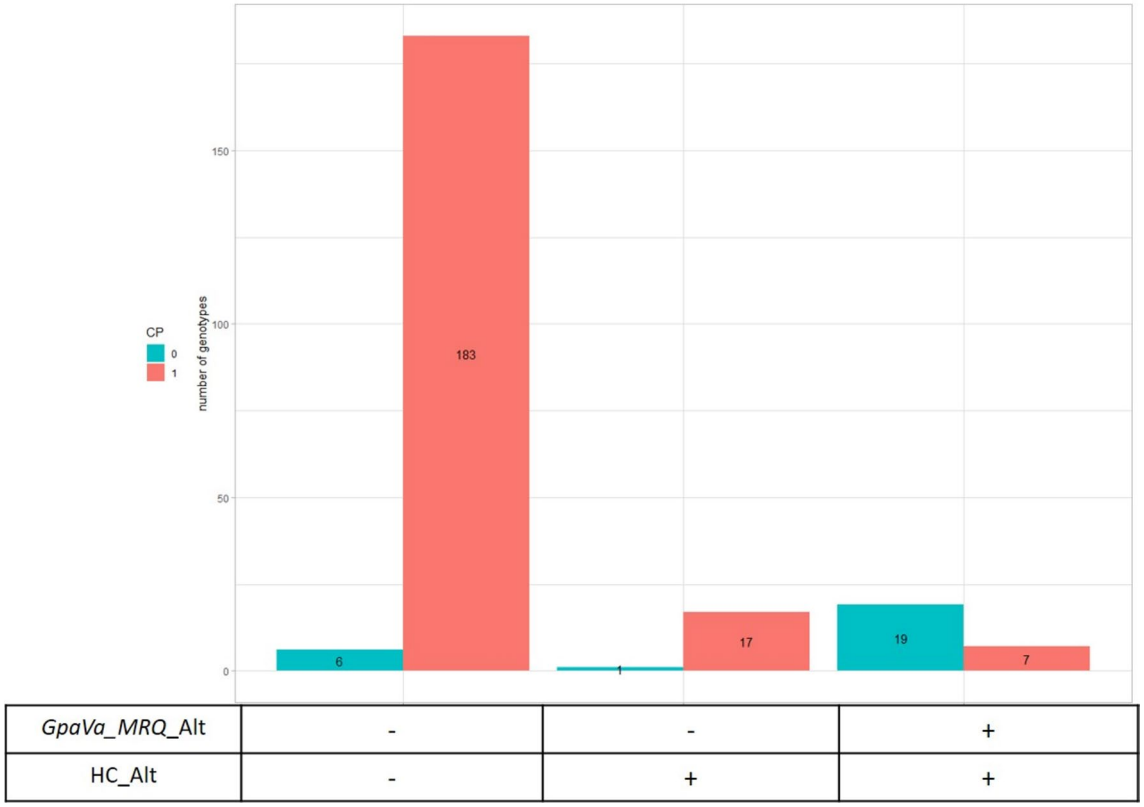


Fig. 3 Number of genotypes of the MRQ panel in two phenotypic classes for resistance to *G. pallida* and according to the presence of the reference (−) or alternate (+) haplotype at *GpaVa_MRQ* or at HC marker. Blue color represents the genotypes without any cyst (CP = 0), red color represents the genotypes with at least one cyst (CP = 1)

also mainly susceptible suggesting that HC is not specific to the resistant haplotype.

The QTL *GpaVb_MRQ,* detected on the southern arm of chromosome V, was composed of four markers covering 0.7 Mb. The pairwise correlation value between each marker of the set and the peak marker (chr05_46793576) is 0.93. This locus explained 15% of the phenotypic variation for CP. The "Alt" haplotype at this locus was found in about 6.5% of the genotypes of the panel.

The QTL *GpaIX_MRQ*, detected on chromosome IX, was composed of five markers covering 2.2 Mb. In this set, two markers were 100% correlated with the marker peak (chr09_65310251) and two other markers were correlated at 96% and 97% with the marker peak, respectively. The locus explained 18% of the phenotypic variation for CP. The "Alt" haplotype at this locus was present in 6% of the genotypes of the panel.

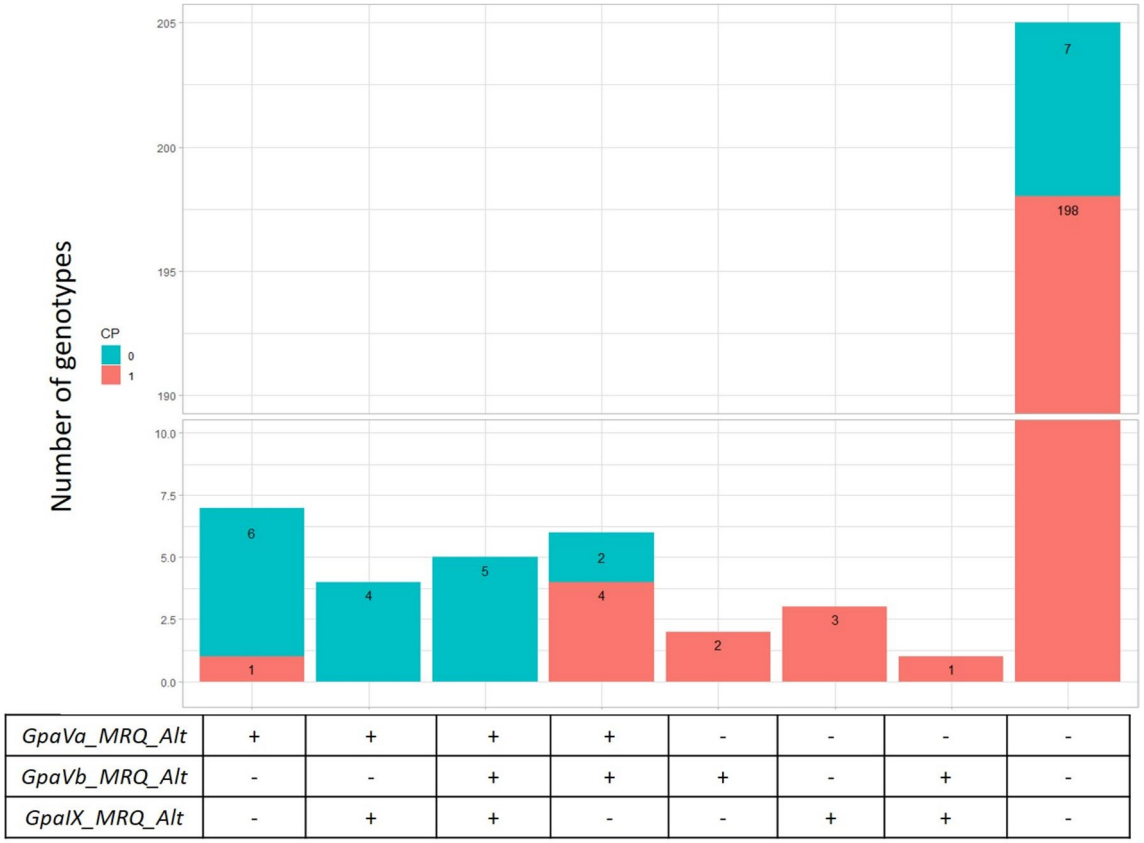
The effects of both QTL *GpaVb_MRQ* and *GpaIX_MRQ* were difficult to evaluate because of the high effect of QTL *GpaVa_MRQ*. Some genotypes carrying the alternate haplotypes at both *GpaVa_MRQ* and *GpaVb_MRQ* were susceptible, suggesting that the presence of the alternate haplotype at *GpaVb_MRQ* attenuated the effect of *GpaVa_MRQ* (Fig. 4). All the genotypes which cumulated alternate haplotypes at least at *GpaVa_MRQ* and *GpaIX_MRQ* presented a high level of resistance (CP = 0). However, the number of genotypes in the different classes was low.

The QTL *GpaXI_MRQ,* detected on chromosome XI, was composed of five markers covering 7.4 Mb and presented pairwise correlation values higher than 0.94. The locus explained 18% of the phenotypic variation for CN1. The "Alt" haplotype at this locus was associated with higher susceptibility. The alternate haplotype was present in 78% of the genotypes presenting at least one cyst.

Taking into account the knowledge we have on the validation panel, even if the SPUD1636 marker (associated with *GpaV*) was not detected in the GWAS analysis, SNP markers in LD with this marker were looked for. The *GpaV_SPUD1636* locus was composed of four markers covering approximately 1 Mb. Pairwise correlation values between each marker of the set and SPUD1636 were higher than 94%.

Development and validation of PACE markers

A total of 13 markers was tested with the PACE technology, including four markers at *GpaVa_MRQ*, five markers



GpaVa_MRQ_Alt	+	+	+	+	-	-	-	-
GpaVb_MRQ_Alt	-	-	+	+	+	-	+	-
GpaIX_MRQ_Alt	-	+	+	-	-	+	+	-

Fig. 4 Number of genotypes of the MRQ panel in two phenotypic classes for resistance to *G. pallida* and according to the presence of the reference (−) or alternate (+) haplotypes at *GpaVa_MRQ*, *GpaVb_MRQ* and *GpaIX_MRQ*. Blue color represents the genotypes without any cyst (CP = 0), red color represents the genotypes with at least one cyst (CP = 1)

at *GpaIX_MRQ* and four markers at *GpaV_SPUD1636*. Eleven of the 13 markers were successfully converted (Table 3). Marker chr09_67484357 (locus *GpaIX_MRQ*) failed to amplify, and marker chr05_4617036 (locus *GpaV_SPUD1636*) did not show any clustering. The other markers demonstrated a rate of concordant calls between GBS and PACE technologies higher than 95% (Table 3). The Kappa concordance test supported these findings, with scores above 0.8 confirming an almost perfect agreement between both technologies.

In the Validation panel, the 11 markers showed a high level of assignment rate, the marker chr05_5433470 (locus *GpaVa_MRQ*) having the lowest rate with 94% of assigned genotypes. In this panel, as in the MRQ panel, few genotypes presented recombinant profiles within the marker sets that were defined. The percentage of recombinant genotypes at *GpaVa_MRQ*, *GpaIX_MRQ* and *GpaV_SPUD1636* were 3.1%, 0.8% and 1.5%, respectively. The alternate haplotype at *GpaVa_MRQ* was present in more than half of the genotypes of the Validation panel, and the alternate haplotypes at *GpaIX_MRQ* and at *GpaV_SPUD1636* were present in a quarter of genotypes.

The level of resistance of the genotypes in the Validation panel depended mainly on the presence of the "Alt" haplotype at the major-effect QTL *GpaVa_MRQ* / *GpaV_SPUD1636* (Fig. 5). However, some genotypes carrying this haplotype at the single *GpaVa_MRQ* or *GpaV_SPUD1636* locus were susceptible to *G. pallida*. The four genotypes carrying the "Alt" haplotype only at *GpaIX_MRQ* exhibited higher levels of resistance than the genotypes carrying only the "Ref" haplotypes at all three loci. All the genotypes which combined the alternate haplotypes at *GpaIX_MRQ* and *GpaVa_MRQ* were resistant or very resistant.

Discussion

In this study, we performed a GWAS on a panel of pre-breeding genotypes of potato, in order to finely identify genetic determinants of resistance to *G. pallida* accumulated in these pre-breeding genotypes and to develop reliable tools and strategies for choosing the genetic loci to be integrated into future varieties by marker-assisted selection.

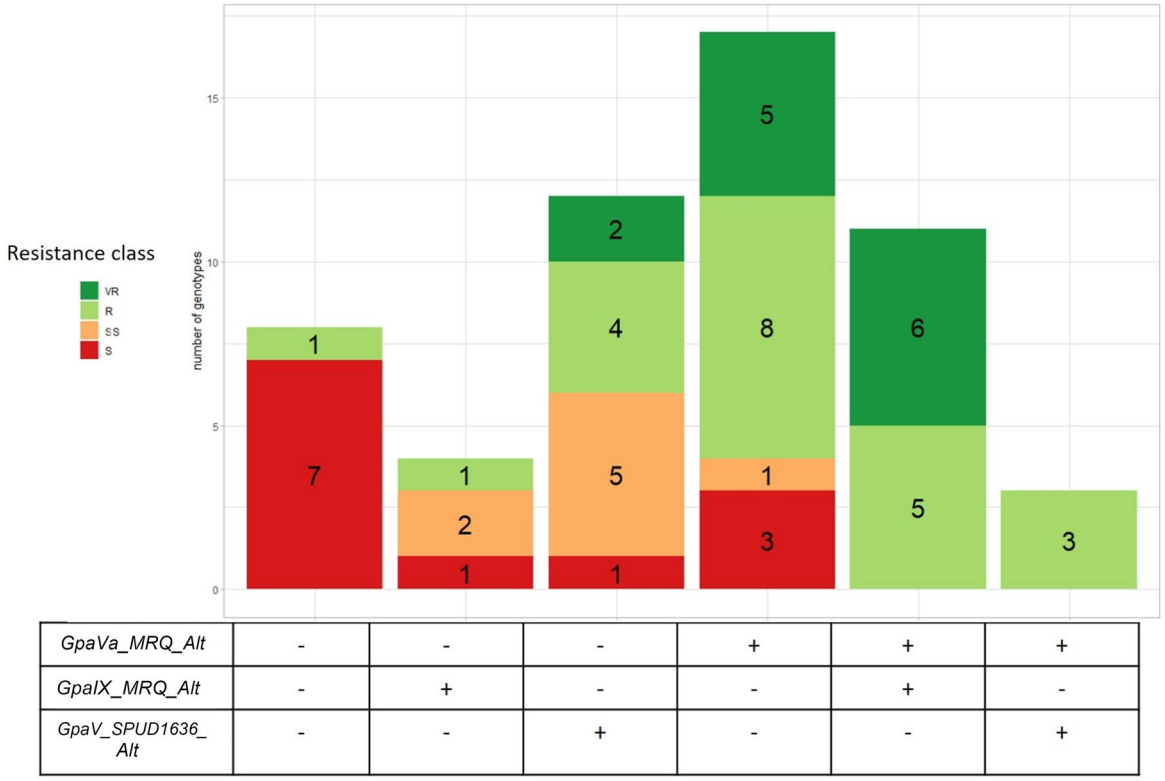
Table 3 **a** Comparison of GBS and PACE genotyping data in the MRQ subset and PACE genotyping data in the Validation panel. Assignment rate: Rate of genotypes of the Validation panel that could be assigned to one genotypic class (AAAA, AAAB, AABB, ABBB, BBBB) **b** PACE marker sequences, A1 corresponds to forward-FAM sequence, A2 to forward-HEX sequence and CP to reverse sequence

a

Locus name	Marker	MRQ panel	Validation panel		
		Concordance calls between GBS and PACE (%)	Assignment rate	% of clones with Alt haplotype	% of clones with only Ref haplotype
GpaVa_MRQ	chr05_4636267	97	1.00	54.7	42.2
	chr05_5423448	96	0.98		
	chr05_5433470	97	0.94		
	chr05_5995827	98	0.97		
GpaIX_MRQ	chr09_65292024	96	0.97	25.0	74.2
	chr09_65310251	95	1.00		
	chr09_66676903	95	1.00		
	chr09_67244173	96	0.98		
GpaV_SPUD1636	chr05_4931756	97	1.00	26.6	71.9
	chr05_5630510	99	1.00		
	chr05_5995936	98	0.98		

b

Locus name	Marker	A1	A2	CP
GpaVa_MRQ	chr05_4636267	GAAGGTGACCAAGTTCAT GCTAGCTGTGTATGCTTA ACTTRGCTATG	GAAGGTCGGAGTCAACGG ATTAGCTGTGTATGCTTA ACTTRGCTATC	TCTGCAAAAATATATCTGGTT TGGTWTC
	chr05_5423448	GAAGGTGACCAAGTTCAT GCTATGAGCTCCCTAACA CGGCTTG	GAAGGTCGGAGTCAACGG ATTATGAGCTCCCTAACA CGGCTTC	GATGGAGAACTTCCGCAA GAATCCAT
	chr05_5433470	GAAGGTGACCAAGTTCAT GCTATAAACCCAGACTAG CTCCTGAGT	GAAGGTCGGAGTCAACGG ATTAAACCCAGACTAGCT CCTGAGC	ACTGAAACTTTTGTCAAC CAATKCTGCTAT
	chr05_5995827	GAAGGTGACCAAGTTCAT GCTTTKTCAAGTCCTTGT TTGGTTAT	GAAGGTCGGAGTCAACGG ATTTKTCAAGTCCTTGT TTGGTTAC	AAGCACRTATCCGTGGCC AGATTATTA
GpaIX_MRQ	chr09_65292024	GAAGGTGACCAAGTTCAT GCTTCTGGATCATCATGC TTGAAAGTC	GAAGGTCGGAGTCAACGG ATTCTTCTGGATCATCAT GCTTGAAAGTT	CGTTAAGGTTGAGTCTTTGCT GAAATCAA
	chr09_65310251	GAAGGTGACCAAGTTCAT GCTGACATGGCASTATCCA CCC	GAAGGTCGGAGTCAACGG ATTGACATGGCASTATCCA CCT	TAAACAGCAAGGTACCRA GGAGT
	chr09_66676903	GAAGGTGACCAAGTTCAT GCTCTCGACCACAGAGAC CATTCT	GAAGGTCGGAGTCAACGG ATTCTCGACCACAGAGAC CATTCC	AGTTCATCCAATAACAGC AGTCGAGTT
	chr09_67244173	GAAGGTGACCAAGTTCAT GCTACAAAGAGCAAAAAG GCGATGAAATTC	GAAGGTCGGAGTCAACGG ATTCAACAAAGAGCAAAA AGGCGATGAAATTT	ATATGGCTACAAGGCACA GTRTCACTA
GpaV_SPUD1636	chr05_4931756	GAAGGTGACCAAGTTCAT GCTTCCTCCACTCTAAAA GATTAYAGCG	GAAGGTCGGAGTCAACGG ATTGTTTCCTCCACTCTA AAAGATTAYAGCA	CTTCAGATCTGGAATACY AACAGATCATA
	chr05_5630510	GAAGGTGACCAAGTTCAT GCTGCCGGGGTTGAAGTG GCTAGA	GAAGGTCGGAGTCAACGG ATTCCGGGGTTGAAGTGG CTAGG	CTGATCTGGTGGTATCKGAT ATGTGAT
	chr05_5995936	GAAGGTGACCAAGTTCAT GCTGCAACAGCCTGCCTT CTGACC	GAAGGTCGGAGTCAACGG ATTAGCAACAGCCTGCCT TCTGACA	CTACAAGCTGCAGTACGT GGYCAT



GpaVa_MRQ_Alt	-	-	-	+	+	+
GpaIX_MRQ_Alt	-	+	-	-	+	-
GpaV_SPUD1636_Alt	-	-	+	-	-	+

Fig. 5 Number of genotypes of the Validation panel in different phenotypic classes for resistance to *G. pallida, i.e.,* VR very resistant, R Resistant, SS Slightly Susceptible and S Susceptible, according to the presence of reference (−) or alternate (+) haplotypes at *GpaVa_MRQ*, *GpaIX_MRQ* and *GpaV_SPUD1636*

Better knowledge of already known QTL and identification of new genomic regions

We identified QTL on chromosomes III, IV, V, IX and XI, which were significantly associated with resistance to *G. pallida*, using a multi locus mixed model. The regions identified on chromosome IV, north of chromosome V and chromosome IX are known to harbor resistance to *G. pallida* inherited from *S. vernei*. The regions identified on chromosomes III and south of chromosome V are known to harbor resistance to *G. rostochiensis*. The chromosome XI is known to host some resistance QTL to *G. pallida* (Gartner et al. 2021).

The most significant region, *GpaVa_MRQ*, is located on the north of chromosome V and explains 54% of the phenotypic variance. Its physical position and the pedigree of the genotypes of the panel carrying resistance at this QTL suggest that this QTL corresponds to the well-known $GpaV_{vrn}$ locus (initially named *Gpa5*) inherited from *S. vernei* (Rouppe van der Voort et al. 2000; Gartner et al. 2021). Indeed, the clone AM 78-3778 is frequent in the pedigree of the genotypes carrying *GpaVa_MRQ* and has three sources of *S. vernei* in its pedigree. In previous reports, this QTL was identified from linkage mapping analysis (Rouppe van der Voort et al. 2000), with the HC marker as the most closely linked marker (Sattarzadeh et al. 2006). In our study, we identified closely linked SNP markers to this QTL and found a more reliable marker set associated with $GpaV_{vrn}$. Indeed, the HC marker can be challenging in certain genetic backgrounds where the favorable allele was amplified in susceptible genotypes (Milczarek et al. 2011; Asano et al. 2012, 2021). Our study also revealed that the HC marker was not always correlated with a high level of resistance, indicating that the favorable allele at this marker is not specific to the resistance allele.

The *GpaIX_MRQ* QTL identified on chromosome IX explained 18% of the phenotypic variance and is located in the same region than the *Gpa6* QTL described by Rouppe van der Voort et al. (2000). In our study, the level of variation explained by this region was similar to the one detected in Rouppe van der Voort et al. (2000) and clones carrying the favorable allele at this QTL also originated from *S. vernei*, suggesting that we detected *Gpa6*.

We also identified a QTL on the north of chromosome IV, a region known to contain the *GpaIV* and *H3* QTL (Bryan et al. 2002, 2004). In the MRQ panel, the genotypes carrying the resistance-associated allele at the significant marker detected on chromosome IV are derived from *S. vernei* or *S.*

andigena, suggesting that these regions are likely the same ones.

On chromosome III, the only gene/QTL which is known to be involved in resistance to PCN is *Gro1.4* which is a resistance gene to *G. rostochiensis* (Kreike et al. 1996). Until now, no resistance to *G. pallida* has been identified on this chromosome. The region we identified is located near *Gro1.4*. However, further studies are needed to determine if the resistance conferred by this QTL is effective against both PCN species.

The QTL we identified on the south of chromosome V explains 15% of the phenotypic variance. The presence of that QTL seems to mitigate the effect of *GpaVa*. This region is known to harbor the *H1* gene, which confers resistance to *G. rostochiensis* but no previous studies have shown the involvement of this region in resistance to *G. pallida*. The presence of the alternate haplotype is not linked to the 57R marker (Finkers-Tomczak et al. 2011) (results not shown), which suggests that this genomic region is independent.

The *GpaXI_MRQ* QTL, explaining 18% of the variation for CN1, is near the centromeric region. Two QTL have already been identified on chromosome XI: $GpaXI_{tar}$ located on the north arm (Adillah Tan et al. 2009) and $GpaXI_{spl}$ located on the south arm (Caromel et al. 2005). Our QTL is approximately at 20 Mb from these two QTL, and the genotypes in the MRQ panel carrying the resistance-associated haplotype at the QTL do not originate from *S. tarijense* or *S. sparsipilum*. Notably, no specific genealogical pattern was observed in association with this QTL, suggesting that this QTL could originate from *S. tuberosum*. The presence of the alternate haplotype at the QTL increased the susceptibility to *G. pallida*. Genes associated with more susceptibility have also been observed for late blight disease (Sun et al. 2016).

Although *GpaVa_MRQ* and *GpaIX_MRQ* are key quantitative resistance loci, they have not been detected with CN1 variable because there was a low number of genotypes carrying these resistance alleles among the genotypes considered for this variable (5 for *GpaVa_MRQ* and 2 for *GpaIX_MRQ*). As a result, significant regions detected with CP variable were not identified with CN1 variable.

Other minor-effect genetic resistance loci may exist in our panel but may have remained undetected in our study. The heritabilities of the traits measured were moderate (for the CN1 variable, $H^2 = 0.48$) to high (for the CP variable, $H^2 = 0.70$), and the MAF and Bonferroni thresholds applied for marker filtering and GWAS detection were quite stringent. In our study, we chose to focus on reliable QTL for application in marker-assisted selection. However, we identified some resistant genotypes carrying no resistance alleles at the detected QTL. Some of them, like INRAE 02 T.127. 5 or INRAE 02P.112. 67, originated from *S. sparsipilum* or *S. spegazzinii* (Caromel et al. 2005). These species were poorly represented in the MRQ panel. The detection of rare alleles is a limitation of GWAS approaches, which highlights the interest of complementary genetic studies with other panels created from crosses involving uncommon sources of resistance.

Marker haplotypes at a relevant *G. pallida* resistance QTL combination, useful for marker-assisted selection

Based on the GWAS results, our study identified favorable haplotypes at QTL combinations associated with increased levels of resistance and developed robust marker sets.

We showed that the combination of alternate haplotypes at the *GpaVa_MRQ* and *GpaIX_MRQ* QTL was associated to a total resistance. Clones carrying both haplotypes exhibited the highest resistance levels in the MRQ and Validation panels. The complementarity of effects can alter the resistance mechanism, particularly for *G. pallida*, where the combination of $GpaV_{spl}$ and $GpaXI_{spl}$ shifts resistance from masculinization to blocking (Caromel et al. 2005). Utilizing *GpaIX* may help sustain the efficacy of the $GpaV_{vrn}$ QTL. To further validate these results, it would be valuable to determine if genotypes combining *GpaVa_MRQ* and *GpaIX_MRQ* QTL remain resistant to $GpaV_{vrn}$ virulent populations.

Marker sets relevant for marker-assisted selection

In this study, we identified and developed robust marker sets to target favorable haplotypes at the QTL detected for marker-assisted selection. The use of a set of markers in high LD to target a QTL region in breeding is relevant in order to efficiently select the favorable allele at the QTL, while limiting the risk of loss of marker-QTL association by recombination. In our study, we defined marker sets in using a fairly strict approach applying a rule of at least four markers in the set, each being in LD with the GWAS marker peak with a value higher than or equal to 0.9. The size of LD-blocks observed in our panel was consistent with the linkage disequilibrium observed in other potato studies (Vos et al. 2017).

The approach applied in this study to define marker sets proved advantageous, as it enabled the development of PACE markers to tag the relevant combination of resistance QTL *GpaVa_MRQ* and *GpaIX_MRQ*. PACE marker sets, each containing at least three markers, were designed to facilitate the selection of specific haplotypes by breeders and reduce the likelihood of recombination events that could disrupt the marker-causal gene linkage. The transferability of these PACE marker sets associated with QTL *GpaV_MRQ* and *GpaIX_MRQ* was validated in a second panel. This validation showed minimal recombination, thereby increasing confidence in these haplotypes as reliable tools for tagging

the resistance QTL. This approach also allowed the development of the *GpaV_SPUD1636* marker set, which, despite having similar genetic limitations as the current marker, offers the advantage of new technology makers. These PACE marker sets will be useful and quickly applicable in breeding programs.

The potential to integrate these two QTL in breeding programs will allow their stacking with other QTL from other species, such as *S. sparsipilum* (Caromel et al. 2005), *S. andigena* (Bradshaw et al. 1998) or *S. spegazzinii* (Gartner et al. 2024). It would be beneficial to pyramid resistance that we detected with other sources to improve resistance levels and durability in future varieties (Price et al. 2023).

The PACE molecular markers developed in this study promise substantial improvements in potato resistance to PCNs. Future research should evaluate these markers in broader genetic pools and diverse environments to confirm their reliability, ensuring their utility in various breeding contexts. This work lays a foundation for developing potato varieties with enhanced resilience to PCNs, contributing to global food security and sustainable agriculture.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00122-024-04794-8.

Acknowledgements We acknowledge the BrACySol BRC (INRAE Ploudaniel, France) for providing the plants that were used in this study.

Author contribution statement J.L. contributed to conceptualization, methodology, analysis, visualization, and writing; F. E. contributed to conceptualization, methodology, supervising, and reviewing; P.L. L. contributed to PACE genotyping, analysis, and writing; S. F. contributed to phenotyping and reviewing; M.P. C. contributed to plant material preparation; S. M. contributed to methodology and reviewing; C. P. contributed to methodology and reviewing; M.L. P.N. contributed to conceptualization, methodology, supervising, and reviewing; M.C. K. contributed to conceptualization, methodology, supervising, reviewing, project administration, and funding acquisition.

Funding The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was funded by ACVNPT (Association des Créateurs de Variétés Nouvelles de Pomme de terre), and CASDAR project GECONEM (2020 -C2018-07).

Data availability Genomic data used in this study are presented in Leuenberger et al (2024), phenotypic data are in Tables S1 and S2. Plant material are maintained in BrACySol BRC (INRAE Ploudaniel, France).

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

Adillah Tan MY, Park T-H, Alles R et al (2009) GpaXItarloriginating from *Solanum tarijense* is a major resistance locus to *Globodera pallida* and is localised on chromosome 11 of potato. Theor Appl Genet 119:1477–1487. https://doi.org/10.1007/s00122-009-1149-4

Alvarez-Morezuelas A, Barandalla L, Ritter E, Ruiz de Galarreta JI (2023) Genome-wide association study of agronomic and physiological traits related to drought tolerance in potato. Plants 12:734. https://doi.org/10.3390/plants12040734

Asano K, Kobayashi A, Tsuda S et al (2012) DNA marker-assisted evaluation of potato genotypes for potential resistance to potato cyst nematode pathotypes not yet invading into Japan. Breed Sci 62:142–150. https://doi.org/10.1270/jsbbs.62.142

Asano K, Shimosaka E, Yamashita Y et al (2021) Improvement of diagnostic markers for resistance to *Globodera pallida* and application for selection of resistant germplasms in potato breeding. Breed Sci 71:354–364. https://doi.org/10.1270/jsbbs.20134

Bastien M, Boudhrioua C, Fortin G, Belzile F (2018) Exploring the potential and limitations of genotyping-by-sequencing for SNP discovery and genotyping in tetraploid potato. Genome 61:449–456. https://doi.org/10.1139/gen-2017-0236

Bhat JA, Yu D, Bohra A et al (2021) Features and applications of haplotypes in crop breeding. Commun Biol 4:1–12. https://doi.org/10.1038/s42003-021-02782-y

Bonnafous F, Fievet G, Blanchet N et al (2018) Comparison of GWAS models to identify non-additive genetic control of flowering time in sunflower hybrids. Theor Appl Genet 131:319–332. https://doi.org/10.1007/s00122-017-3003-4

Bradshaw JE, Hackett CA, Meyer RC et al (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. https://doi.org/10.1007/s001220050886

Bryan J, McLean K, Bradshaw E et al (2002) Mapping QTLs for resistance to the cyst nematode *Globodera pallida* derived from the wild potato species *Solanum vernei*. TAG Theor Appl Genet Theor Angew Genet 105:68–77. https://doi.org/10.1007/s00122-002-0873-9

Bryan GJ, McLean K, Pande B et al (2004) Genetical dissection of H3-mediated polygenic PCN resistance in a heterozygous autotetraploid potato population. Mol Breed 14:105–116. https://doi.org/10.1023/B:MOLB.0000037999.13581.9c

Caromel B, Mugniéry D, Lefebvre V et al (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. https://doi.org/10.1007/s00122-003-1211-6

Caromel B, Mugniéry D, Kerlan M-C et al (2005) Resistance quantitative trait loci originating from *Solanum sparsipilum* act independently on the sex ratio of *Globodera pallida* and together for developing a necrotic reaction. Mol Plant-Microbe Interactions® 18:1186–1194. https://doi.org/10.1094/MPMI-18-1186

Dalton E, Griffin D, Gallagher TF et al (2013) The effect of pyramiding two potato cyst nematode resistance loci to *Globodera pallida* Pa2/3 in potato. Mol Breed 31:921–930. https://doi.org/10.1007/s11032-013-9845-9

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Ernst K, Kumar A, Kriseleit D et al (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 J 31:127–136. https://doi.org/10.1046/j.1365-313X.2002.01341.x

Felcher KJ, Coombs JJ, Massa AN et al (2012) Integration of two diploid potato linkage maps with the potato genome sequence. PLoS ONE 7:e36347. https://doi.org/10.1371/journal.pone.0036347

Finkers-Tomczak A, Bakker E, de Boer J et al (2011) Comparative sequence analysis of the potato cyst nematode resistance locus H1 reveals a major lack of co-linearity between three haplotypes in potato (*Solanum tuberosum* ssp.). Theor Appl Genet 122:595–608. https://doi.org/10.1007/s00122-010-1472-9

Foot MA (1977) Laboratory rearing of potato cyst nematode; a method suitable for pathotyping and biological studies. N Z J Zool 4:183–186. https://doi.org/10.1080/03014223.1977.9517951

Fournet S, Kerlan MC, Renault L et al (2013) Selection of nematodes by resistant plants has implications for local adaptation and cross-virulence. Plant Pathol 62:184–193. https://doi.org/10.1111/j.1365-3059.2012.02617.x

Gartner U, Hein I, Brown LH et al (2021) Resisting potato cyst nematodes with resistance. Front Plant Sci. https://doi.org/10.3389/fpls.2021.661194

Gartner U, Armstrong MR, Sharma SK et al (2024) Characterisation and mapping of a *Globodera pallida* resistance derived from the wild potato species *Solanum spegazzinii*. TAG Theor Appl Genet Theor Angew Genet 137:106. https://doi.org/10.1007/s00122-024-04605-0

Hasan N, Choudhary S, Naaz N et al (2021) Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes. J Genet Eng Biotechnol 19:128. https://doi.org/10.1186/s43141-021-00231-1

Kassambara A, Mundt F (2020) factoextra: Extract and visualize the results of multivariate data analyses

Khlestkin VK, Erst T, Rozanova I et al (2020) Genetic loci determining potato starch yield and granule morphology revealed by genome-wide association study (GWAS). PeerJ 8:e10286. https://doi.org/10.7717/peerj.10286

Klaassen MT, Willemsen JH, Vos PG et al (2019) Genome-wide association analysis in tetraploid potato reveals four QTLs for protein content. Mol Breed 39:151. https://doi.org/10.1007/s11032-019-1070-8

Kreike CM, Kok-Westeneng AA, Vinke JH, Stiekema WJ (1996) Mapping of QTLs involved in nematode resistance, tuber yield and root development in Solanum sp. Theor Appl Genet 92:463–470. https://doi.org/10.1007/BF00223694

Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159–174. https://doi.org/10.2307/2529310

Lê S, Josse J, Husson F (2008) FactoMineR: An R package for multivariate analysis. J Stat Softw 25:1–18. https://doi.org/10.18637/jss.v025.i01

Leuenberger J, Sharma SK, McLean K et al (2024) A genomic dataset integrating genotyping-by-sequencing, SolCAP array and PCR marker data on tetraploid potato advanced breeding lines. Front Plant Sci. https://doi.org/10.3389/fpls.2024.1384401

Marhadour S, Prodhomme C (2023) Chapter 15 - Recent trends in genetics studies and molecular breeding of potato. In: Çalişkan ME, Bakhsh A, Jabran K (eds) Potato production worldwide. Academic Press, pp 273–301

Milczarek D, Flis B, Przetakiewicz A (2011) Suitability of molecular markers for selection of potatoes resistant to Globodera spp. Am J Potato Res 88:245–255. https://doi.org/10.1007/s12230-011-9189-0

Muñoz F, Sanchez L (2019) breedR: Statistical Methods for Forest Genetic Resources Analysts

Mwangi J, Niere B, Finckh M et al (2019) Reproduction and life history traits of a resistance breaking *Globodera pallida* population. J Nematol 51:1–13. https://doi.org/10.21307/jofnem-2019-079

Paal J, Henselewski H, Muth J et al (2004) Molecular cloning of the potato Gro1-4 gene conferring resistance to pathotype Ro1 of the root cyst nematode *Globodera rostochiensis*, based on a candidate gene approach. Plant J 38:285–297. https://doi.org/10.1111/j.1365-313X.2004.02047.x

Palomares-Rius JE, Hasegawa K, Siddique S, Vicente CSL (2021) Editorial: protecting our crops—approaches for plant parasitic nematode control. Front Plant Sci. https://doi.org/10.3389/fpls.2021.726057

Pandey J, Scheuring DC, Koym JW, Vales MI (2022) Genomic regions associated with tuber traits in tetraploid potatoes and identification of superior clones for breeding purposes. Front Plant Sci 13:952263. https://doi.org/10.3389/fpls.2022.952263

Pham GM, Hamilton JP, Wood JC et al (2020) Construction of a chromosome-scale long-read reference genome assembly for potato. GigaScience 9:giaa100. https://doi.org/10.1093/gigascience/giaa100

Pilet-Nayel M-L, Moury B, Caffier V et al (2017) Quantitative resistance to plant pathogens in pyramiding strategies for durable crop protection. Front Plant Sci 8:1838. https://doi.org/10.3389/fpls.2017.01838

Price JA, Coyne D, Blok VC, Jones JT (2021) Potato cyst nematodes *Globodera rostochiensis* and G. pallida. Mol Plant Pathol 22:495–507. https://doi.org/10.1111/mpp.13047

Price J, Preedy K, Young V et al (2023) Stacking host resistance genes to control *Globodera pallida* populations with different virulence. Eur J Plant Pathol. https://doi.org/10.1007/s10658-023-02761-5

Pulavarty A, Singh A, Smyth D et al (2022) Sustainable management of the potato cyst nematode, *Globodera rostochiensis*, with two microbial fermentation products. Front Plant Sci. https://doi.org/10.3389/fpls.2022.987059

Rasheed A, Hao Y, Xia X et al (2017) Crop breeding chips and genotyping platforms: progress, challenges, and perspectives. Mol Plant 10:1047–1064. https://doi.org/10.1016/j.molp.2017.06.008

Rosyara UR, De Jong WS, Douches DS, Endelman JB (2016) Software for genome-wide association studies in autopolyploids and its application to potato. Plant Genome. https://doi.org/10.3835/plantgenome2015.08.0073

Rouppe van der Voort J, van der Vossen E, Bakker E et al (2000) Two additive QTLs conferring broad-spectrum resistance in potato to *Globodera pallida* are localized on resistance gene clusters. Theor Appl Genet 101:1122–1130. https://doi.org/10.1007/s001220051588

Sattarzadeh A, Achenbach U, Lübeck J et al (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. https://doi.org/10.1007/s11032-006-9026-1

Segura V, Vilhjálmsson BJ, Platt A et al (2012) An efficient multi-locus mixed-model approach for genome-wide association studies in

structured populations. Nat Genet 44:825–830. https://doi.org/10.1038/ng.2314

Semagn K, Babu R, Hearne S, Olsen M (2014) Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. Mol Breed 33:1–14. https://doi.org/10.1007/s11032-013-9917-x

Sun K, Wolters A-MA, Vossen JH et al (2016) Silencing of six susceptibility genes results in potato late blight resistance. Transgenic Res 25:731–742. https://doi.org/10.1007/s11248-016-9964-2

The Potato Genome Sequencing Consortium (2011) Genome sequence and analysis of the tuber crop potato. Nature 475:189–195. https://doi.org/10.1038/nature10158

Tibbs Cortes L, Zhang Z, Yu J (2021) Status and prospects of genome-wide association studies in plants. Plant Genome 14:e20077. https://doi.org/10.1002/tpg2.20077

Van der Plank JE (1968) Disease resistance in plants. Academic Press, New York, p 206

van Eck HJ, Vos PG, Valkonen JPT et al (2017) Graphical genotyping as a method to map Ny(o, n)stoand Gpa5 using a reference panel of tetraploid potato cultivars. Theor Appl Genet 130:515–528. https://doi.org/10.1007/s00122-016-2831-y

VanRaden PM (2008) Efficient methods to compute genomic predictions. J Dairy Sci 91:4414–4423. https://doi.org/10.3168/jds.2007-0980

Vexler L, Leyva-Perez MD, Konkolewska A, Clot CR, Byrne S, Griffin D, Ruttink T, Hutten RC, Engelen C, Visser RG, Prigge V (2024) QTL discovery for agronomic and quality traits in diploid potato clones using PotatoMASH amplicon sequencing. G3 Genes Genomes Genet 14(10):164. https://doi.org/10.1093/g3journal/jkae164

von Maydell D (2023) PCR Allele competitive extension (PACE). In: Shavrukov Y (ed) Plant genotyping: methods and protocols. Springer, New York, pp 263–271

Vos PG, Paulo MJ, Voorrips RE et al (2017) Evaluation of LD decay and various LD-decay estimators in simulated and SNP-array data of tetraploid potato. Theor Appl Genet 130:123–135. https://doi.org/10.1007/s00122-016-2798-8

Wainer J, Dinh Q (2021) Taxonomy, morphological and molecular identification of the potato cyst nematodes *Globodera Pallida* and G. rostochiensis. Plants 10:184. https://doi.org/10.3390/plants10010184

Yu J, Pressoir G, Briggs WH et al (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208. https://doi.org/10.1038/ng1702