

University of Dundee

**RLP/K enrichment sequencing; a novel method to identify receptor-like protein (RLP) and receptor-like kinase (RLK) genes**

Lin, Xiao; Armstrong, Miles; Baker, Katie; Wouters, Doret; Visser, Richard G. F.; Wolters, Pieter J.

*Published in:*  
New Phytologist

*DOI:*  
[10.1111/nph.16608](https://doi.org/10.1111/nph.16608)

*Publication date:*  
2020

*Licence:*  
CC BY

*Document Version*  
Publisher's PDF, also known as Version of record

[Link to publication in Discovery Research Portal](#)

*Citation for published version (APA):*

Lin, X., Armstrong, M., Baker, K., Wouters, D., Visser, R. G. F., Wolters, P. J., Hein, I., & Vleeshouwers, V. G. A. (2020). RLP/K enrichment sequencing; a novel method to identify receptor-like protein (RLP) and receptor-like kinase (RLK) genes. *New Phytologist*, 227(4), 1264-1276. <https://doi.org/10.1111/nph.16608>

**General rights**

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Methods

## RLP/K enrichment sequencing; a novel method to identify receptor-like protein (RLP) and receptor-like kinase (RLK) genes

Xiao Lin<sup>1</sup> , Miles Armstrong<sup>2</sup> , Katie Baker<sup>2</sup> , Doret Wouters<sup>1</sup>, Richard G. F. Visser<sup>1</sup> , Pieter J. Wolters<sup>1</sup> , Ingo Hein<sup>2,3\*</sup>  and Vivianne G. A. A. Vleeshouwers<sup>1\*</sup> 

<sup>1</sup>Plant Breeding, Wageningen University and Research, Droeendaalsesteeg 1, 6708 PB, Wageningen, the Netherlands; <sup>2</sup>Cell and Molecular Sciences, The James Hutton Institute, Dundee, DD2 5DA, UK; <sup>3</sup>Division of Plant Sciences, School of Life Sciences, University of Dundee at the James Hutton Institute, Dundee, DD2 5DA, UK

Authors for correspondence:

Ingo Hein

Tel: +44 1382 568869

Email: ingo.hein@hutton.ac.uk

Vivianne G. A. A. Vleeshouwers

Tel: +31 317 487460

Email: vivianne.vleeshouwers@wur.nl

Received: 26 February 2020

Accepted: 27 March 2020

New Phytologist (2020)

doi: 10.1111/nph.16608

**Key words:** genotyping by sequencing (GBS), pattern recognition receptor (PRR), *Phytophthora infestans*, potato, receptor-like kinase (RLK), receptor-like protein (RLP), RenSeq, RLP/K enrichment sequencing (RLP/KSeq).

### Summary

- The identification of immune receptors in crop plants is time-consuming but important for disease control. Previously, resistance gene enrichment sequencing (RenSeq) was developed to accelerate mapping of nucleotide-binding domain and leucine-rich repeat containing (NLR) genes. However, resistances mediated by pattern recognition receptors (PRRs) remain less utilized.
- Here, our pipeline shows accelerated mapping of PRRs. Effectoromics leads to precise identification of plants with target PRRs, and subsequent RLP/K enrichment sequencing (RLP/KSeq) leads to detection of informative single nucleotide polymorphisms that are linked to the trait.
- Using *Phytophthora infestans* as a model, we identified *Solanum microdontum* plants that recognize the apoplastic effectors INF1 or SCR74. RLP/KSeq in a segregating *Solanum* population confirmed the localization of the INF1 receptor on chromosome 12, and led to the rapid mapping of the response to SCR74 to chromosome 9. By using markers obtained from RLP/KSeq in conjunction with additional markers, we fine-mapped the SCR74 receptor to a 43-kbp *G-LecRK* locus.
- Our findings show that RLP/KSeq enables rapid mapping of PRRs and is especially beneficial for crop plants with large and complex genomes. This work will enable the elucidation and characterization of the nonNLR plant immune receptors and ultimately facilitate informed resistance breeding.

### Introduction

To protect themselves against pathogens, plants have evolved two layers of defence (Jones & Dangl, 2006). The first layer is formed by extracellular receptors on the plant cell surface that are often referred to as pattern recognition receptors (PRRs). These surface receptors typically represent receptor-like proteins (RLPs) and receptor-like kinases (RLKs), which can recognize apoplastic effectors, microbe-associated molecular patterns (MAMPs) from plant pathogens and danger-associated plant breakdown products (DAMPs). The second layer of defence is mounted upon recognition of cytoplasmic effectors by internal immune receptors that typically encode for resistance (*R*) genes of the nucleotide-binding domain and leucine-rich repeat (NLR) class. Stacking and pyramiding *R* genes and PRRs is believed to contribute to more durable plant disease resistance (Dangl *et al.*, 2013).

Potato is an important food crop. However, the global yield of potato is threatened by potato late blight, which is caused by the oomycete pathogen *Phytophthora infestans* that led to the great Irish famine in the mid-1840s (Haverkort *et al.*, 2008). Traditionally, breeding for late blight resistance in potato has relied on introducing *R* genes from wild *Solanum* species into potato cultivars (Vleeshouwers *et al.*, 2011a; Jo *et al.*, 2014). However, these NLRs are often quickly defeated by fast-evolving *P. infestans* isolates in the field (Wastie, 1991; Fry, 2008). Another, currently largely unexploited layer of immunity occurs at the surface of plant cells. This apoplastic immunity is believed to generally provide a broader spectrum of resistance and is based on RLP/RLK-mediated recognition of MAMPs or apoplastic effectors. Some MAMPs, like Nep1-like proteins (NLP), are conserved among different pathogen kingdoms (Gijzen & Nürnberger, 2006; Oome *et al.*, 2014). Other examples of well characterized MAMPs are flagellin and elicitors, from bacteria and oomycetes, respectively (Felix *et al.*, 1999; Derevnina *et al.*, 2016). INF1 is a

\*These authors contributed equally to this work.

well-studied elicitor from *P. infestans* that triggers defence responses upon recognition by ELR, an RLP from *Solanum microdontum* residing on chromosome 12 (Du *et al.*, 2015). Other types of apoplastic effectors are extremely diverse and include small cysteine-rich proteins such as SCR74 from *P. infestans* (Liu *et al.*, 2005). Cloning and characterizing plant surface immune receptors, including the receptor of SCR74, will further our understanding of plant immunity and help to engineer crops with more durable disease resistance.

Recent advances in sequencing technologies have facilitated whole-genome sequencing and enabled genotyping by sequencing (GBS). This development has led to the emergence of several novel approaches for map-based cloning, such as genomic resequencing (Zou *et al.*, 2012; Zhu *et al.*, 2017), bulked segregant RNA-seq (Ramirez-Gonzalez *et al.*, 2015), Indel-seq (Singh *et al.*, 2017), and QTL-seq (Takagi *et al.*, 2013). In addition, when targeting certain types of gene families (e.g. *NLRs*), target enrichment sequencing significantly reduces the complexity of the genome before sequencing (Hodges *et al.*, 2007; Jupe *et al.*, 2013). *R* gene enrichment sequencing (RenSeq) aided the reannotation and mapping of *NLR* genes in potato. All *NLR* genes from the potato reference genome DM1-3, v.4.03 (doubled monoploid *Solanum tuberosum* group phureja clone) were predicted and an RNA bait library was generated to represent these *NLRs* (Jupe *et al.*, 2013). This work led to the accelerated genetic mapping of late blight *R* genes *Rpi-ber2*, *Rpi-rzc1*, *Rpi-ver1* from *Solanum berthaultii*, *Solanum ruiz-ceballosii* and *Solanum verrucosum*, respectively (Jupe *et al.*, 2013; Chen *et al.*, 2018). When combined with single-molecule real-time (SMRT) PACBIO sequencing of larger DNA fragments, RenSeq generates a true sequence representation of full-length *NLR* genes which enabled the rapid cloning of *Rpi-amr3* from *Solanum americanum* (Witek *et al.*, 2016). RenSeq has also been successfully applied to other crops and has led to the cloning of two stem rust resistance genes, *Sr22* and *Sr45*, from hexaploid bread wheat (Steuernagel *et al.*, 2016). Furthermore, used as a diagnostic tool and referred to as dRenSeq, the methodology enables the identification of known functional *NLRs* in potatoes (Van Weymers *et al.*, 2016; Jiang *et al.*, 2018; Armstrong *et al.*, 2019). These successful advances in enrichment sequencing indicate that, with adaption and optimization, the sequence capture technology can be applied to other types of immune receptors, such as RLPs and RLKs. Consistent with other genome reduction technologies such as RenSeq, GenSeq and PenSeq (Jupe *et al.*, 2013; Strachan *et al.*, 2019; Thilliez *et al.*, 2019), we refer to this adaptation as RLP/KSeq.

In this study, we established a pipeline to accelerate the identification of surface receptors that perceive apoplastic effectors, by using the potato–*Phytophthora infestans* pathosystem as a proof of concept. We developed a pipeline (Fig. 1) that consists of two steps: effectoromics, that is screening wild *Solanum* species to identify plants that recognize the apoplastic effectors INF1 and SCR74; and RLP/KSeq, to accelerate the genetic mapping of the underlying immune receptors through bulked segregant analysis (BSA). Ultimately, we fine-mapped the SCR74 receptor to a 43-kbp *G-LecRK* locus.

## Materials and Methods

### Plant material

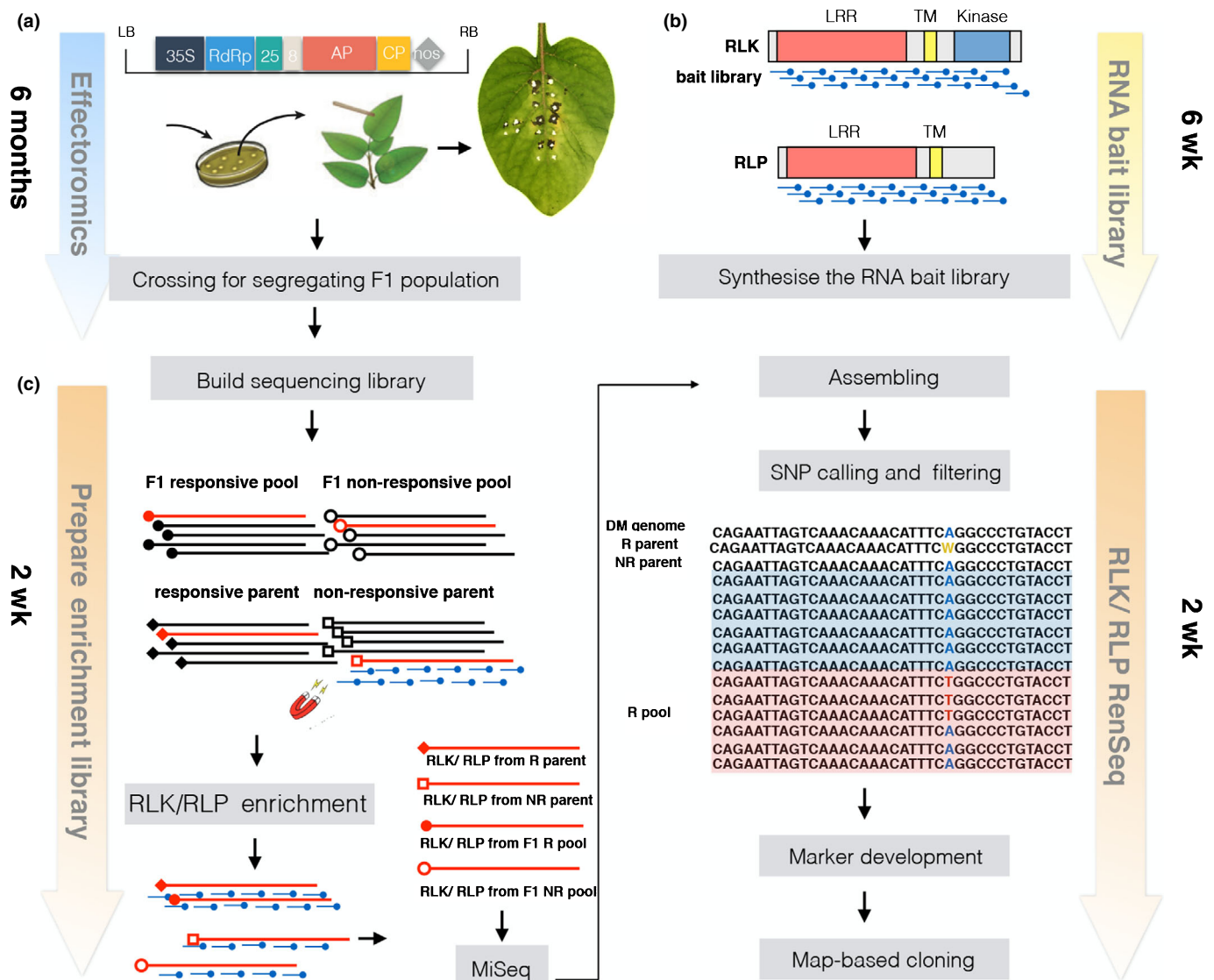
*Solanum* genotypes used in this study are listed in Fig. 2 and Supporting Information Table S1 (Vleeshouwers *et al.*, 2011b). These plants were maintained *in vitro* on MS20 medium at 25°C, as described by Du *et al.* (2014). Top shoots of plants were cut and clonally propagated *in vitro* 2 wk before transfer to soil in a climate-controlled glasshouse compartment with a 22°C : 18°C, day : night regime under long-day conditions. The F<sub>1</sub> population 7026 was generated by crossing *Solanum microdontum* ssp. *gigantophyllum* (GIG362-6) with *S. microdontum* (MCD360-1). The plants were grown in a crossing glasshouse until flowering. Flowers from GIG362-6 were emasculated before they were fully opened and pollinated with pollen that was collected from MCD360-1. After 4–5 wk, the ripe berries were removed from the plants. The seeds were collected and cleaned by water and dried on filter paper. Seeds were sown on MS20 medium or were soaked on filter paper after 3–4 months of dormancy. Gibberellic acid (GA3) was used for breaking dormancy if necessary.

### Cloning of effectors for PVX agroinfection

*Inf1* (XM\_002900382.1) and *Scr74-B3b* (AY723717.1) were cloned into the potato virus X (PVX) vector pGR106, and electrotransformed into *Agrobacterium tumefaciens* strain GV3101. Recombinant *A. tumefaciens* strain GV3101 carrying the effector constructs were grown for 2 d in LB medium at 28°C with kanamycin (50 µg ml<sup>-1</sup>) for selection. From the suspension, 1 ml of *Agrobacterium* culture was plated out onto LBA plates supplemented with kanamycin (50 µg ml<sup>-1</sup>) and incubated at 28°C for 2 d more. The *Agrobacterium* culture was collected from the Petri dishes with a plate spreader and used to inoculate 3- to 4-wk-old plants through toothpick inoculation (Takken *et al.*, 2000; Du *et al.*, 2014). Per leaf, two spots were inoculated for each construct and three leaves were used per plant. In total, three replicated plants were used for each genotype. Cell death responses were scored 2 wk post-infection on a range from 0 (no response) to 10 (strong response).

### Design of customized RLP/ RLK enrichment library

In total, 444 *RLP* genes and 533 *RLK* genes were predicted from the DM genome by HMMER, BLASTP, and INTERPRO, from both Potato Genome Sequencing Consortium (PGSC) and International Tomato Annotation Group annotations. Sequences of predicted *RLK* and *RLP* genes from DM are summarized in Notes S1. The respective homologues of the 444 and 533 *RLK* genes from the *Solanum chacoense* (M6) genome (Leisner *et al.*, 2018) are in Notes S2. Further, 18 known *RLP/RLK* genes from other *Solanaceae* species were included (Notes S3). All *RLP* and *RLK* genes were included in the enrichment bait-library design and represented by 120 bp fragments allowing for two times coverage (60 bp overlap). Duplicated oligonucleotides were removed. Unique RNA oligonucleotides were synthesized to



**Fig. 1** Overview of the effectormomics and receptor-like protein/kinase enrichment sequencing (RLP/KSeq) pipeline for the fast identification and mapping of surface immune receptors. (a) Predicted *Phytophthora infestans* apoplastic effectors are cloned into the binary potato virus X (PVX) vector pGR106 and transformed into *Agrobacterium tumefaciens* for functional screening by PVX agroinfection. Agroinfected leaves are scored at 10–14 d post-infection (dpi) for occurrence of cell death phenotypes. Responsive (R) and nonresponsive (NR) genotypes are crossed to create segregating F<sub>1</sub> populations. (b) Prediction of the RLP and RLK genes from the reference genome enables the design and synthesis of the RLP/RLK bait library for bespoke target enrichment sequencing in the selected plant species. (c) An F<sub>1</sub> population is screened for segregation of the recognition phenotype and pooled, based on their response pattern. Responsive and nonresponsive pools as well as the respective parents are subjected to enrichment sequencing. RLP/KSeq-derived reads are mapped to the reference genome, and single nucleotide polymorphisms (SNPs) linked to the recognition phenotype identified. Candidate markers are tested on the segregating population by SNP genotyping technologies such as LightScanner.

generate a customized MYbaits enrichment library (Arbor Biosciences Inc., Ann Arbor, MI, USA) comprising 57 020 probes (Notes S4).

#### Preparation of sequencing libraries and target capture

Genomic DNA was isolated from GIG362-6, MCD360-1 and the F<sub>1</sub> progenies using the DNeasy Plant Mini Kit (Qiagen). Equal amounts of DNA were pooled from 30 responsive as well as 30 nonresponsive progenies for the INF1 recognition phenotype and 29 responsive and 30 nonresponsive individuals for the

SCR74 response phenotype, respectively. DNA concentrations were measured using a Qubit fluorometer (ThermoFisher, Dubuque, IA, USA). The NEBNext Ultra<sup>TM</sup> II FS DNA Library Prep Kit (New England Biolabs, Ipswich, MA, USA) was used for fragmentation/adaptor ligation and indexing of samples. The bio-analyzer with a high sensitivity DNA chip was used for detecting the size of DNA after fragmentation. DNA from parents and pools was enriched for RLPs and RLKs with the customized MYbaits custom kit detailed earlier (Notes S4) (Arbor Biosciences Inc.), following a hybridization period of 24 h. Postenrichment PCR was performed, and products were quantified by Qubit.

| Solanum species                       | Genotype | Ploidy (EBN) | Country | INF1 | SCR74 | CRN2 | pGR106 |
|---------------------------------------|----------|--------------|---------|------|-------|------|--------|
| <i>S. brachistotrichum</i>            | BST325-1 | 2x (1EBN)    | MEX     | 2    | 8     | 10   | 0      |
| <i>S. lesteri</i>                     | LES358-2 | 2x           | MEX     | 10   | 10    | 4    | 0      |
| <i>S. demissum</i>                    | DMS585-1 | 6x (4EBN)    | MEX     | 4    | 0     | 9    | 0      |
| <i>S. demissum</i>                    | DMS582-1 | 6x (4EBN)    | MEX     | 7    | 2     | 7    | 0      |
| <i>S. edinense</i>                    | EDN150-4 | 5x           | MEX     | 10   | 0     | 10   | 0      |
| <i>S. edinense</i>                    | EDN151-1 | 5x           | MEX     | 10   | 4     | 10   | 0      |
| <i>S. semidemissum</i>                | SEM295-1 | 6x           | MEX     | 10   | 9     | 10   | 2      |
| <i>S. hougasii</i>                    | HOU271-1 | 6x (4EBN)    | MEX     | 0    | 6     | 9    | 0      |
| <i>S. hougasii</i>                    | HOU654-1 | 6x (4EBN)    | MEX     | 0    | 7     | 10   | 0      |
| <i>S. hougasii</i>                    | HOU655-1 | 6x (4EBN)    | MEX     | 0    | 10    | 9    | 0      |
| <i>S. papita</i>                      | PTA370-1 | 4x (2EBN)    | MEX     | 2    | 10    | 10   | 0      |
| <i>S. papita</i>                      | PTA370-2 | 4x (2EBN)    | MEX     | 1    | 10    | 8    | 0      |
| <i>S. papita</i>                      | PTA370-5 | 4x (2EBN)    | MEX     | 2    | 10    | 10   | 0      |
| <i>S. papita</i>                      | PTA767-1 | 4x (2EBN)    | MEX     | 4    | 7     | 10   | 0      |
| <i>S. papita</i>                      | PTA767-8 | 4x (2EBN)    | MEX     | 2    | 9     | 8    | 2      |
| <i>S. polytrichon</i>                 | PLT378-2 | 4x (2EBN)    | MEX     | 0    | 10    | 10   | 0      |
| <i>S. polytrichon</i>                 | PLT378-4 | 4x (2EBN)    | MEX     | 0    | 10    | 10   | 0      |
| <i>S. stoloniferum</i>                | STO389-4 | 4x (2EBN)    | MEX     | 7    | 3     | 10   | 0      |
| <i>S. stoloniferum</i>                | STO607-1 | 4x (2EBN)    | MEX     | 0    | 8     | 9    | 1      |
| <i>S. chacoense</i>                   | CHC262-4 | 2x (2EBN)    | BOL     | 3    | 8     | 10   | 1      |
| <i>S. chacoense</i>                   | CHC543-5 | 2x (2EBN)    | BOL     | 10   | 0     | 10   | 0      |
| <i>S. astleyi</i>                     | AST114-5 | 2x (2EBN)    | BOL     | 6    | 4     | 9    | 0      |
| <i>S. megistacrolobum toralapanum</i> | TOR704-4 | 2x (2EBN)    | BOL     | 8    | 0     | 5    | 1      |
| <i>S. huancabambense</i>              | HCB353-8 | 2x (2EBN)    | PER     | 7    | 2     | 10   | 4      |
| <i>S. huancabambense</i>              | HCB354-1 | 2x (2EBN)    | PER     | 8    | 1     | 10   | 0      |
| <i>S. mochiquense</i>                 | MCQ186-2 | 2x (1EBN)    | PER     | 8    | 0     | 10   | 0      |
| <i>S. microdontum</i>                 | MCD360-1 | 2x (2EBN)    | ARG     | 8    | 0     | 10   | 0      |
| <i>S. microdontum gigantophyllum</i>  | GIG362-6 | 2x (2EBN)    | ARG     | 0    | 10    | 10   | 0      |
| <i>S. microdontum gigantophyllum</i>  | GIG712-6 | 2x (2EBN)    | BOL     | 10   | 0     | 9    | 0      |
| <i>S. okadae/ venturi</i>             | OKA366-2 | 2x (2EBN)    | ARG     | 6    | 1     | 10   | 0      |
| <i>S. okadae/ venturi</i>             | OKA367-1 | 2x (2EBN)    | ARG     | 4    | 2     | 10   | 0      |
| <i>S. okadae/ venturi</i>             | OKA741-1 | 2x (2EBN)    | ARG     | 7    | 6     | 9    | 2      |
| <i>S. okadae/ venturi</i>             | VNT250-2 | 2x (2EBN)    | ARG     | 7    | 2     | 10   | 0      |
| <i>S. tuquerrense</i>                 | TUQ299-4 | 4x (4EBN)    | ECU     | 4    | 1     | 6    | 0      |

**Fig. 2** *Solanum* species show specific response to INF1 and SCR74 after potato virus X (PVX) agroinfection. *Solanum* genotypes that respond to either pGR106-INF1 or pGR106-SCR74, or both, are indicated. The empty vector pGR106 and the vector containing the CRN2 cell death-inducer, pGR106-CRN2, are included as negative and positive controls, respectively. The responses are scored from 0 to 10 and presented as a heat map ranging from no response (0–2, blank), weak response (3–4 yellow), medium response (5–6 orange), and strong response (7–10, red). Experiments were independently repeated at least three times. The ploidy level and endosperm balance number (EBN) are shown. Countries of origin: MEX, Mexico; BOL, Bolivia; PER, Peru; ARG, Argentina; ECU, Ecuador.

Paired-end sequencing was performed on a single Illumina MiSeq platform (San Diego, CA, USA) lane for six individually indexed samples including the INF1 and SCR74 responsive and nonresponsive bulks as well as the parents of the 7026 population, GIG362-6 and MCD360-1.

All RLP/RLK-enriched Illumina MiSeq raw reads were deposited at NCBI Sequence Read Archive (SRA) under accession PRJNA396439.

#### Read mapping and single nucleotide polymorphism (SNP) calling

Paired-end Illumina MiSeq reads were quality- and adapter-trimmed with FASTP (doi: 10.1093/bioinformatics/bty560) to a minimum base quality of 20. The trimmed reads were then mapped to the DM (v.4.03) or SOLYNTUS (v.1.0) reference genomes (<https://www.plantbreeding.wur.nl/Solyntus/>) using

BOWTIE2 (v.2.2.1) (Langmead & Salzberg, 2012) in very-sensitive end-to-end mode. Discordant and mixed mappings were disabled and the maximum insert was set to 1000 bp. Two score-min parameters were used in different mapping runs: 'L,-0.3,-0.3' and 'L,-0.18,-0.18', approximately equal to 5% and 3% mismatch rates respectively or 'L,-0.54,-0.54' for the Solyntus reference (9% mismatch). The Binary Alignment Map (BAM) files for the bulks were sorted, merged and indexed using SAMTOOLS (v.0.1.18; Li *et al.*, 2009), as were the BAM files for the parents. PILEUP files were generated for the bulk and parents using SAMTOOLS mpileup with default settings and piped into VARS-CAN MPILEUP2SNP (v.2.3.7; (Koboldt *et al.*, 2012)) with --strand-filter 0 and --output-vcf 1 for variant calling.

### Diagnostic RLP/KSeq

A dRenSeq-type analysis was conducted to ascertain the presence and sequence integrity of the known functional target gene *ELR*. The mapping condition for the diagnostic analysis of the RLP/KSeq-derived reads was as described previously (Armstrong *et al.*, 2019), and adapted for RLP/KSeq. For this, the *ELR* sequence was used as reference (GenBank no. MK388409.1).

### Read coverage and on target estimation

The percentage of reads on target was calculated as the proportion of reads mapping to a targeted *RLP/RLK* region in the DM reference (Notes S1). The mean read coverage to *RLP/RLK* genes was calculated from the previously generated BAM files using BEDTOOLS coverage (Table 1).

### SNP filtering

Single nucleotide polymorphisms were filtered using custom Java code (Notes S5) to retain informative SNPs present in both bulks and parents. SNPs were filtered based on the expected allele ratio for responsive/nonresponsive bulks/plants (Rr, responsive; rr, nonresponsive). To be retained, each SNP had a minimum read depth of 50 and alternate allele ratios reflecting the expected genotype: 0–10% or 90–100% alternate allele for nonresponsive

and 40–60% alternate allele for responsive bulks/plants. BEDTOOLS intersect (v.2.20.1; (Quinlan & Hall, 2010)) was used to extract SNPs present in both bulks and parents (informative SNPs) and to relate the informative SNP locations to annotated *RLP/RLK* genes. The number of parental, bulk and informative SNPs and variant genes were plotted in 1 Mb bins over each chromosome and visualized using R.

### High-resolution melt (HRM) marker development and single sequence repeat (SSR) markers

The BAM and VCF files for the filtered informative SNPs were visualized using GENEIOUS R10 (Kearse *et al.*, 2012) (<http://www.geneious.com>). Primers were designed in GENEIOUS R10 for the PCR products to contain the informative SNP(s) and a size between 80 and 150 bp. Primers flanking the informative SNPs were manually selected on the conserved sequences of both parents, responsive (R) and nonresponsive (NR) bulks. The HRM markers were tested on the parents and the F<sub>1</sub> progenies with the following protocol for a 10 µl reaction mixture: (1 µl template (20 ng gDNA), 1 µl dNTP (5 mM), 0.25 µl forward primer and 0.25 µl reverse primer (10 Mm), 1 µl LCGreen<sup>®</sup> Plus+ (BioFire, Salt Lake City, Utah, USA), 2 µl 5x Phire Buffer, 0.06 µl Phire taq, 4.44 µl MQ water (Millipore Corp., Billerica, MA, USA). Black 96-well microtiter PCR plates with white wells were used and 20 µl mineral oil was added to prevent evaporation. The protocol for PCR cycling is as follows: 95°C for 3 min (95°C for 10 s, 60°C for 15 s, 72°C for 30 s) with 40 cycles, then 72°C for 2 min followed by 94°C for 40 s. The LightScanner<sup>®</sup> System (BioFire) was used for measuring and analysing the melting curve. The primers used in this study are listed in Table S2. A further 78 SSR markers described in Milbourne *et al.* (1998) were used in this study (Table S3).

### Bacterial artificial chromosome (BAC) library

A BAC library of plant GIG362-6 was generated by Bio S&T (Saint-Laurent, Quebec, Canada). A BAC clone that spans the mapping interval was isolated using molecular markers (Table S2) and subsequently sequenced using PACBIO sequencing

**Table 1** Receptor-like protein/kinase enrichment sequencing (RLP/KSeq) read statistics: enriched reads are mapped to the DM genome v.4.03 at 3%, 5%, mismatch rates and the number of reads that map to target genes are specified.

|                          | PE reads | Total reads | Mismatch | Reads mapped | %         | On target | %         | Average read depth |
|--------------------------|----------|-------------|----------|--------------|-----------|-----------|-----------|--------------------|
| MCD360-1                 | 2284 289 | 4568 578    | 3        | 1757 896     | 38.477 97 | 870 460   | 49.517 15 | 54.26              |
|                          |          |             | 5        | 2541 388     | 55.627 55 | 1189 339  | 46.7988   | 72.88              |
| GIG362-6                 | 2589 095 | 5178 190    | 3        | 2024 698     | 39.1005   | 1020 145  | 50.385 05 | 64.03              |
|                          |          |             | 5        | 2957 160     | 57.107 99 | 1397 279  | 47.250 71 | 86.25              |
| INF1 responsive bulk     | 2251 593 | 4503 186    | 3        | 1669 192     | 37.066 91 | 902 139   | 54.046 45 | 55.3               |
|                          |          |             | 5        | 2450 878     | 54.425 42 | 1210 312  | 49.382 79 | 74.39              |
| INF1 nonresponsive bulk  | 2345 320 | 4690 640    | 3        | 1741 082     | 37.118 22 | 952 555   | 54.710 52 | 59.12              |
|                          |          |             | 5        | 2564 836     | 54.679 87 | 1282 185  | 49.990 92 | 79.74              |
| SCR74 responsive bulk    | 2598 380 | 5196 760    | 3        | 1970 580     | 37.9194   | 1057 408  | 53.659 73 | 65.75              |
|                          |          |             | 5        | 2905 738     | 55.914 42 | 1421 961  | 48.936 31 | 88.58              |
| SCR74 nonresponsive bulk | 2502 842 | 5005 684    | 3        | 1818 864     | 36.335 97 | 969 584   | 53.307 12 | 59.66              |
|                          |          |             | 5        | 2691 420     | 53.767 28 | 1311 170  | 48.716 66 | 80.93              |

(INRA-CNRGV). The GenBank accession number is MT270812.

## Results

### A wide range of wild *Solanum* species respond to apoplastic effectors of *P. infestans*

To explore the recognition spectra of apoplastic effectors from *P. infestans*, transient effectoromics screens with INF1 elicitor and SCR74 were performed on a wide range of wild *Solanum* genotypes (Fig. 1a). In total, 100 *Solanum* genotypes were screened for responses to INF1 and SCR74 by PVX agroinfection. An empty vector and the general cell death-inducing crinkling and necrosis-inducing protein (CRN2) were included as negative and positive controls, respectively. An overview of all tested plants, including responsive as well as nonresponsive plants, is presented in Table S1. A set of 34 *Solanum* genotypes showed specific cell death responses to INF1 and/or SCR74 at 2 wk after agroinfection (Fig. 2). These responsive plants belong to 17 different wild *Solanum* species, vary in ploidy levels as well as endosperm balance numbers, and originate from different geographic origins (Fig. 2). In most cases, the specific effector responses were clear and highly reproducible (i.e. clear cell death phenotypes scores > 7). In some cases, we observed more variability (cell death phenotypes scores ranging from 4 to 6), but these variations were also observed for the positive control CRN2 in some genotypes, which suggest that these plants were less amenable to the PVX-based transient expression system. As expected, response to INF1 elicitor was confirmed in MCD360-1 (Fig. 2), which is the source of the elicitor receptor ELR (Du *et al.*, 2015). In addition, other *Solanum* genotypes were also found to respond to INF1 (Fig. 2; Table S1). Similarly, SCR74 was recognized in various plants including GIG362-6 (Fig. 2; Table S1). In conclusion, responses to INF1 and SCR74 are widely found in wild *Solanum* species, which suggests that surface receptors that recognize these effectors are present in these plants.

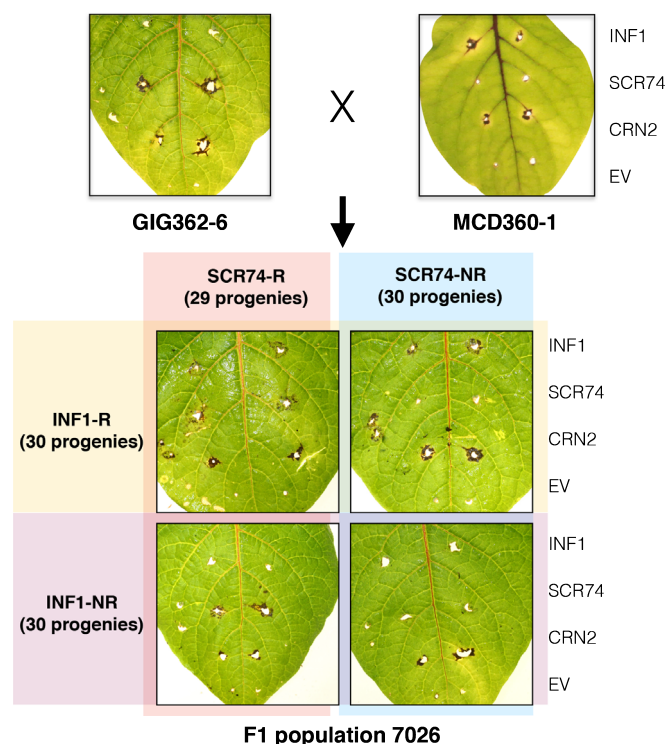
### Response to INF1 and SCR74 segregates independently in *S. microdontum*

To genetically map the gene encoding the immune receptor that recognizes SCR74 and to confirm the location of the INF1 receptor (ELR), a mapping population was developed (Fig. 1a). We crossed MCD360-1 with GIG362-6 and generated the F<sub>1</sub> population 7026 (Fig. 3). From this population, 100 progenies were tested for responses to INF1 and SCR74 by PVX agroinfection. The population segregated for clear responses to INF1, with 53 responsive vs 41 nonresponsive offspring clones, which is close to a 1 : 1 segregation ( $\chi^2 = 1.532$ ,  $P = 0.216$ ). Reproducible segregation for responses to SCR74 was also observed at a near 1 : 1 ratio ( $\chi^2 = 0.36$ ,  $P = 0.549$ ), as 47 responsive vs 53 nonresponsive offspring genotypes were identified. Importantly, the responses to SCR74 were independent of the responses to INF1. Both segregation ratios are consistent with two different dominant loci that mediate the responses to INF1 and SCR74, respectively.

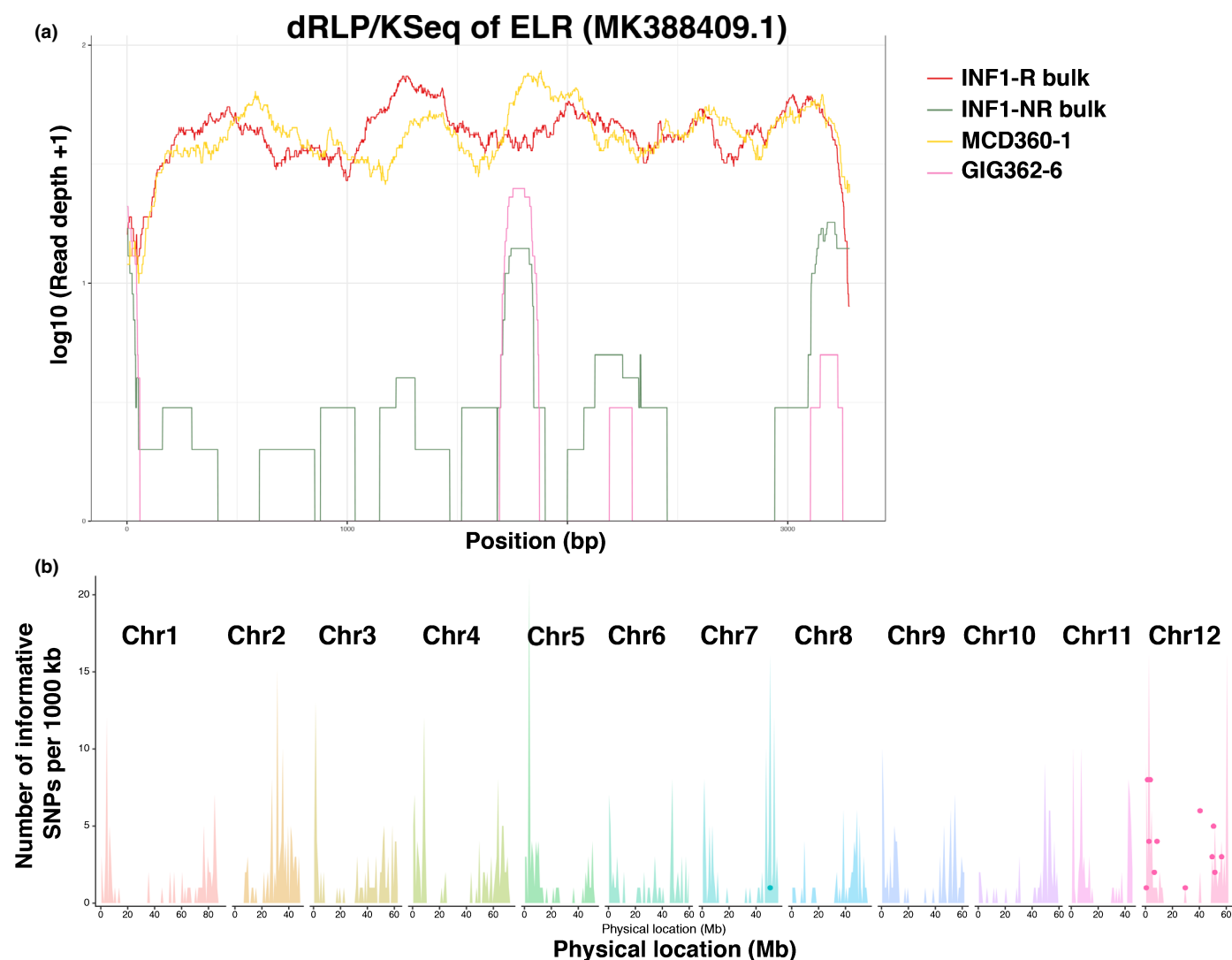
### Designing the RLP/RLK bait library for target enrichment sequencing

For mapping the gene that confers recognition of SCR74, we developed an RLP/KSeq approach, based on adapting previously described RenSeq targeted enrichment technology to nonNLR genes (Jupe *et al.*, 2013) (Fig. 1b). As the INF1 receptor ELR was originally cloned from MCD360-1 (Du *et al.*, 2015), we used this genotype and the segregating progeny as a positive control throughout this study.

To design a comprehensive bait library for *Solanum* RLPs and RLKs, we combined 301 LRR-RLK and 404 LRR-RLP genes previously predicted in potato (Andolfo *et al.*, 2013) with *de novo* identified genes. A combination of BLASTP, MEME and Pfam searches was utilized to predicted 533 RLK genes and 444 RLP genes from the potato reference genome DM1-3, v.4.03 (Notes S1), including 70 RLK with WAX or WAX-EGF domain, 38 RLK with malectin domain, 11 RLK



**Fig. 3** Independent segregation of responses to SCR74 and INF1 in F<sub>1</sub> population 7026 of *Solanum microdontum*. *Solanum microdontum* ssp. *gigantophyllum* GIG362-6 (SCR74 responsive) was crossed with *S. microdontum* MCD360-1 (INF1 responsive) and progeny plants were assessed for phenotypic responses to INF1 and SCR74 through agroinfection with pGR106-INF1 and pGR106-SCR74. The empty vector pGR106 and the vector containing the CRN2 cell death-inducer, pGR106-CRN2, were included as negative and positive controls, respectively. For INF1, 30 responsive (INF1-R) and 30 INF1 nonresponsive (INF1-NR) progeny plants were identified. Similarly for SCR74, 29 and 30 SCR74 responsive (SCR74-R) and SCR74 nonresponsive (SCR74-NR) progeny plants were selected for the receptor-like protein/kinase enrichment sequencing (RLP/KSeq), respectively. Representative images are shown at 14 d post-infection (dpi).



**Fig. 4** *ELR* was recovered and independently mapped to chromosome 12 by receptor-like protein/kinase enrichment sequencing (RLP/KSeq). (a) Diagnostic RLP/KSeq for *ELR*. The x-axis depicts the coding DNA sequence (CDS) of *ELR* from start to stop and the y-axis indicates the read coverage of functional *ELR* with RLP/Kseq-derived reads mapped to the reference under highly stringent conditions on a log scale. The yellow and red horizontal lines indicate full-length *ELR* sequence from *Solanum microdontum* MCD360-1 and INF1 responsive bulk without any sequence polymorphisms, respectively. The green and pink lines show a low and discontinuous read-coverage from *S. microdontum* ssp. *gigantophyllum* GIG362-6 and INF1 nonresponsive bulk, respectively. (b) Mapping of *ELR*. The x-axis represent the physical positions of the 12 individual potato chromosomes and the y-axis the number of RLP/receptor-like kinases (RLKs) or single nucleotide polymorphisms (SNPs) per 1 Mb bin. The background colour spikes represent the number and position of annotated RLP/RLKs and the coloured dots depict the position of significant and linked SNPs in a 1 MB bin. The peak in chromosome 12 indicates various SNPs that are linked with *ELR*, which confers response to INF1.

with antifungal domain, six *RLK* with ANK repeat, 11 *LysM* *RLKs*, 24 *L-LecRK*, 103 *G-LecRK*, one *C-LecRK* and 22 other *RLKs* with transmembrane domain. Additionally, 18 known *Solanaceae* *RLP/RLK* genes from were included (Notes S3) alongside the *RLP/RLK* homologs from *Solanum chacoense* (M6) (Leisner *et al.*, 2018; Notes S2).

A customized target enrichment RNA bait library with 2× coverage comprising 57 020 120-mer biotinylated RNA oligo probes was synthesized (MYbaits custom kit; Arbor Biosciences Inc.) (Notes S4). The long RNA baits can tolerate mismatches like SNPs and indels (Clark *et al.*, 2011) and were used for the mapping of the INF1 and SCR74 receptors (Fig. 1a,b).

### Bulked segregant analysis (BSA) and RLP/K enrichment

To map the genes that mediate response to INF1 and SCR74 using RLP/KSeq, we used a BSA approach. Normally, for mapping one gene, it would require two pools (i.e. responsive and nonresponsive) plus the two parents (Fig. 1c). In this case, as we multiplex for two target genes, we created four bulked pools. These comprised response to INF1 (INF-R: 30 plants), no response to INF1 (INF1-NR: 30 plants), response to SCR74 (SCR74-R: 29 plants), no response to SCR74 (SCR74-NR: 30 plants), progeny individuals, respectively (Fig. 3). DNA was isolated from each clone and then pooled before indexing. DNA

from the parents GIG362-6 and MCD360-1 was individually indexed and included in the enrichment.

### Mapping reads to the reference genome and SNP calling

The *RLP/RLK* enriched DNA libraries from the bulks and parents were sequenced with Illumina 2 × 250 bp chemistry on a MiSeq platform (Fig. 1c). The number of raw reads that passed quality control ranged from 4503 186 to 5196 760 in different samples/pools (Table 1). High-quality paired-end reads were mapped to the potato reference genome (DM v.4.03) using BOWTIE2. To compensate for differences between the potato reference DM and *S. microdontum*, two mismatch rates, 3% and 5%, were used for the read mapping (Table 1). The mapping rates ranged from 36% to 57%, with reads on target accounting for 46–55%, depending on the mismatch rate (Table 1). The resulting coverage of known *RLP/RLK* genes was calculated and ranged from ×54 to ×89. To enable the identification of informative SNPs whilst ensuring sufficient accuracy, a 5% mismatch rate was used for further analysis. SNPs were called by SAMTOOLS and VARSCAN from different samples, and the output SNPs were filtered by a custom java script (Notes S5; Chen *et al.*, 2018).

### Diagnostic analysis of RLP/KSeq-derived reads confirms presence and sequence integrity of INF1 receptor ELR

To validate our targeted enrichment sequencing approach and to confirm that RLP/KSeq specifically yields sequence representation of expected RLPs/RLKs, we used diagnostic mapping of enriched samples to previously characterized, functional gene sequences as a control. In line with dRenSeq (Van Weymers *et al.*, 2016; Armstrong *et al.*, 2019), we refer to this approach as dRLP/KSeq. As a proof of concept, we assessed the sequence representation of the known INF1 receptor ELR that was expected to be present in the INF1 responsive parent of the population 7026, MCD360-1, as well as in the INF1 responsive bulk, but expected to be absent in the nonresponsive parent, GIG362-6, and the non-responsive bulk.

In line with this expectation, dRLP/KSeq revealed continuous coverage of ELR in the progenitor parent of the INF1 receptor and the responsive bulk. Indeed, a very similar nucleotide representation profile was observed for both samples and only the very 5' and 3' regions of ELR are not resolved owing to a lack of flanking sequences in the reference that prevented the mapping of RLP/KSeq-derived reads that extend from the gene into the 5' and 3' untranslated regions, respectively (Fig. 4a). By contrast, functional ELR sequence representation in the nonresponsive parent and bulk was very limited and discontinuous, which is in accordance with the absence of the function receptor sequence in these samples. The partial coverage observed hints at the present of nonfunctional ELR homologues.

### De novo mapping of the INF1 response using unrelated potato reference genomes coincides with the physical position of the ELR receptor and identifies linked SNPs in closely related homologues

Following the successful dRLP/KSeq analysis, we assessed the suitability of using RLP/KSeq-derived reads for the mapping of receptors using ELR as an example. SNPs from the population parents alongside INF1 nonresponsive and responsive bulks were called and filtered for the expected ratios of heterozygosity as described by Chen *et al.*, (2018). In short, for a single dominant gene segregating in a diploid population, the allele frequencies were set at 0–10% or 90–100% for the INF1 nonresponsive bulk and parent as well as 40–60% for INF1 responsive bulk and responsive parent. The SNPs from the bulks were independently validated through comparison with parental SNPs, and only the accordant SNPs at the correct ratio were maintained as informative SNPs (Table S4). Allowing for a 5% mismatch rate for the positioning of RLP/KSeq-enriched reads and determined SNPs from *S. microdontum* against the *S. phureja* reference genome (DM), 99 SNPs passed the filter criteria in the bulks and 4323 SNPs in the parents. Among those, 48 SNPs were shared in both bulks and parents (Table S4). The number of informative genic SNPs per 1 Mb interval was placed on the 12 chromosomes of potato. With the exception of one significant SNP on chromosome 7, the remaining 47 SNPs were positioned on Chr12. The SNPs were found to localize in two major locations on chromosome 12, one near the bottom and one at the top of chromosome 12 where ELR resides (Fig. 4b). The majority of SNPs were localized in two *RLP/RLK* loci that correspond to 19 polymorphic genes. Intriguingly, the sequence that is most similar to *ELR* in the DM reference genome has not been placed on any linkage group and is currently found in the unassembled chromosome 00. This highlights some remaining ambiguity within the DM reference genome which does not contain functional ELR.

Thus, we also mapped the reads to the recently released but largely unannotated potato genome Solyntus (v.1.0) (see the Materials and Methods section). Essentially, we observed a somewhat similar distribution of informative SNPs as seen in DM (Table S5). However, in Solyntus, the most similar sequence to *ELR* is 92.8% identical and spans the physical position between base pairs 2491 817 and 2495 104 on chromosome 12. Allowing for a 9% mismatch rate, we observed an informative SNP at position 2493 665 in this gene (Table S5). In summary, despite the absence of a true *ELR* gene in DM and Solyntus, RLP/KSeq led to the correct mapping of the ELR locus on chromosome 12 and identified an informative SNP within the closest homologue of ELR in Solyntus.

### RLP/KSeq accelerates mapping of the SCR74 response gene on chromosome 9

To map the single dominant gene that confers the response to SCR74, the same SNP filtering approach was performed as shown for ELR in the previous section. The SNPs that meet 0–5% or 95–100% allele frequency in the SCR74 nonresponsive

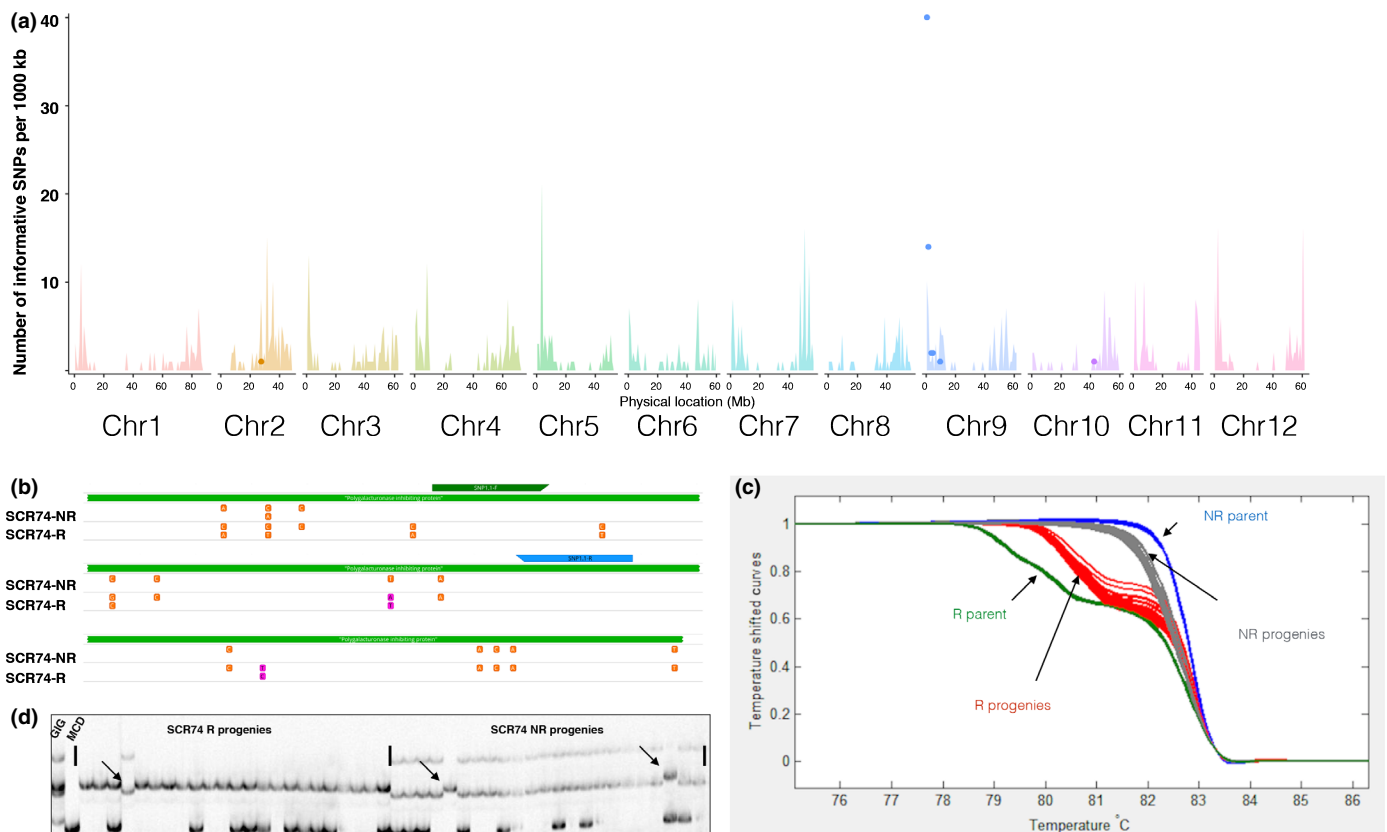
bulk and parent as well as 45–55% allele frequency in the SCR74 responsive bulk and parent were identified and then independently corroborated between bulks and the parental material. This resulted in the identification of 61 informative SNPs, of which 60 could be placed on chromosome 9. The SNPs correspond to 16 polymorphic genes (Fig. 5a; Table S4).

One of the identified SNPs, RLP/KSeq-snp1.1 (A → T), corresponds to PGSC0003DMG400008492, a polygalacturonase-inhibiting protein (PGIP) gene that resides at position 0.16 Mb on chromosome 9 (Fig. 5b). This SNP displayed a 59% frequency in the responsive bulk and 0% or 100% frequency in the nonresponsive bulk and was used to independently corroborate the mapping position of the receptor on potato linkage group 9. We converted the SNP to HRM marker, RLP/KSeq-snp1.1, and tested it on the mapping parents and the 56 progeny genotypes of the F<sub>1</sub> population. Our result indicates that the SCR74 receptor is linked to this marker which resides on chromosome 9 (Fig. 5c).

To further confirm our RLP/KSeq methods, 78 SSR markers dispersed over all 12 potato chromosomes were tested on 56 F<sub>1</sub> progeny of population 7026 (Table S3). SSR marker STM1051 was found to be linked to the SCR74 responsive phenotype, and three recombination events were detected (Fig. 5d). This marker resides in position 6.15 Mb of DM chromosome 9 and independently confirms the RLP/KSeq mapping analysis for the SCR74 response (Fig. 5c). Consequently, the SCR74 response gene was mapped to a 10.7 cM region on potato chromosome 9 between RLP/KSeq-snp1.1 and STM1051, which spans a 5.99 Mb physical distance based on the DM genome.

### Fine-mapping of candidate SCR74 response gene to a 43 kbp *G-LecRK* locus

To fine-map the SCR74 response gene, we first genotyped 1500 progenies of population 7026 with flanking markers RLP/KSeq-snp1.1 and STM1051. As a result, 120 recombinants were

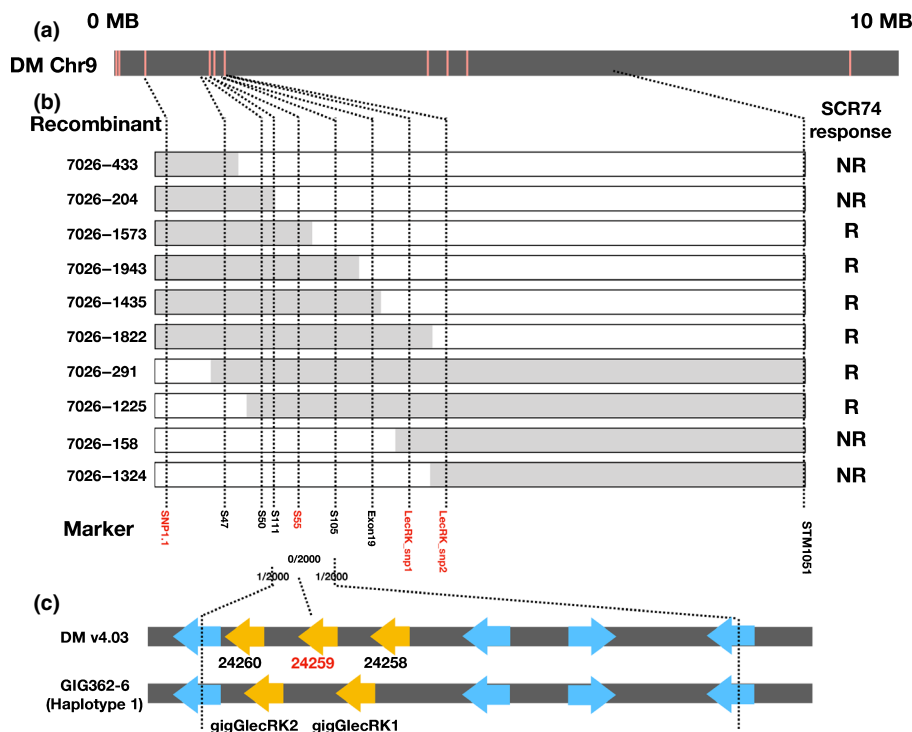


**Fig. 5** The gene conferring response to SCR74 was mapped to chromosome 9 by receptor-like protein/kinase enrichment sequencing (RLP/KSeq). (a) Mapping of SCR74. The x-axis represent the physical positions of the 12 individual potato chromosomes and the y-axis the number of RLP/receptor-like kinases (RLKs) or single nucleotide polymorphisms (SNPs) per 1 Mb bin. The background colour spikes represent the number and position of annotated RLP/RLKs genes and the coloured dots depict the position of significant and linked SNPs to SCR74 in a 1 Mb bin. The significant accumulation of SNPs on the top of chromosome 9 indicate the map position of SCR74 receptor. (b) Within the identified region on chromosome 9, a polygalacturonase-inhibiting protein (PGIP, PGSC0003DMG400006492, green bar) contains 15 SNPs (orange). Two of these SNPs (pink) show a near 1 : 1 frequency in the SCR74-B3b responsive pool and, together with marker RLP/KSeq-snp1.1 (green and blue arrow), flank the interval. (c) Melting curves of the high-resolution melting (HRM) marker RLP/KSeq-snp1.1 tested on the mapping parents and progenies. (d) Single sequence repeat (SSR) marker STM1051 on chromosome 9 is linked with the SCR74 response. The mapping parents *Solanum microdontum* ssp. *gigantophyllum* GIG362-6 and *S. microdontum* MCD360-1, as well as the responsive progenies and nonresponsive progenies were tested with STM1051 and three recombination events (arrow) were found. This figure is reproduced from Domazakis *et al.* (2017).

identified. To further narrow down the mapping interval, we genotyped 500 additional progenies and developed more SNP markers for genes predicted to reside in this interval and for which RLP/KSeq had identified SNPs (Fig. 6a). The latter included two *L-LecRK* genes, PGSC0003DMG400008822 and PGSC0003DMG400008897, and a *G-LecRK* gene, PGSC0003DMG400024259 (Table S4). By testing those RLP/KSeq markers and other SNP markers developed within this interval (Table S2), we located the candidate gene between marker S111 and S105 (Fig. 6b). In DM, the mapping interval contains eight genes, including three receptor-like kinases with a G-type lectin domain (*G-LecRK*) genes, a putative reticulate-related 1 like gene, a serine/threonine-protein kinase ATG1c-like (autophagy-related protein) gene, a prenylated rab acceptor family gene and an uracil phosphoribosyltransferase encoding gene. Remarkably, of the markers, S55, was derived from the RLP/KSeq analysis and locates within a *G-LecRK* gene (PGSC0003DMG400024259). This marker displays perfect linkage and cosegregates with the SCR74 response (0 recombinants out of 2000 F<sub>1</sub> progeny; Fig. 6b). To obtain the physical representation of GIG362-6, a BAC library was generated. A BAC clone that covers the mapping interval was isolated (Fig. 6c); unlike in DM, only two *G-LecRK* genes are located in this region. The physical distance between the two flanking markers in the GIG362-6 responsiveness haplotype is 43 kbp.

## Discussion

In this paper, we present a work flow that combines RLP/KSeq with effectoromics of apoplastic effectors, to rapidly map plant surface immune receptors (Fig. 1). We used potato and *P. infestans* as a model system. We screened wild potato species that mount specific cell death response to the apoplastic effectors INF1 and/or SCR74 of *P. infestans*. *Solanum microdontum* MCD360-1, which responds to INF1, and GIG362-6, which responds to SCR74, were crossed in order to generate a population that segregates for both responses independently. In parallel, we designed bait libraries based on predicted *RLP* and *RLK* genes from the potato genome. We subjected pools of genomic DNA derived from responding vs nonresponding genotypes to a BSA RLP/K enrichment sequencing approach, using a bespoke bait library to enrich for genomic DNA representing our target genes. This approach quickly led to the identification of SNPs that are linked to the phenotype and could be used as molecular markers to genetically map the genes encoding the putative RLP/RLK genes. Here, we have shown that RLP/KSeq can successfully identify informative SNPs in the ELR receptor that underpins INF1 responses, obtain full-length sequence representation of ELR in responsive parent and bulks through dRLP/KSeq analysis, and fine-mapped a novel receptor for SCR74 response to a 43 kbp interval containing two *G-LecRK* genes.



**Fig. 6** Fine-mapping the SCR74 response gene to a *G-LecRK* locus. (a) Graphical representation of the top 10 Mb of chromosome 9 of *S. tuberosum* Group Phureja DM1-3 (DM) (black bar) and the location of informative single nucleotide polymorphisms (SNPs) (red lines). (b) Overview of recombinant screening, showing representative recombinants with their genotyping and phenotyping results abbreviated as (R) for response and (NR) for nonresponse to SCR74. The position of the responsive haplotype of *Solanum microdontum* ssp. *gigantophyllum* GIG362-6 (grey bar), the nonresponsive haplotype of *S. microdontum* MCD360-1 (white bar) and the markers (dotted lines) are indicated. Receptor-like protein/kinase enrichment sequencing (RLP/KSeq)-derived markers are shown in red. (c) The *G-LecRK* locus of DM1-3 (v.4.03) and GIG362-6 based on the sequencing of a bacterial artificial chromosome (BAC) clone that spans the SCR74-responsive interval on chromosome 9. The locations of *G-LecRK* (yellow arrows) and other predicted genes (blue arrows) are marked.

With continuous advances of sequencing technology, genotyping by sequencing has already been applied to clone plant genes in multiple crops (Huang *et al.*, 2009; Austin *et al.*, 2011; Mascher *et al.*, 2014; Pandey *et al.*, 2017). However, when the genome size is large, or when certain types of genes are expected, targeted enrichment sequencing becomes a preferential option, as it can dramatically reduce the genome complexity (Hodges *et al.*, 2007). RenSeq and its descendants, such as dRenSeq, MutRenSeq, SMRT RenSeq and AgRenSeq, have been demonstrated to be powerful tools to clone plant disease resistance genes (Steuernagel *et al.*, 2016; Van Weymers *et al.*, 2016; Witek *et al.*, 2016; Armstrong *et al.*, 2019; Arora *et al.*, 2019). However, they all target *NLR* genes. RLP/KSeq can complement the RenSeq toolbox by targeting additional types of plant immune receptors, including RLPs/RLKs, that also function as effective immune receptors (Boutrot & Zipfel, 2017).

Effectoromics has proven to be a medium to high-throughput approach to identify plants carrying *R* genes as well as surface immune receptors (Vleeshouwers *et al.*, 2011a; Du *et al.*, 2015; Domazakis *et al.*, 2017). The specificity and robustness of effector responses enable us to identify multiple receptors in a single segregating population (Fig. 3). Another advantage of combining the enrichment sequencing with effectoromics is that targeted libraries can be used for PRR or NLR, depending on the matching effector response. Effectoromics was pioneered for the potato–late blight pathosystem and has been successfully applied in various other *Solanaceae*, such as *Nicotiana benthamiana*, tomato and pepper (Takken *et al.*, 2000; Oh *et al.*, 2009; Lee *et al.*, 2014). Beyond *Solanaceae*, the approach has been used in other plants such as sunflower (Gascuel *et al.*, 2016), as well as in various plant pathogens such as fungi, nematodes and insects (Catanzariti *et al.*, 2006; Sacco *et al.*, 2009; Hogenhout & Bos, 2011). This demonstrates the wide application of the effectoromics strategy for pathogens with well-characterized genomes.

To summarize, our newly developed pipeline enables us to rapidly identify plants carrying novel immune receptors and to genetically map the genes responsible for the phenotype. This strategy complements the current RenSeq toolbox and will help us to understand the first layer of the plant immune system and ultimately to develop more durable disease resistance in plants.

## Acknowledgements

This work was supported by a NWO-VIDI grant 12378 (XL, DW, VGA AV), Short Term Scientific Mission (STSM) of COST Action SUSTAIN FA1208 (XL), China Scholarship Council (CSC) (XL), the Rural & Environment Science & Analytical Services Division of the Scottish Government (IH) and the Biotechnology and Biological Sciences Research Council (BBSRC) through awards BB/L008025/1 (IH), BB/K018299/1 (IH) and BB/S015663/1 (IH). We thank Dr Hendrik Rietman for performing effectoromics screens, Isolde Bertram-Pereira for culturing *Solanum* plants, and Henk Smid and Harm Wiegersma for help in the glasshouse. We thank Dr Helene Berges and

Caroline Callot from the French Plant Genomic Resource Center (INRA-CNRCV) for their help to sequence the BAC clones. We thank Prof. Dr Evert Jacobsen for reviewing the manuscript. The authors declare that they have no competing interests.

## Author contributions

XL, VGA AV and IH planned and designed the research; XL, KB, MA and DW performed the experiment; and XL, MA, VGA AV, IH, RGFV and PJW wrote the manuscript.

## ORCID

Miles Armstrong  <https://orcid.org/0000-0002-9441-476X>  
 Katie Baker  <https://orcid.org/0000-0002-6265-9252>  
 Ingo Hein  <https://orcid.org/0000-0002-0128-2084>  
 Xiao Lin  <https://orcid.org/0000-0002-3848-0244>  
 Richard G. F. Visser  <https://orcid.org/0000-0002-0213-4016>  
 Vivianne G. A. A. Vleeshouwers  <https://orcid.org/0000-0002-8160-4556>  
 Pieter J. Wolters  <https://orcid.org/0000-0001-5941-9189>

## References

- Andolfo G, Sanseverino W, Rombauts S, Van de Peer Y, Bradeen JM, Carpato D, Frusciante L, Ercolano MR. 2013. Overview of tomato (*Solanum lycopersicum*) candidate pathogen recognition genes reveals important *Solanum* R locus dynamics. *New Phytologist* 197: 223–237.
- Armstrong MR, Vossen J, Lim TY, Hutten RCB, Xu J, Strachan SM, Harrower B, Champouret N, Gilroy EM, Hein I. 2019. Tracking disease resistance deployment in potato breeding by enrichment sequencing. *Plant Biotechnology Journal* 17: 540–549.
- Arora S, Steuernagel B, Gaurav K, Chandramohan S, Long Y, Matny O, Johnson R, Enk J, Periyannan S, Singh N *et al.* 2019. Resistance gene cloning from a wild crop relative by sequence capture and association genetics. *Nature Biotechnology* 37: 139–143.
- Austin RS, Vidaurre D, Stamatiou G, Breit R, Provart NJ, Bonetta D, Zhang JF, Fung P, Gong YC, Wang PW *et al.* 2011. Next-generation mapping of *Arabidopsis* genes. *Plant Journal* 67: 715–725.
- Boutrot F, Zipfel C. 2017. Function, discovery, and exploitation of plant pattern recognition receptors for broad-spectrum disease resistance. *Annual Review of Phytopathology* 55: 257–286.
- Catanzariti AM, Dodds PN, Lawrence GJ, Ayliffe MA, Ellis JG. 2006. Haustorially expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell* 18: 243–256.
- Chen X, Lewandowska D, Armstrong MR, Baker K, Lim TY, Bayer M, Harrower B, McLean K, Jupe F, Witek K *et al.* 2018. Identification and rapid mapping of a gene conferring broad-spectrum late blight resistance in the diploid potato species *Solanum verrucosum* through DNA capture technologies. *TAG. Theoretical and Applied Genetics* 131: 1287–1297.
- Clark MJ, Chen R, Lam HY, Karczewski KJ, Chen R, Euskirchen G, Butte AJ, Snyder M. 2011. Performance comparison of exome DNA sequencing technologies. *Nature Biotechnology* 29: 908–914.
- Dangl JL, Horvath DM, Staskawicz BJ. 2013. Pivoting the plant immune system from dissection to deployment. *Science* 341: 746–751.
- Derevnina L, Dagdas YF, De la Concepcion JC, Bialas A, Kellner R, Petre B, Domazakis E, Du J, Wu CH, Lin X *et al.* 2016. Nine things to know about elicitors. *New Phytologist* 212: 888–895.
- Domazakis E, Lin X, Aguilera-Galvez C, Wouters D, Bijsterbosch G, Wolters PJ, Vleeshouwers VG. 2017. Effectoromics-based identification of cell surface receptors in potato. *Methods in Molecular Biology* 1578: 337–353.

- Du J, Rietman H, Vleeshouwers VG. 2014. Agroinfiltration and PVX agroinfection in potato and *Nicotiana benthamiana*. *Journal of Visualized Experiments* 83: 50971.
- Du J, Verzaux E, Chaparro-Garcia A, Bijsterbosch G, Keizer LC, Zhou J, Liebrand TW, Xie C, Govers F, Robatzek S *et al.* 2015. Elicitor recognition confers enhanced resistance to *Phytophthora infestans* in potato. *Nature Plants* 1: 15034.
- Felix G, Duran JD, Volko S, Boller T. 1999. Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *The Plant Journal* 18: 265–276.
- Fry W. 2008. *Phytophthora infestans*: the plant (and Rgene) destroyer. *Molecular Plant Pathology* 9: 385–402.
- Gascuel Q, Buendia L, Pecrix Y, Blanchet N, Munos S, Vear F, Godiard L. 2016. RXLR and CRN effectors from the sunflower downy mildew pathogen *Plasmopara halstedii* induce hypersensitive-like responses in resistant sunflower lines. *Frontiers in Plant Science* 7: 1887.
- Gijzen M, Nürnberger T. 2006. Nep1-like proteins from plant pathogens: recruitment and diversification of the NPP1 domain across taxa. *Phytochemistry* 67: 1800–1807.
- Haverkort AJ, Boonekamp PM, Hutten R, Jacobsen E, Lotz LAP, Kessel GJT, Visser RGF, van der Vossen EAG. 2008. Societal costs of late blight in potato and prospects of durable resistance through cisgenic modification. *Potato Research* 51: 47–57.
- Hodges E, Xuan Z, Balija V, Kramer M, Molla MN, Smith SW, Middle CM, Rodesch MJ, Albert TJ, Hannon GJ *et al.* 2007. Genome-wide *in situ* exon capture for selective resequencing. *Nature Genetics* 39: 1522–1527.
- Hogenhout SA, Bos JIB. 2011. Effector proteins that modulate plant–insect interactions. *Current Opinion in Plant Biology* 14: 422–428.
- Huang XH, Feng Q, Qian Q, Zhao Q, Wang L, Wang AH, Guan JP, Fan DL, Weng QJ, Huang T *et al.* 2009. High-throughput genotyping by whole-genome resequencing. *Genome Research* 19: 1068–1076.
- Jiang R, Li J, Tian Z, Du J, Armstrong M, Baker K, Tze-Yin Lim J, Vossen JH, He H, Portal L *et al.* 2018. Potato late blight field resistance from QTL dP109c is conferred by the NB-LRR gene *R8*. *Journal of Experimental Botany* 69: 1545–1555.
- Jo KR, Kim CJ, Kim SJ, Kim TY, Bergervoet M, Jongsma MA, Visser RGF, Jacobsen E, Vossen JH. 2014. Development of late blight resistant potatoes by cisgene stacking. *BMC Biotechnology* 14: 50.
- Jones JDG, Dangl JL. 2006. The plant immune system. *Nature* 444: 323–329.
- Jupe F, Witek K, Verweij W, Sliwka J, Pritchard L, Etherington GJ, Maclean D, Cock PJ, Leggett RM, Bryan GJ *et al.* 2013. Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. *The Plant Journal* 76: 530–544.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C *et al.* 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647–1649.
- Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, Wilson RK. 2012. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Research* 22: 568–576.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9: 357–359.
- Lee HA, Kim SY, Oh SK, Yeom SI, Kim SB, Kim MS, Kamoun S, Choi D. 2014. Multiple recognition of RXLR effectors is associated with nonhost resistance of pepper against *Phytophthora infestans*. *New Phytologist* 203: 926–938.
- Leisner CP, Hamilton JP, Crisovan E, Manrique-Carpintero NC, Marand AP, Newton L, Pham GM, Jiang J, Douches DS, Jansky SH *et al.* 2018. Genome sequence of M6, a diploid inbred clone of the high-glycoalkaloid-producing tuber-bearing potato species *Solanum chacoense*, reveals residual heterozygosity. *The Plant Journal* 94: 562–570.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25: 2078–2079.
- Liu ZY, Bos JIB, Armstrong M, Whisson SC, da Cunha L, Torto-Alalibo T, Win J, Avrova AO, Wright F, Birch PRJ *et al.* 2005. Patterns of diversifying selection in the phytotoxin-like *sc74* gene family of *Phytophthora infestans*. *Molecular Biology and Evolution* 22: 1159–1159.
- Mascher M, Jost M, Kuon JE, Himmelbach A, Assfalg A, Beier S, Scholz U, Graner A, Stein N. 2014. Mapping-by-sequencing accelerates forward genetics in barley. *Genome Biology* 15: R78.
- Milbourne D, Meyer RC, Collins AJ, Ramsay LD, Gebhardt C, Waugh R. 1998. Isolation, characterisation and mapping of simple sequence repeat loci in potato. *Molecular and General Genetics* 259: 233–245.
- Oh SK, Young C, Lee M, Oliva R, Bozkurt TO, Cano LM, Win J, Bos JIB, Liu HY, van Damme M *et al.* 2009. In planta expression screens of *Phytophthora infestans* RXLR effectors reveal diverse phenotypes, including activation of the *Solanum bulbocastanum* disease resistance protein Rpi-blb2. *Plant Cell* 21: 2928–2947.
- Oome S, Raaymakers TM, Cabral A, Samwel S, Böhm H, Albert I, Nürnberger T, Van den Ackerveken G. 2014. Nep1-like proteins from three kingdoms of life act as a microbe-associated molecular pattern in Arabidopsis. *Proceedings of the National Academy of Sciences, USA* 111: 16955–16960.
- Pandey MK, Khan AW, Singh VK, Vishwakarma MK, Shasidhar Y, Kumar V, Garg V, Bhat RS, Chitkineni A, Janila P *et al.* 2017. QTL-seq approach identified genomic regions and diagnostic markers for rust and late leaf spot resistance in groundnut (*Arachis hypogaea* L.). *Plant Biotechnology Journal* 15: 927–941.
- Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26: 841–842.
- Ramirez-Gonzalez RH, Segovia V, Bird N, Fenwick P, Holdgate S, Berry S, Jack P, Caccamo M, Uauy C. 2015. RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. *Plant Biotechnology Journal* 13: 613–624.
- Sacco MA, Koropacka K, Grenier E, Jaubert MJ, Blanchard A, Govere A, Smant G, Moffett P. 2009. The cyst nematode SPRYSEC protein RBP-1 elicits Gpa2- and RanGAP2-dependent plant cell death. *PLoS Pathogens* 5: e1000564.
- Singh VK, Khan AW, Saxena RK, Sinha P, Kale SM, Parupalli S, Kumar V, Chitkineni A, Vechalapu S, Kumar CVS *et al.* 2017. Indel-seq: a fast-forward genetics approach for identification of trait-associated putative candidate genomic regions and its application in pigeonpea (*Cajanus cajan*). *Plant Biotechnology Journal* 15: 906–914.
- Steuernagel B, Periannan SK, Hernandez-Pinzon I, Witek K, Rouse MN, Yu GT, Hatta A, Ayliffe M, Bariana H, Jones JDG *et al.* 2016. Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nature Biotechnology* 34: 652–655.
- Strachan SM, Armstrong MR, Kaur A, Wright KM, Lim TY, Baker K, Jones J, Bryan G, Blok V, Hein I. 2019. Mapping the H2 resistance effective against *Globodera pallida* pathotype Pa1 in tetraploid potato. *TAG. Theoretical and Applied Genetics* 132: 1283–1294.
- Takagi H, Abe A, Yoshida K, Kosugi S, Natsume S, Mitsuoka C, Uemura A, Utsushi H, Tamiru M, Takuno S *et al.* 2013. QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *The Plant Journal* 74: 174–183.
- Takken FLW, Luderer R, Gabriels SHEJ, Westerink N, Lu R, de Wit PJGM, Joosten MHJ. 2000. A functional cloning strategy, based on a binary PVX-expression vector, to isolate HR-inducing cDNAs of plant pathogens. *Plant Journal* 24: 275–283.
- Thilliez GJA, Armstrong MR, Lim TY, Baker K, Jouet A, Ward B, van Oosterhout C, Jones JDG, Huitema E, Birch PRJ *et al.* 2019. Pathogen enrichment sequencing (PenSeq) enables population genomic studies in oomycetes. *New Phytologist* 221: 1634–1648.
- Van Weymers PSM, Baker K, Chen XW, Harrower B, Cooke DEL, Gilroy EM, Birch PRJ, Thilliez GJA, Lees AK, Lynott JS *et al.* 2016. Utilizing "Omic" technologies to identify and prioritize novel sources of resistance to the oomycete pathogen *Phytophthora infestans* in potato germplasm collections. *Frontiers in Plant Science* 7: 672.
- Vleeshouwers VG, Finkers R, Budding D, Visser M, Jacobs MM, van Berloo R, Pel M, Champouret N, Bakker E, Krenek P *et al.* 2011b. SolRgene: an online

- database to explore disease resistance genes in tuber-bearing *Solanum* species. *BMC Plant Biology* 11: 116.
- Vleeshouwers VG, Raffaele S, Vossen JH, Champouret N, Oliva R, Segretin ME, Rietman H, Cano LM, Lokossou A, Kessel G *et al.* 2011a. Understanding and exploiting late blight resistance in the age of effectors. *Annual review of Phytopathology* 49: 507–531.
- Wastie RL. 1991. Breeding for resistance. In: Ingram DS, Williams DS, eds. *Phytophthora infestans, the cause of late blight of potato. Advances in plant pathology, vol. 7*. London, UK & San Diego, CA, USA: Academic Press, 193–224.
- Witek K, Jupe F, Witek AI, Baker D, Clark MD, Jones JDG. 2016. Accelerated cloning of a potato late blight-resistance gene using RenSeq and SMRT sequencing. *Nature Biotechnology* 34: 656–660.
- Zhu J, Chen J, Gao F, Xu C, Wu H, Chen K, Si Z, Yan H, Zhang T. 2017. Rapid mapping and cloning of the *virescent-1* gene in cotton by bulked segregant analysis-next generation sequencing and virus-induced gene silencing strategies. *Journal of Experimental Botany* 68: 4125–4135.
- Zou G, Zhai G, Feng Q, Yan S, Wang A, Zhao Q, Shao J, Zhang Z, Zou J, Han B *et al.* 2012. Identification of QTLs for eight agronomically important traits using an ultra-high-density map based on SNPs generated from high-throughput sequencing in sorghum under contrasting photoperiods. *Journal of Experimental Botany* 63: 5451–5462.

## Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

**Notes S1** FASTA file of the 977 RLP RLK genes from DM genome used for generating the RLP/KSeq enrichment library.

**Notes S2** FASTA file of the 977 RLP RLK genes from M6 genome used for generating the RLP/KSeq enrichment library.

**Notes S3** Eighteen additional known RLP and RLK genes from Solanaceae species.

**Notes S4** The 2x bait library used in this study.

**Notes S5** Java script for calling the informative SNPs.

**Table S1** All tested *Solanum* genotypes for INF1 and SCR74 response by PVX agroinfection.

**Table S2** Primers used in this study.

**Table S3** SSR markers used in this study.

**Table S4** Summary of the SNP calling outputs under 5% mismatch criterion.

**Table S5** Using the Solyntus genome for mapping and SNP calling.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.