



Contents lists available at ScienceDirect

The Crop Journal

journal homepage: www.keaipublishing.com/en/journals/the-crop-journal/

Transcriptional profiling during infection of potato NLRs and *Phytophthora infestans* effectors using cDNA enrichment sequencing

Amanpreet Kaur^{a,b,c}, Vikrant Singh^a, Stephen Byrne^c, Miles Armstrong^b, Thomas M. Adams^a, Brian Harrower^a, Eleanor Gilroy^a, Ewen Mullins^{c,*}, Ingo Hein^{a,b,*}

^aThe James Hutton Institute, Invergowrie, Dundee DD25DA, Scotland, UK

^bUniversity of Dundee, Dundee DD14HN, Scotland, UK

^cCrop Research Centre, Teagasc, Oak Park, Carlow R93 XE12, Ireland

ARTICLE INFO

Article history:

Received 12 March 2024

Revised 20 August 2024

Accepted 18 September 2024

Available online 17 October 2024

Keywords:

RxLR effector

NLRs

Late blight

Potato

cDNA sequencing

RenSeq

PenSeq

RNAseq

ABSTRACT

An accurate assessment of host and pathogen gene expression during infection is critical for understanding the molecular aspects of host–pathogen interactions. Often, pathogen-derived transcripts are difficult to ascertain at early infection stages owing to the unfavourable transcript representation compared to the host genes. In this study, we compare two sequencing techniques, RNAseq and enrichment sequencing (RenSeq and PenSeq) of cDNA, to investigate gene expression patterns in the doubled monoploid potato (DM) infected with the late blight pathogen *Phytophthora infestans*. Our results reveal distinct advantages of cDNA RenSeq and PenSeq over traditional RNAseq in terms of target gene representation and transcriptional quantification at early infection stages. Throughout the infection time course, cDNA enrichment sequencing enables transcriptomic analyses for more targeted host and pathogen genes. For highly expressed genes that were sampled in parallel by both cDNA enrichment and RNAseq, a high level of concordance in expression profiles is observed, indicative of at least semi-quantitative gene expression representation following enrichment.

© 2024 Crop Science Society of China and Institute of Crop Science, CAAS. Production and hosting by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The potato late blight pathogen *Phytophthora infestans* causes about 20% of crop losses annually, resulting in a cost of \$6.7 billion globally [1]. In the absence of durable resistances, most currently grown commercial potato varieties are vulnerable to the pathogen and often protected through repeated chemical applications, which can exceed 15 sprays per crop annually in conditions that are favourable for pathogen proliferation [2]. Furthermore, the withdrawal of fungicides or the development of fungicide resistance in *P. infestans* can lead to the failure of widely used chemical control measures, which can take a decade to develop and be approved [3]. This has added to the cost and complexity of potato crop protection methods, whilst the cost of chemical-based late blight control has risen by about 25% in developed nations [4]. To identify additional control strategies, a more detailed understanding of

the molecular mechanisms involved in *P. infestans* pathogenicity is needed.

P. infestans is a hemibiotrophic pathogen that requires living host tissue in the initial biotrophic phase and which progresses to a necrotrophic phase leading to sporulation in a susceptible host [5]. Important features of *P. infestans* infection are haustoria which are indicative of a closely entwined relationship with the host [6]. These specialised invasion organs act as a junction to deliver apoplastic or cytoplasmic effectors which manipulate host metabolism and also suppress host defences [7]. In *P. infestans*, many cytoplasmic effectors have a canonical structure comprising an N-terminal signal peptide, an RxLR motif and a C-terminal effector domain; thus known as RxLRs [8]. As a defence mechanism, plants employ intracellular disease resistance genes that respond to effector function [9]. Known and recognised *P. infestans* RxLRs include, for example, *Avr1*, *Avr2*, *Avr3a*, *Avr3b*, *Avrblb1*, *Avrblb2*, *Avrvnt1*, *Avramr1*, and *Avramr3* (reviewed in [10]). These, in turn, are recognised by the host genes *R1*, *R2*, *R3a*, *R3b*, *Rpi-blb1*, *Rpi-blb2*, *Rpi-vnt1*, *Rpi-amr1*, and *Rpi-amr3*, all of which belong to the family of nucleotide-binding, leucine-rich-repeat (NLR) disease resistance genes (reviewed in [10]). In the Solanaceae plants tomato and

* Corresponding authors.

E-mail addresses: ewen.mullins@teagasc.ie (E. Mullins), Ingo.Hein@hutton.ac.uk (I. Hein).

potato, NLRs are often expressed constitutively at low levels, enabling plants to remain poised for pathogen detection, albeit tissue specific regulatory patterns have been observed [11].

The application of PenSeq to genomic DNA has provided insight into the allelic diversity of *P. infestans* effectors [12] and has also been described for environmental samples [13]. As effectors accounts for <1% of the *P. infestans* genome, targeting less than 0.5 Mb of the 240 Mb *P. infestans* genome, PenSeq reduces the genome complexity and enables evolutionary studies with higher statistical power [12]. However, genomic DNA-based PenSeq does not provide any information about effector expression levels. Therefore, cDNA PenSeq was developed, which allowed identification of 47 additional *P. infestans* RxLRs expressed during infection [14].

Similarly, RenSeq, when applied to genomic DNA has proven to be a versatile tool. In potatoes, RenSeq has been used to map and clone new resistances such as *Rpi-ver1*, *Rpi-blb4*, *Rpi-amr1*, and *Rpi-amr3* from the wild potato species *S. verrucosum*, *S. bulbocastanum* and *S. americanum*, respectively [14–17]. RenSeq has also been used to track the deployment of functional NLRs in cultivars (dRenSeq; [18]), and to represent full-length NLRs by the development of long-read enrichment sequencing with PacBio, SMRT-RenSeq [17]. Recently, RenSeq-based association studies have been reported in the form of AgRenSeq [19], SMRT-AgRenSeq [20] and SMRT-AgRenSeq-d, to rapidly identify candidate genes associated with elusive resistances in potato cultivars [21]. RenSeq has also been applied to cDNA to provide information about low expressed NLRs which otherwise cannot be identified by RNAseq [22].

Although the role of PenSeq and RenSeq in exploring effector diversity and the identification of novel resistances has been well established, it remains unknown if cDNA enrichment sequencing of effectors and NLRs can provide critical differential expression insight into early infection time points which often remain elusive in standard RNAseq studies. To address this question, we infected microshoots of the *Solanum tuberosum* group Phureja clone DM 1–3 516 R44 with the virulent *P. infestans* isolate W9928C. Extracted cDNA samples were subjected, in parallel, to RNAseq, PenSeq, and RenSeq-base sequencing. This is the first comparative study to establish the quantitative capabilities of RenSeq and PenSeq enrichment techniques.

2. Material and methods

2.1. Sample preparation

Tissue culture plants of the doubled monoploid potato *S. tuberosum* group Phureja DM 1–3 516 R44 (commonly known as DM) were maintained on MS20 medium (Murashige and Skoog medium, Duchefa, containing 20 g L⁻¹ sucrose; [23]). Plants were kept in a growth room at a light intensity of 110 μmol m⁻² s⁻¹, a temperature of 18 ± 2 °C, and a photoperiod of 16 h/8 h light/dark [23].

For the experiment, healthy three-week-old DM plantlets with fully expanded leaves were selected. *In vitro* shoots along with roots were gently removed from the media and dipped for one minute in a zoospore suspension of *P. infestans* isolate W9928C adjusted to 4 × 10⁶ spores mL⁻¹ [24]. Dip-inoculated microshoots were gently blotted on sterile paper towels and again planted in fresh MS20 media in vented tissue culture grade glass containers (Generon, UK). The infected plants were kept in darkness for 16 h and then incubated under the growth conditions mentioned above. The disease severity was recorded by counting the number of leaves showing disease symptom in 24-h intervals. The leaf samples from four independent replicates were collected after 0, 24, 48, 72 h post infection (hpi) and immediately immersed in liquid Nitrogen before storing at –70 °C for further processing (Fig. S1).

Additional samples were incubated for 96 h for visual inspection only.

2.2. RNA isolation and cDNA synthesis

From each individual replicate taken at 0, 24, 48 and 72 hpi, leaf samples were crushed to a fine powder and RNA was extracted combining TRI reagent (Sigma-Aldrich) and RNeasy kit (Qiagen, Germany) protocols. Specifically, 400 mg of ground sample was resuspended in 2 mL of TRI reagent and vortexed after addition of 10 μL β-mercaptoethanol. The slurry was left to stand at room temperature for 5 min before centrifugation at 10,000×g for 10 min at 4 °C. Chloroform was added to the supernatant (0.2 mL mL⁻¹), kept at room temperature for 5 min, and centrifuged. Isopropanol (0.5 mL mL⁻¹) was added to the aqueous phase and the solution was transferred to a QIA RNeasy spin column for binding and washing in RPE buffer twice. RNA was eluted in RNase free water (50 μL), and the integrity was checked using a Bioanalyzer 2100 (Agilent).

2.3. Sequencing

A pipeline was developed for parallel processing of the samples for RNAseq and enrichment sequencing (PenSeq and RenSeq; Fig. S1). Quality checked RNA (RIN value ≥ 8) was divided into two parts; one part was processed at the James Hutton Institute's Genomics facility for generating RNA sequencing libraries using the stranded Illumina mRNA Prep kit, and Integrated DNA technology (IDT) RNA unique dual UD Indices (Illumina) as recommended, with 100 ng total RNA per sample. Libraries were quality checked on a fluorimeter (Qubit) and on a Bioanalyzer 2100 prior to pooling equimolar amounts before sequencing. Sequencing was conducted on a NextSeq 2000 (Illumina) sequencer at loading concentration of 750 pmol L⁻¹ using a P3 200 kit, generating paired-end 100 bp reads.

In parallel, second strand cDNA was synthesised from the RNA using both oligo(dT) and random hexamer primers following the manufacturer's guideline (Superscript IV, Thermo Fisher Scientific). Individually indexed libraries were prepared for PenSeq and RenSeq by combining equimolar amounts of cDNA from each time point and replicate. Enrichment sequencing was conducted using the myBaits v5.02 protocol in combination with the PenSeq and RenSeq bait library [12,25] at Arbor BioSciences, MI, USA and the captures were pooled for sequencing on an Illumina NovaSeq 6000 platform on a partial S4 PE150 lane.

2.4. Data analysis

Reads obtained from RNAseq and cDNA enrichment sequencing (RenSeq/PenSeq) were subjected to quality control and adapter trimming using fastp [26] at 99% base call accuracy (default settings allowing phred quality 20). Trimmed reads were mapped to the host (*S. tuberosum* group Phureja clone DM1–3 516 R44 v6.1) and pathogen (*P. infestans* T30–4) reference genomes using a bowtie based splice aware mapper, HISAT2, under stringent conditions allowing 1% mismatch rate (–score-min L,-0.06,-0.06; [27]). The reads were also quantified against reference transcriptomes with Salmon running in alignment mode for quantification of the bam files. The quantification files were analysed for differential gene expression using the 3D RNAseq App [28]. Read counts and transcript per million reads (TPMs) were established using tximport R package version 1.10.0 and lengthScaledTPM method [29] with inputs of transcript quantifications from Salmon [30]. Low expressed transcripts and genes were filtered based on analysing

the data mean–variance trend. The Trimmed Means of M values (TMM) method was used to normalise the gene and transcript read counts to \log_2 CPM [31] and a principal component analysis (PCA) was performed to identify any batch effects. The limma R package was used for gene and transcript expression comparison [32,33]. The \log_2 fold change (FC) of gene and transcript abundance was calculated and a *t*-test was used to determine significance of expression changes. *P*-values of multiple testing were adjusted with the Benjamini-Hochberg (BH) procedure to correct false discovery rate (FDR < 0.05; [34]), according to the built-in protocol of 3D-RNAseq app [28]. A gene was considered significantly differential expressed if it had an adjusted *P*-value of < 0.01 and \log_2 FC \geq 1.

To filter NLRs from the expressed genes, NLR annotation was conducted by running NLR tracker [35] on coding sequence (CDS) of DM v6.1. The sequences of retrieved NLRs were subjected to a BLAST analysis with the previously reported 755 NLRs by [25]. Only 100% matching sequences with available reference gene IDs were retained for the expression analysis yielding 583 NLRs.

GO enrichment analysis of expressed genes was performed using the 'Database for annotation, visualisation and integrated discovery' (DAVID) functional analysis tool [36]. GO terms with *P* value < 0.05 were considered significantly associated (enriched) for a pathway.

To calculate the sequencing coverage of NLRs and pathogen targets (RxLR and non-RxLR), bed files generated from genome assemblies of DM v4.03 [25] and *P. infestans* T30-4 [12] with coordinates of 755 NLRs and 579 effectors (RxLRs and non-RxLRs) were used as a reference. Sequencing depth and coverage of target genes were calculated using bedtools v2.31.0 and compared between RNAseq and cDNA enrichment (RenSeq and PenSeq). Data were further processed to establish correlations among the datasets and time points using Pearson's correlation. Before analysing the TPM values were transformed using \log_{10} (TPM + 1).

3. Results

Microshoots of DM infected with *P. infestans* isolate W9928C showed progressive late blight disease symptoms (Fig. 1). Visible disease symptoms developed at 48 h post infection (hpi) which

coincides with the transition between the biotrophic and necrotrophic phases. At 72 hpi, most of the leaves showed lesions which proceeded to complete necrosis at 96 hpi. Thus, 96 hpi samples were excluded from the sequence-based analyses. Leaf samples from the infected plants were collected at 0, 24, 48 and 72 hpi for RNAseq and RenSeq/PenSeq-based cDNA enrichment sequencing, respectively. Following Illumina sequencing, 1.13 billion paired end reads (100 bp) showing an overall Q30 > 94% were obtained for RNAseq (Table S1). For PenSeq and RenSeq-based enrichment sequencing, a total of 369 million reads passed the filter (PF; Q30) with read counts ranging between 19 million to 28 million per sample (Table S1).

Following QC and adaptor trimming, more than 832 million reads from RNAseq and over 363 million reads from enrichment sequencing were retained (Table S1). These reads were mapped to the potato (*S. tuberosum* group Phureja DM 1–3 516 R44 v6.1) and *P. infestans* (T30-4) reference genomes at 1% mismatch rates. For RNAseq, 749 million reads out of the 832 million reads (89.98%) mapped to the potato reference genome whereas only 1.6 million (0.2%) mapped to *P. infestans* reference genome (Table 1). For the indexed RenSeq and PenSeq enrichment sequencing, a total of 363 million reads were mapped to reference genomes, out of which 280 million (76.99%) mapped to potato (RenSeq) and 8.3 million (2.28%) mapped to *P. infestans* (PenSeq). Importantly, in the RNAseq analysis the percentage of pathogen derived reads did not increase during the infection time course whereas the PenSeq analysis displayed a steady increase of detectable pathogen transcripts from 1.04% at 0 hpi, 1.47% at 24 hpi, 2.52% at 48 hpi and 4.04% at 72 hpi (Table 1).

3.1. Targeted genes are highly represented by RenSeq and PenSeq reads

The mapped RNAseq, RenSeq and PenSeq reads were filtered for the 579 target genes in *P. infestans* (including 438 RxLR and 141 non-RxLR effectors; [12]) and 755 potato NLRs [25] using bedtools. The number of *P. infestans* targets and NLRs expressed in RNAseq and enrichment (RenSeq and PenSeq) datasets were evaluated (Table S2). Out of a total of 579 *P. infestans* effectors, an average of 316 genes (289 at 0 hpi, 332 at 24 hpi, 331 at 48 hpi and 312 at 72 hpi) were represented by cDNA-PenSeq reads indicative of

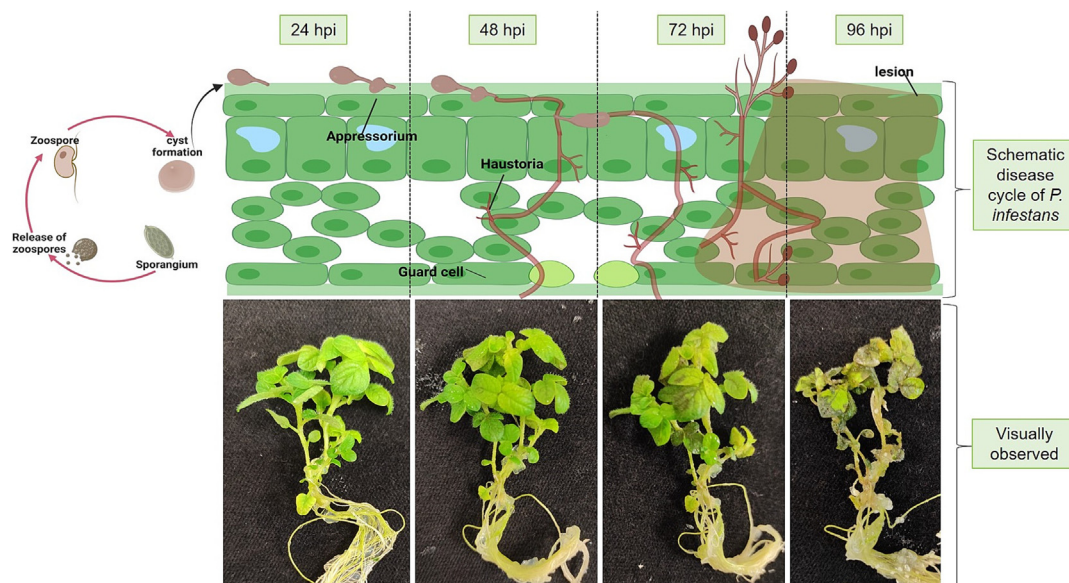


Fig. 1. Diagrammatic representation of various stages of *Phytophthora infestans* life cycles and disease progression on *in vitro* *Solanum tuberosum* group Phureja DM 1–3 516 R44 (commonly known as DM) plants.

Table 1
Number of RNAseq, RenSeq and PenSeq reads mapped to the host and pathogen genomes.

Sequencing techniques	Total reads	Reads mapped to DM v6.1	% reads mapped to DM v6.1	Reads mapped to <i>P. infestans</i> T30-4	% reads mapped to <i>P. infestans</i> T30-4
RNAseq					
0	192,017,649	190,972,662	99.4250	527,619	0.27
24	225,017,104	222,731,572	98.9925	258,190	0.11
48	214,620,790	205,952,655	95.9950	429,980	0.20
72	201,262,977	129,825,714	64.9825	483,111	0.24
Total	832,918,520	749,482,603	89.9800*	1,698,900	0.21*
RenSeq					
0	95,324,230	80,302,072	84.2400		
24	85,202,507	73,723,461	86.5300		
48	88,339,763	75,416,568	85.3700		
72	95,095,897	50,796,658	53.4200		
Total	363,962,397	280,238,759	76.9900*		
PenSeq					
0	95,324,230			992,390	1.04
24	85,202,507			1,248,629	1.47
48	88,339,763			2,227,074	2.52
72	95,095,897			3,842,593	4.04
Total	363,962,397			8,310,686	2.27*

*average across the time points. *Solanum tuberosum* group Phureja DM 1–3 516 R44–v6.1 (DM v6.1) and *Phytophthora infestans* strain T30-4 genomes were used as references and reads obtained at different infection timepoints were mapped using HISAT2 [27]. 0, 24, 48, and 72 indicate hours post infection.

their expression, whereas an average of 167 *P. infestans* target genes (146 at 0 hpi, 181 at 24 hpi, 192 at 48 hpi and 150 at 72 hpi) were detected by RNAseq reads (Fig. 2A; Table S2). Thus, a two-fold increase in the number of expressed target genes was observed in PenSeq in comparison with RNAseq. Importantly, all genes identified by RNAseq were also represented by cDNA-based PenSeq. A comparison of the number of effectors (RxLRs and non-RxLRs) identified by PenSeq and RNAseq at different time points revealed that enrichment sequencing identified 129, 156, 186 and 193 unique RxLRs and 21, 16, 19 and 20 non-RxLRs at 0 hpi, 24 hpi, 48 hpi and 72 hpi, respectively (Fig. 2B; Table S3). Among these genes, 188 RxLRs and 13 non-RxLRs were uniquely present in these datasets (Fig. S2A, B; Table S4). Additionally, RenSeq identified expression of 48 more NLRs compared to RNAseq (Fig. S2C).

To evaluate NLR expression, the reads were mapped to the lat-est transcript assembly of potato (DM v6.1). NLR sequences identi-

fied using NLR tracker were subjected to BLAST using the previously reported 755 NLRs in DM v4.3 [25] and we could determine the position of 583 NLRs with existing reference transcript IDs. Out of 583 NLRs, 56 belong to the helper and 527 belong to the sensor NLR groups (Fig. S2b). Among these indexed NLRs, an average of 571 NLRs was detectable in cDNA RenSeq across all the infection points whereas only 429 NLRs were expressed according to RNAseq data (Fig. 2C; Table S5). Critically, all genes identified as expressed by RNAseq were also detectable by cDNA RenSeq.

3.2. Enriched mapping resulted in higher expression values of the target genes.

The length scaled transcript per million (TPM) values of the target genes in *P. infestans* and potato were highest in the enriched datasets (PenSeq and RenSeq) in comparison to RNAseq

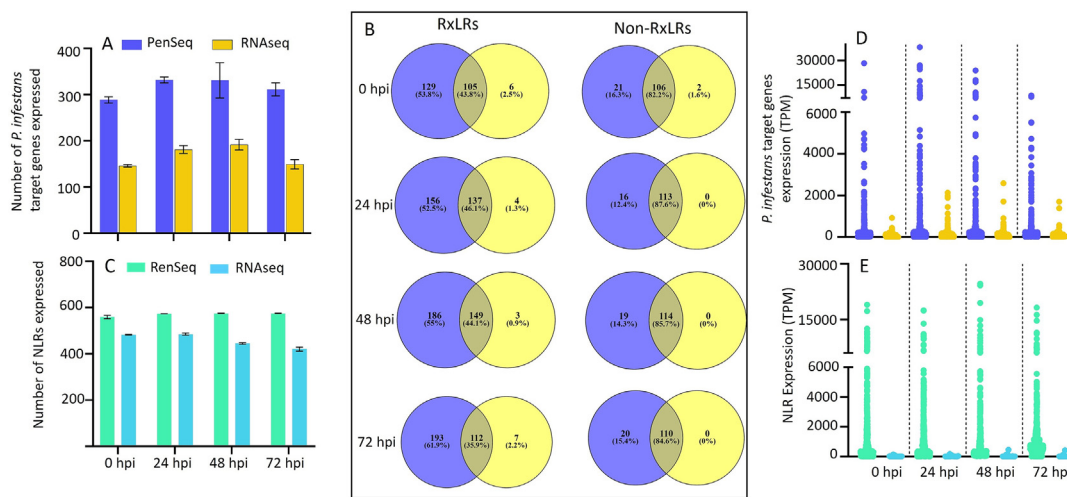


Fig. 2. A comparison between RNAseq and PenSeq enrichment sequencing for the expression of pathogen effectors (RxLR and non-RxLR effectors) at different time points of infection. (A) Number of effectors expressed in the RNAseq and PenSeq datasets following mapping to the *Phytophthora infestans* transcriptome with high stringency, allowing a 1% mismatch rate. (B) Venn diagrams demonstrating the number of effectors (RxLRs and non-RxLRs) identified in the PenSeq dataset but not in RNAseq. The PITG numbers of unique RxLRs and non-RxLRs are detailed in Table S3. (C) Number of NLRs expressed in the RNAseq and RenSeq datasets following mapping to the *Solanum tuberosum* group Phureja DM 1–3 516 R44–v6.1 (DM v6.1) transcriptome at a 1% mismatch rate. (D, E) A comparison between the expression values (Lengthscaled Transcripts Per Million; TPM values) of effectors and NLRs between RNAseq and enrichment techniques (PenSeq and RenSeq).

(Figs. 2D, E, S2a D, E). For *P. infestans*, an average of 64.94% of expressed effectors (RxLRs and non-RxLRs) yielded TPM values greater than 100 in the PenSeq dataset, whereas this was limited to 21.59% in RNAseq derived reads (Fig. 2D; Tables S2, S6). A similar trend was observed for potato, where an average of 72.22% of expressed NLRs have TPM values greater than 100 in the RenSeq-based analysis. In contrast, only 0.19% expressed NLRs in RNAseq reads showed TPM values greater than 100 (Fig. 2E; Tables S5, S7).

3.3. High correlation between RNAseq and enrichment datasets validates concordance between RNAseq and enrichment techniques

To assess if cDNA targeted enrichment sequencing retains quantitative transcriptional properties, expression patterns of genes that were detectable in RNAseq and PenSeq/RenSeq-derived datasets were compared. Across the infection time points, 579 pathogen effector genes and 583 host NLRs were used to calculate a Pearson correlation coefficient (Fig. S3). The computed correlation coefficient for pathogen target genes expressed in four replicates was in the range of 0.84–0.98 and 0.78–0.87 within and between the PenSeq and RNAseq datasets, respectively (Fig. S3A; Table S8). A similar high correlation was observed between RenSeq and RNAseq datasets (Fig. S3B; Table S8). This shows a strong consensus between the replicates as well as the techniques.

A quantitative comparison of pathogen target gene expression levels [$\log_{10}(\text{TPM} + 1)$] between PenSeq/RenSeq and RNAseq indicated a significant correlation among the techniques (Fig. S3C, D).

The expression profiles of pathogen target genes and NLRs as detected by RenSeq and PenSeq were subjected to hierarchical clustering to group the genes into clusters with similar expression patterns (Fig. 3). The expression values of the selected genes from RenSeq and PenSeq when compared with RNAseq showed similar expression patterns (Fig. 3). To highlight our observations, the known effectors linked with the biotrophic phase [PITG_04314 (*PexRD24*), PITG_14371 (*Avr3a*), PITG_20300 (*Avrblb2*), PITG_21388 (*Avrblb1*), PITG_03192, PITG_04089] and the closest homologs of functional resistance genes [Soltu.DM.04G006260 [*Rpi-blb3*], Soltu.DM.04G006450 [*R2*], Soltu.DM.04G008180 [*Rpi-amr3*], Soltu.DM.05G005840 [*R1*], Soltu.DM.06G001240 [*Rpi-blb2*], Soltu.DM.08G020910 [*Rpi-blb1*], Soltu.DM.09G029510 [*Rpi-vnt1*], Soltu.DM.09G029560 [*R9a*], Soltu.DM.09G030280, [*R8*] Soltu.DM.11G024990 [*R3a*)] were compared in their expression values in PenSeq, RenSeq and RNAseq datasets. The expression trends of the genes at different infection time points were very similar between the techniques, although the genes from cDNA enrichment showed higher values (Fig. 3).

To further validate our observations, previously transcriptionally characterised *P. infestans* marker genes were specifically queried in the data sets. These include the cellulose synthase gene (*CesA4*; PITG_16984) associated with cyst germination and appressorium formation [37], haustorial membrane protein (*HMP1*; PITG_00375) associated with biotrophy [38], and the Necrosis inducing Phytophthora Protein 1 (*NPP1*; PITG_00957), a necrotrophy marker which shows a transition between the biotrophic and necrotrophic phases of infection ([39]; Fig. 4A). For the *CesA4* gene (PITG_16984), expression was highest at 24 hpi with a TPM value of 865.17 in the case of PenSeq and 46.93 in RNAseq (Table S6). The TPM values decrease to 620.17 at 48 hpi in the case of PenSeq and to 28.68 in RNAseq. These observations are in accordance with [37] which showed that there is an increase in *CesA* gene expression during the initial phases of infection. For *HMP1*, cDNA PenSeq expression was the highest at 24 hpi with a TPM value of 1335.51. This decreased to 754.01 at 48 hpi and to 531.56 at 72 hpi (Table S6). The expression profile of *HMP1*, as detected by PenSeq, coincides with the results of the RNAseq anal-

ysis where a decrease in the TPM value from 85.22 at 24 hpi to 34.95 at 72 hpi was observed (Table S6). For *NPP1*, PenSeq and RNAseq showed that the expression declined after the onset of the necrotrophic phase, which in our experiment occurred at around 48 hpi. These observations are in line with previous studies which also showed that the expression level of *NPP1* is highest at the beginning of necrotrophic phase [40]. As the difference between the expression values was very high, the TPM values were transformed using $\log_{10}(\text{TPM} + 1)$ for plotting the graphs to compare gene expression between PenSeq and RNAseq.

3.4. PenSeq provides more information about expressed genes and biological processes than RNAseq

To better understand the overall changes in effector expression during the different phases of infection, genes were clustered into functional components and pathways using DAVID bioinformatics resources. Of the 579 effectors analysed in the study, 575 were annotated in the Gene Ontology (GO) database GO term analysis. Expressed genes from the PenSeq dataset at different time points clustered into 44, 51, 51, and 51 major groups of enriched GO terms. These include biological process: 10, 15, 15, 15; cellular components: 12, 12, 12, 12 and molecular functions: 20, 24, 24, 24 at 0, 24, 48 and 72 hpi, respectively (Fig. S4). The GO terms of expressed genes in the RNAseq dataset were clustered into 35, 45, 36, 35 major groups, comprised of biological process (8, 13, 9, 8), cellular component (10, 12, 12, 11), and molecular function (17, 20, 15, 16) at 0, 24, 48 and 72 hpi, respectively (Table S9). In the biological process category, lipid, cellulose and amino acid biosynthesis were predominant subcategories, while in the molecular function category, hydrolase activity, membrane transporter, and peptidase inhibition were enriched in both PenSeq and RNAseq datasets (Fig. S4).

To further evaluate whether PenSeq was able to provide additional insight into the *P. infestans* – potato interaction, the expression levels of non-RxLRs genes and a known RxLRs identified exclusively in the PenSeq dataset (from at least at one timepoint) (Figs. 4B, S2A; Table S10) were compared with RNAseq (Fig. 4B). This analysis revealed that genes involved in infection initiation processes such as energy generation through carbohydrate metabolism (PITG_01112), host cell penetration and sporulation (PITG_03540, PITG_05902, PITG_13473, PITG_19286), inhibition of plant proteases (PITG_22920, PITG_22950, PITG_13292, PITG_01369, PITG_23012, PITG_22692, PITG_23195), and suppression of host defence [PITG_13940, PITG_13956, PITG_23008, PITG_21949, PITG_21949, PITG_07387 (*Avr4*), PITG_14737, PITG_21422, PITG_20300 (*Avrblb2*)], were exclusively identified by PenSeq, particularly at early infection time point (0 hpi) (Fig. 4B; Table S6).

Overall, these results show that enrichment techniques retain the quantitative expression nature in the data and allow the detection of a greater number of low abundant transcripts compared to RNAseq. Not surprisingly, where genes were only detectable at low expression levels in the RNAseq dataset, the correlation to cDNA enriched profiles was less robust.

As mentioned previously, host and pathogen genes that were represented in the RenSeq and PenSeq bait libraries and identified as expressed by normal RNAseq, were also identified by RenSeq and PenSeq. However, in terms of differential expression, PenSeq identified almost twice as many differentially expressed effectors in comparison to RNAseq in all time points (Fig. 5A; Table S11). A modulation of expression of 187, 201 and 159 pathogen target genes (average of 4 replicates) at 24, 48 and 72 hpi, was identified by PenSeq, while RNAseq identified 87, 84 and 66 target genes only. Critically, most of the differentially expressed genes (DEGs) identified by RNAseq were also identified by PenSeq and for

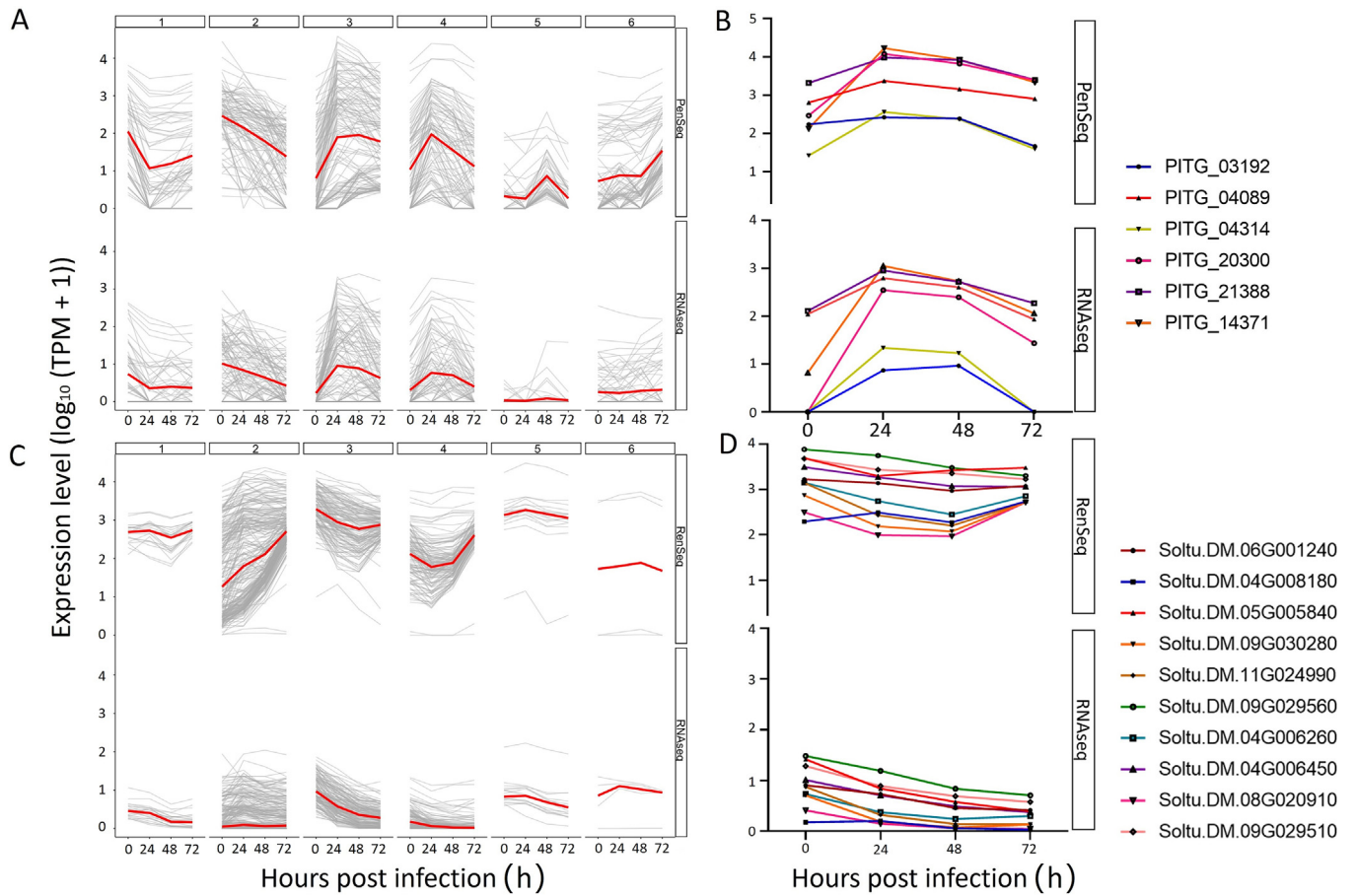


Fig. 3. A comparison between expression profiles of effectors (RxLR and non-RxLR) and NLR genes. (A) Expression patterns of effectors in the PenSeq and RNASeq datasets. (B) A subset of effectors linked with the biotrophic phase of infection was plotted to compare expression levels. (C) Expression patterns of NLRs in the RenSeq and RNASeq datasets. (D) A subset of NLRs (homologues of commercially important *R* genes) was plotted. Transcripts Per Million (TPM) values were transformed to $\log_{10}(\text{TPM} + 1)$ before plotting.

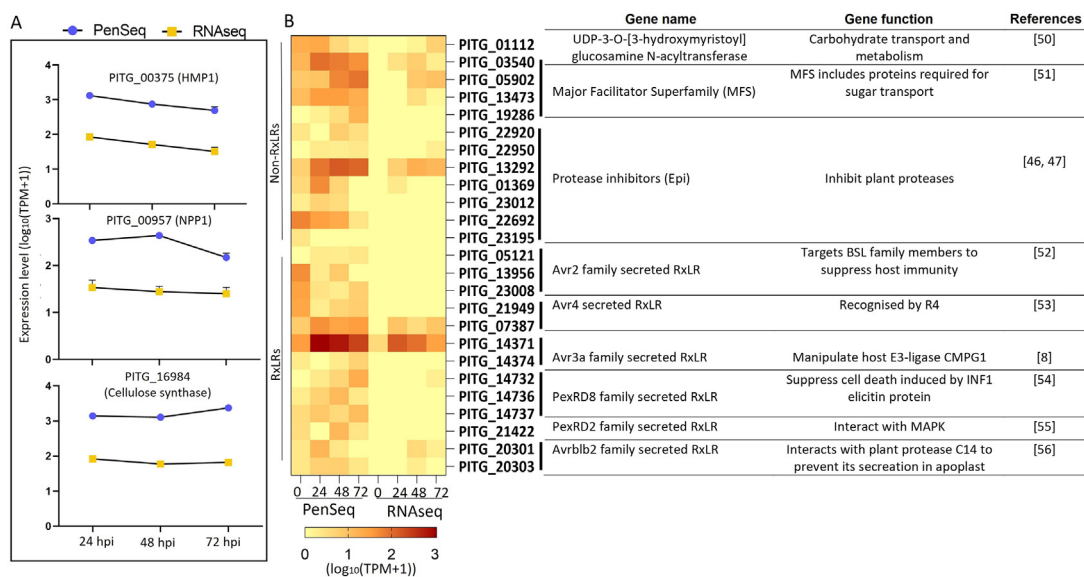


Fig. 4. Expression profiles of selected genes and their role in pathogenicity. (A) Expression profiles of marker genes for infection initiation through appressorium formation (cellulose synthase *CesA4*), biotrophy (Haustorium Membrane Protein; HMP1), and necrotrophy (Necrosis-Inducing Phytophthora Protein; NPP1-like) as detected by PenSeq and RNASeq during different infection time points. (B) Heatmap of selected effectors (expressed only in the PenSeq dataset, primarily during the initial stages of infection) along with their function [8,46,47,50–56]. TPM values were normalised to $\log_{10}(\text{TPM} + 1)$ for plotting expression profiles. The TPM values of the effectors used to create the heatmap are provided in Table S10.

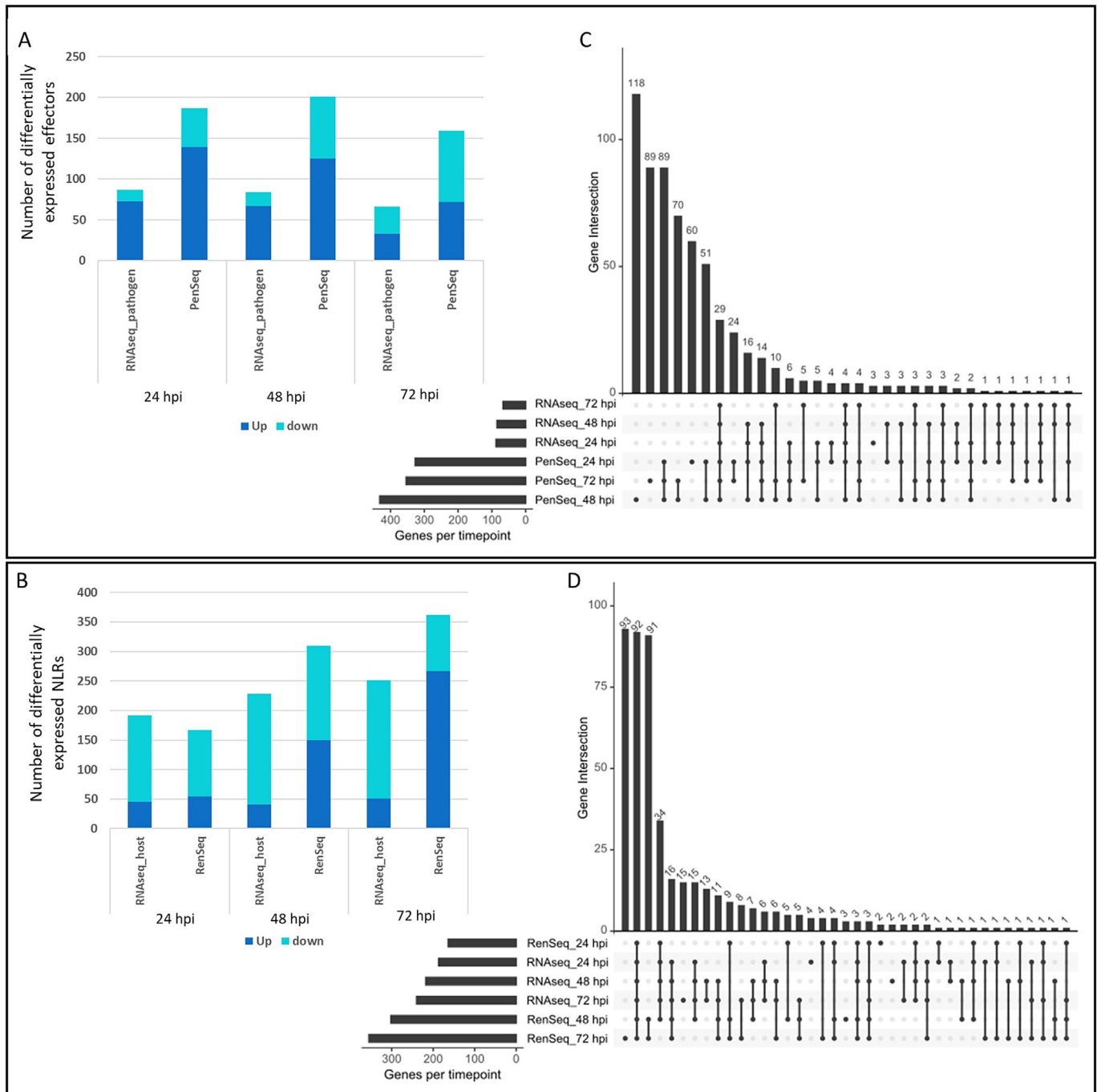


Fig. 5. A comparison between the differential expression of effectors (RxLR and non-RxLR) and NLRs in RNAseq and PenSeq/RenSeq datasets. The \log_2 Fold change (FC) in effector abundance was calculated. A *t*-test was used to determine the significance of the expression changes. A gene was considered significantly differentially expressed if it had an adjusted *P*-value < 0.01 and $\log_2FC \geq 1$. (A, B) Variation in the number of effectors and NLRs differentially expressed between the RNAseq and PenSeq/RenSeq datasets at different infection time points. (C, D) An overlap between differential effector/NLR expression identified by target enrichment (PenSeq/RenSeq) and RNAseq throughout infection is presented as an UpSet plot.

99.64% of the genes (upregulated or downregulated) the directions of expression changes were identical (Table S11). Since PenSeq was able to identify 43.33% more DEGs in comparison to RNAseq, an average overlap of 39.73% was found between the two techniques (Fig. 5C; Table 2).

Similarly, RenSeq identified 163, 301 and 355 differentially expressed NLRs at 24, 48 and 72 hpi, whereas RNAseq identified 192, 228 and 251 differentially expressed NLRs at these time points (Fig. 5B; Tables S2, S12). In total, 277 NLRs with significantly modulated expression were identified by RNAseq and the RenSeq datasets with expression concordance. Further, 207 and 43 NLRs were

predicted to be differentially expressed exclusively by RenSeq and RNAseq respectively (Fig. 5D).

3.5. cDNA enrichments yield higher target gene coverage

Target gene sequence representation was calculated for the expressed effectors and NLRs as a percentage of gene length covered by the reads (Fig. S5). In line with the dRenSeq-type analyses, cDNA-derived RenSeq and PenSeq reads showed the highest average number of target genes with 100% coverage in comparison to RNAseq reads (Table 3). During infection, out of 289, 332, 331

Table 2
Concordance of differentially expressed (DE) pathogen target genes (RxLR and non-RxLRs) and NLRs between RNAseq and enrichment (PenSeq and RenSeq) approaches.

Hours post infection (h)	DE Pathogen target genes identified in PenSeq	Overlapping DEGs with RNASeq	% of overlapping DE pathogen target genes	DE NLRs identified in RenSeq	Overlapping DE NLRs with RNASeq	% of overlapping DE NLRs
24	187	83	44.39	163	138	84.66
48	201	77	38.31	301	183	60.80
72	159	58	36.48	355	146	41.13
Average	182	73	39.73	273	156	62.19

and 312 (average of 4 replicates) effectors expressed at 0, 24, 48 and 72 hpi; a complete (100%) coverage was observed for 99, 132, 100 and 58 effectors (average of 4 replicates) in the PenSeq data, respectively. This number was limited to 9, 15, 19 and 12 effectors (average of 4 replicates) with 100% coverage at various infection time points for RNAseq (Table 3). Similarly, RenSeq data showed 100% coverage in almost twice the number of NLRs detectable by RNAseq during all the time points (Table 3). Complete coverage was observed for 595, 607, 667 and 690 NLRs expressed (average of 4 replicates) at 0, 24, 48 and 72 hpi in RenSeq dataset, whereas only 378, 356, 291 and 237 NLRs (average of 4 replicates) were 100% covered with RNAseq reads (Table 3). Similar to the expression data, cDNA enrichment sequencing provides a more robust analysis of the transcription across larger parts of the genes, including at early infection time points.

4. Discussion

Transcriptomics is a powerful method to enhance our understanding of molecular events occurring during infection. Although parallel expression studies of the host and pathogen during disease progression have provided insights into plant-pathogen interactions, the abundance of host biomass, and thus expressed genes during the early phases of infection, often limit detection and expression profiling of pathogen genes [41,42]. To overcome this limitation, we introduce cDNA-based target enrichment sequencing, RenSeq and PenSeq, for parallel NLRome and effectorome transcriptome studies during potato late blight infection. The use of cDNA RenSeq and PenSeq has been described previously to aid the annotation of expressed host and pathogen genes [14,22]. Here, we further assess if cDNA enrichment technologies enable quantitative transcriptomic analyses at early and late infection time-points.

Targeted enrichment sequencing in the form of RenSeq and PenSeq is facilitated through an excess of RNA-derived probes that drive the hybridisation to target sequences, which are represented through cDNA in this study. It is this molar excess of baits that is key to retaining the quantitative nature in the gene expression profiles post enrichment sequencing. Our comparative analysis between PenSeq and RNAseq reflects the hemibiotrophic nature of the infection. Previously utilised marker genes for the late blight infection progression such as *HMP1*, *NPP1* and known *Avr* genes [38,39,43] with a biotrophic expression profile, revealed concor-

dance in their transcriptional behaviour for both RNAseq and PenSeq analyses (Fig. 4A; Table S6).

In our study, a spore suspension of *P. infestans* was used to infect the *in-vitro* DM plantlets. It is well documented that zoospores, upon contact with host, encyst and germinate, forming an appressorium at the tip of the germ tube which then develops into a penetration peg [44]. Further, *P. infestans* develops haustoria that exhibit a close entwinement with the host cells and through which most of the molecule exchange occurs [6]. GO term analysis showed enrichment of GO terms such as GO:0035879, GO:0072488, and GO:0090481 throughout the PenSeq dataset (Fig. S4). These GO terms indicate involvement of plasma membrane lactate transport, ammonium transmembrane transport and sugar transport pathways from the beginning of the infection process. This information was absent from the RNAseq dataset, where only GO:0090481 (sugar transport) was enriched at 0 hpi, while GO:0072488 (ammonium transmembrane transport) was enriched at 24 hpi and GO:0035879 (plasma membrane lactate transport) was found only at 48 hpi. Separately, Grenville-Briggs et al. [37,45] have shown the elevation of amino acid and carbohydrate biosynthesis during infection. In agreement with these reports, our study showed the enrichment of GO:0009245, GO:0030244, and GO:1901607, which are involved in lipid, cellulose and amino acid biosynthesis pathways in both PenSeq and RNAseq datasets (Fig. S4; Table S9).

During infection, plants release various proteases (e.g., P69B) into the apoplast to inhibit pathogenic proteins [46]. In response, and to combat this defence mechanism, *P. infestans* secretes extracellular protease inhibitors [47]. Among the 17 protease inhibitors identified in our study, eight (PITG_22920, PITG_22692_PITG_23195, PITG_23012, PITG_05430, PITG_16827, PITG_23032, PITG_09175) were identified only by PenSeq (Fig. 4B; Table S6). In addition to non-RxLRs, 11 known RxLRs or their homologues were exclusively represented in PenSeq and were not identified in the RNAseq dataset (Fig. 4B). The capacity of PenSeq to provide additional information of molecular interactions during the early phases of plant-pathogen interactions is a distinct advantage over conventional RNAseq.

In terms of biology, PenSeq and RNAseq revealed that the progression of infection in the tissue culture plant DM is very rapid, which is in line with the highly susceptible phenotype of DM and the lack of functional NLRs to control late blight [48]. This susceptibility is further promoted using tissue culture derived plants.

Table 3
Comparison between RNAseq and enrichment techniques for coverage of pathogen effector genes (RxLRs and non-RxLRs) and host NLRs.

Hours post infection (h)	Average number of effectors with coverage above 0		Average number of effectors with 100% coverage		Average number of NLRs with coverage above 0		Average number of NLRs with 100% coverage	
	PenSeq	RNAseq	PenSeq	RNAseq	RenSeq	RNAseq	RenSeq	RNAseq
0	289	146	99	9	705	630	595	378
24	332	181	132	15	705	637	607	356
48	331	192	100	19	705	612	667	291
72	312	150	58	12	705	605	690	237

Coverage values for NLRs were calculated using all the 755 reported NLRs identified in Jupe et al. 2013 [25] (according to DM v4.03). Average was taken for 4 replicates. Values for number of genes were round to nearest whole numbers.

Consequently, in our replicated experiments, the visual transition from biotrophy to necrotrophy occurred at approximately 48 h. These observations are in line with a previous report on *P. infestans* – host interactions [45]. No significant disease symptoms were observable at 24 hpi whereas later time points at 48 and 72 hpi revealed progressive tissue collapse until complete plant death at 96 hpi (Fig. 1).

In line with this, cDNA PenSeq and RNAseq transcriptomic analyses showed that during the biotrophic phase the expression of the Haustorial Membrane Protein (*HMP1*; PITG_00375) was highest at 24 hpi (Fig. 4A; Table S6). Similarly, *NPP1* (PITG_00957), a necrotrophic marker which shows the transition between the biotrophic and necrotrophic phases of infection [39], was highest at 48 hpi and after that it reduced significantly. This expression profile is in line with the observation made by [39], where in infections of tomato the expression of *NPP1* decreased after 48 hpi.

The high-level of expression correlation for transcripts that are detectable by RenSeq, PenSeq and RNAseq provides strong evidence that quantitative information is retained through the enrichment sequencing process. The correlation between RNAseq and enrichment datasets ranged between $r = 0.78$ – 0.87 , for PenSeq and $r = 0.62$ – 0.86 for RenSeq, both within and between replicates (Fig. S3; Table S8). The hierarchical clustering analysis further confirmed the similarity of expression patterns between PenSeq, RenSeq, and RNAseq data, while highlighting genes with no or low expression in the latter (Fig. 3). Importantly, any gene identified as being expressed by the RNAseq analysis was also identified in the RenSeq or PenSeq analyses. However, RenSeq and PenSeq identified additional host and pathogen genes, respectively, that were below the threshold of expression detection in the RNAseq analysis. Notably, PenSeq was able to identify an average of 316 expressed genes throughout the time course, which is a two-fold increase compared to RNAseq, which identifies an average of 167 expressed genes only (Fig. 2A; Table S2). Targeted enrichment sequencing ensures minimal off-target sequencing and provides sufficient read depth and sequence coverage for identification of low expressed target genes [49]. In the present study, we observed that 89.98% of RNAseq reads and 76.99% of RenSeq reads map to the potato reference transcriptome. For PenSeq, averaged across all time points, an almost twelve-fold increase in the percentage of reads mapped to the *P. infestans* reference genome was observed in comparison with RNAseq. It is noteworthy that whilst in the RNAseq dataset, the reads mapped to the reference transcriptome remains almost constant throughout the infection time points at 0.2%, an increase in mapped reads from 1.04% at 0 hpi to 4.04% at 72 hpi was observed in PenSeq dataset (Table 1). This demonstrates the ability of PenSeq to capture transcripts from the pathogen during infection which enables the identification and quantification of expressed genes, many of which elude detection by RNAseq. In addition, the percentage coverage of targeted genes was more comprehensive for RenSeq and PenSeq compared to RNAseq, achieving complete coverage for about 80% of NLRs and 40% of effectors expressed at 24 hpi (Fig. S5; Table 3).

Although limited to targeted genes of interest, our study highlights the advantages of cDNA-based enrichment sequencing techniques, PenSeq and RenSeq compared to RNAseq, for not only ascertaining gene expression, but also their transcriptional quantification. Specifically, these techniques offer insight into the detection and transcriptional profile of low expression genes with increased transcript coverage. An in-depth understanding of the molecular events occurring during the initial stages of infection is important for informing disease control strategies. Using enrichment sequencing for parallel host-pathogen gene expression studies facilitates the analysis of expressed effectors (RxLRs and non-RxLRs) and resistance mechanisms during infection. In this study, we intentionally used DM as a model plant due to its high disease

susceptibility, which aids in identify effector expression required for compatible interactions. Further, the genome assembly of DM is of a high quality as are, NLR gene models. However, our approach can also be replicated for other potato varieties or cultivars to gain further insights into host resistances and effector diversity in prevalent pathogen populations. This will be a key strategy to support potato breeding programs aimed at stacking *R* genes to deliver durable resistance to *P. infestans*.

CRedit authorship contribution statement

Amanpreet Kaur: Data curation, Formal analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing. **Vikrant Singh:** Data curation, Investigation, Writing – original draft. **Stephen Byrne:** Formal analysis. **Miles Armstrong:** Investigation, Methodology. **Thomas M. Adams:** Writing – review & editing. **Brian Harrower:** Investigation, Methodology. **Eleanor Gilroy:** Writing – review & editing. **Ewen Mullins:** Funding acquisition, Project administration, Supervision, Writing – review & editing. **Ingo Hein:** Conceptualization, Funding acquisition, Project administration, Resources, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Rural & Environment Science & Analytical Services (RESAS) Division of the Scottish Government through project JHI-B1-1, the Biotechnology and Biological Sciences Research Council (BBSRC) through awards BB/S015663/1; BB/X009068/1, and through a Research Leaders 2025 fellowship funded by European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement no. 754380. The authors acknowledge the Research/Scientific Computing teams at The James Hutton Institute and NIAB for providing computational resources and technical support for the "UK's Crop Diversity Bioinformatics HPC" (BBSRC grant BB/S019669/1), use of which has contributed to the results reported within this paper.

Appendix A. Supplementary data

Supplementary data for this article can be found online at <https://doi.org/10.1016/j.cj.2024.09.013>.

References

- [1] M. Nowicki, M.R. Foolad, M. Nowakowska, E.U. Kozik, Potato and Tomato late blight caused by *Phytophthora infestans*: an overview of pathology and resistance breeding, *Plant Dis.* 96 (2012) 4–17.
- [2] S. Troussieux, A. Gilgen, J.L. Souche, A new biocontrol tool to fight potato late blight based on *Willarta magna* C2c Maky Lysate, *Plants* 11 (2022) 2756.
- [3] A.A. Ivanov, E.O. Ukladov, T.S. Golubeva, *Phytophthora infestans*: an overview of methods and attempts to combat late blight, *J. Fungi* 12 (2021) 1071.
- [4] S.M. Dong, S.Q. Zhou, Potato late blight caused by *Phytophthora infestans*: From molecular interactions to integrated management strategies, *J. Integr. Agric.* 21 (2022) 3456–3466.
- [5] S.J. Lee, J.K.C. Rose, Mediation of the transition from biotrophy to necrotrophy in hemibiotrophic plant pathogens by secreted effector proteins, *Plant Signal. Behav.* 5 (2010) 769–772.
- [6] S.C. Whisson, P.C. Boevink, S. Wang, P.R. Birch, The cell biology of late blight disease, *Curr. Opin. Microbiol.* 34 (2016) 127–135.
- [7] P.C. Boevink, P.R.J. Birch, D. Turnbull, S.C. Whisson, Devastating intimacy: the cell biology of plant–*Phytophthora* interactions, *New Phytol.* 228 (2020) 445–458.
- [8] J.I.B. Bos, T.D. Kanneganti, C. Young, C. Cakir, E. Huitema, J. Win, M.R. Armstrong, P.R.J. Birch, S. Kamoun, The C-terminal half of *Phytophthora infestans* RXLR effector AVR3a is sufficient to trigger R3a-mediated

- hypersensitivity and suppress INF1-induced cell death in *Nicotiana benthamiana*, *Plant J.* 48 (2006) 165–176.
- [9] P.N. Dodds, J.P. Rathjen, Plant immunity: towards an integrated view of plant-pathogen interactions, *Nat. Rev. Genet.* 11 (2010) 539–548.
- [10] P. Paluchowska, J. Śliwka, Z. Yin, Late blight resistance genes in potato breeding, *Planta* 255 (2022) 127.
- [11] J.K. von Dahlen, K. Schulz, J. Nicolai, L.E. Rose, Global expression patterns of R-genes in tomato and potato, *Front. Plant Sci.* 14 (2023) 1216795.
- [12] G.J.A. Thilliez, M.R. Armstrong, T.Y. Lim, K. Baker, A. Jouet, B. Ward, C. van Oosterhout, J.D.G. Jones, E. Huitema, P.R.J. Birch, I. Hein, Pathogen enrichment sequencing (PenSeq) enables population genomic studies in oomycetes, *New Phytol.* 221 (2019) 1634–1648.
- [13] A. Jouet, D.G.O. Saunders, M. McMullan, B. Ward, O. Furzer, F. Jupe, V. Cevik, I. Hein, G.J.A. Thilliez, E. Holub, C. van Oosterhout, J.D.G. Jones, *Albugo candida* race diversity, ploidy and host-associated microbes revealed using DNA sequence capture on diseased plants in the field, *New Phytol.* 221 (2019) 1529–1543.
- [14] X. Lin, T. Song, S. Fairhead, K. Witek, A. Jouet, F. Jupe, A.I. Witek, H.S. Karki, V.G. A.A. Vleeshouwers, I. Hein, J.D.G. Jones, Identification of AvrM1 from *Phytophthora infestans* using long read and cDNA pathogen-enrichment sequencing (PenSeq), *Mol. Plant Pathol.* 21 (2020) 1502–1512.
- [15] X. Chen, D. Lewandowska, M.R. Armstrong, K. Baker, T.Y. Lim, M. Bayer, B. Harrower, K. McLean, F. Jupe, K. Witek, A.K. Lees, J.D. Jones, G.J. Bryan, I. Hein, Identification and rapid mapping of a gene conferring broad-spectrum late blight resistance in the diploid potato species *Solanum verrucosum* through DNA capture technologies, *Theor. Appl. Genet.* 131 (2018) 1287–1297.
- [16] J. Li, A. Kaur, B. Harrower, M. Armstrong, D. Dou, X. Wang, I. Hein, Identification and mapping of *Rpi-blb4* in diploid wild potato species *Solanum bulbocastanum*, *Crop J.* 11 (2023) 1828–1835.
- [17] K. Witek, F. Jupe, A.I. Witek, D. Baker, M.D. Clark, J.D.G. Jones, Accelerated cloning of a potato late blight-resistance gene using RenSeq and SMRT sequencing, *Nat. Biotechnol.* 34 (2016) 656–660.
- [18] M.R. Armstrong, J. Vossen, T.Y. Lim, R.C.B. Hutten, J. Xu, S.M. Strachan, B. Harrower, N. Champouret, E.M. Gilroy, I. Hein, Tracking disease resistance deployment in potato breeding by enrichment sequencing, *Plant Biotechnol. J.* 17 (2019) 540–549.
- [19] S. Arora, B. Steuernagel, S. Chandramohan, Y. Long, O. Matny, R. Johnson, J. Enk, S. Periyannan, M.A.M. Hatta, N. Athiyannan, J. Cheema, G. Yu, N. Kangara, S. Ghosh, L.J. Szabo, J. Poland, H. Bariana, J.D.G. Jones, A.R. Bentley, M. Ayliffe, E. Olson, S.S. Xu, B.J. Steffenson, E. Lagudah, B.B.H. Wulff, Resistance gene discovery and cloning by sequence capture and association genetics, *Nat. Biotechnol.* 37 (2019) 139–143.
- [20] N.M. Vendelbo, K. Mahmood, B. Steuernagel, B.B.H. Wulff, P. Sarup, M.S. Hovmøller, A.F. Justesen, P.S. Kristensen, J. Orabi, A. Jahoor, Discovery of resistance genes in rye by targeted long-read sequencing and association genetics, *Cells* 11 (2022) 1273.
- [21] Y. Wang, L.H. Brown, T.M. Adams, Y.W. Cheung, J. Li, V. Young, D.T. Todd, M.R. Armstrong, K. Neugebauer, A. Kaur, B. Harrower, S. Oome, X. Wang, M. Bayer, I. Hein, SMRT-AgRenSeq-d in potato (*Solanum tuberosum*) as a method to identify candidates for the nematode resistance *Gpa5*, *Hortic. Res.* 10 (2023) uhad211.
- [22] G. Andolfo, F. Jupe, K. Witek, G.J. Etherington, M.R. Ercolano, J.D.G. Jones, Defining the full tomato NB-LRR resistance gene repertoire using genomic and cDNA RenSeq, *BMC Plant Biol.* 14 (2014) 120.
- [23] L.J.M. Ducreux, W.L. Morris, M.A. Taylor, S. Millam, *Agrobacterium*-mediated transformation of *Solanum phureja*, *Plant Cell Rep.* 24 (2005) 10–14.
- [24] A.M. Michalska, S. Sobkowiak, B. Flis, E. Zimnoch-Guzowska, Virulence and aggressiveness of *Phytophthora infestans* isolates collected in Poland from potato and tomato plants identified no strong specificity, *Eur. J. Plant Pathol.* 144 (2016) 325–336.
- [25] F. Jupe, K. Witek, W. Verweij, J. Śliwka, L. Pritchard, G.J. Etherington, D. Maclean, P.J. Cock, R.M. Leggett, G.J. Bryan, L. Cardle, I. Hein, J.D.G. Jones, 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, *Plant J.* 76 (2013) 530–544.
- [26] S. Chen, Y. Zhou, Y. Chen, J. Gu, fastp: an ultra-fast all-in-one FASTQ preprocessor, *Bioinformatics* 34 (2018) i884–i890.
- [27] D. Kim, J.M. Paggi, C. Park, C. Bennett, S.L. Salzberg, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype, *Nat. Biotechnol.* 37 (2019) 907–915.
- [28] W. Guo, N.A. Tzioutziou, G. Stephen, I. Milne, C.P.G. Calixto, R. Waugh, J.W.S. Brown, R. Zhang, 3D RNA-seq: a powerful and flexible tool for rapid and accurate differential expression and alternative splicing analysis of RNA-seq data for biologists, *RNA Biol.* 18 (2021) 1574–1587.
- [29] C. Sonesson, M.I. Love, M.D. Robinson, Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences, *F1000Research* 4 (2016) 1521.
- [30] R. Patro, G. Duggal, M.I. Love, R.A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression, *Nat. Methods* 14 (2017) 417–419.
- [31] J.H. Bullard, E. Purdom, K.D. Hansen, S. Dudoit, Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments, *BMC Bioinformatics* 11 (2010) 94.
- [32] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, G.K. Smyth, Limma powers differential expression analyses for RNA-sequencing and microarray studies, *Nucleic Acids Res.* 43 (2015) e47.
- [33] C.W. Law, Y. Chen, W. Shi, G.K. Smyth, Voom: precision weights unlock linear model analysis tools for RNA-seq read counts, *Genome Biol.* 15 (2014) R29.
- [34] Y. Benjamini, D. Yekutieli, The control of the false discovery rate in multiple testing under dependency, *Ann. Stat.* 29 (2001) 1165–1188.
- [35] J. Kourelis, T. Sakai, H. Adachi, S. Kamoun, RefPlantNLR is a comprehensive collection of experimentally validated plant disease resistance proteins from the NLR family, *PLoS Biol.* 19 (2021) e3001124.
- [36] B.T. Sherman, M. Hao, J. Qiu, X. Jiao, M.W. Baseler, H.C. Lane, T. Imamichi, W. Chang, DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update), *Nucleic Acids Res.* 50 (2022) W216–W221.
- [37] L.J. Grenville-Briggs, V.L. Anderson, J. Fugelstad, A.O. Avrova, J. Bouzenzana, A. Williams, S. Wawra, S.C. Whisson, P.R.J. Birch, V. Bulone, P. Van West, Cellulose synthesis in *Phytophthora infestans* is required for normal appressorium formation and successful infection of potato, *Plant Cell* 3 (2008) 720–738.
- [38] A.O. Avrova, P.C. Boevink, V. Young, L.J. Grenville-Briggs, P. van West, P.R.J. Birch, S.C. Whisson, A novel *Phytophthora infestans* haustorium-specific membrane protein is required for infection of potato, *Cell. Microbiol.* 10 (2008) 2271–2284.
- [39] A.P. Zuluaga, J.C. Vega-Arreguín, Z. Fei, L. Ponnala, S.J. Lee, A.J. Matas, S. Patev, W.E. Fry, J.K.C. Rose, Transcriptional dynamics of *Phytophthora infestans* during sequential stages of hemibiotrophic infection of tomato, *Mol. Plant Pathol.* 17 (2016) 29–41.
- [40] H.B. Kardile, S.G. Karkute, C. Challam, N.K. Sharma, R.M. Shelake, P.G. Kawar, V. U. Patil, R. Deshmukh, V. Bhardwaj, K.N. Chourasia, S.D. Valluri, Hemibiotrophic *Phytophthora infestans* modulates the expression of *SWEET* genes in potato (*Solanum tuberosum* L.), *Plants* 12 (2023) 3433.
- [41] D.P. Martin, J. Miya, J.W. Reeser, S. Roychowdhury, Targeted RNA sequencing assay to characterize gene expression and genomic alterations, *J. Vis. Exp.* 2016 (2016) 1–9.
- [42] Z. Wang, M. Gerstein, M. Snyder, RNA-Seq: A revolutionary tool for transcriptomics, *Nat. Rev. Genet.* 1 (2009) 57–63.
- [43] J. Zhou, Y. Qi, J. Nie, L. Guo, M. Luo, H. McLellan, P.C. Boevink, P.R.J. Birch, Z. Tian, A *Phytophthora* effector promotes homodimerization of host transcription factor StKNOX3 to enhance susceptibility, *J. Exp. Bot.* 73 (2022) 6902–6915.
- [44] S.Y.A. Rodenburg, M.F. Seidl, D. de Ridder, F. Govers, Genome-wide characterization of *Phytophthora infestans* metabolism: a systems biology approach, *Mol. Plant Pathol.* 19 (2018) 1403–1413.
- [45] L.J. Grenville-Briggs, A.O. Avrova, C.R. Bruce, A. Williams, S.C. Whisson, P.R.J. Birch, P. Van West, Elevated amino acid biosynthesis in *Phytophthora infestans* during appressorium formation and potato infection, *Fungal Genet. Biol.* 42 (2005) 244–256.
- [46] M. Tian, B. Benedetti, S. Kamoun, A second Kazal-like protease inhibitor from *Phytophthora infestans* inhibits and interacts with the apoplast pathogenesis-related protease P69B of tomato, *Plant Physiol.* 138 (2005) 1785–1793.
- [47] M. Tian, E. Huitema, L. Da Cunha, T. Torto-Alalibo, S. Kamoun, A Kazal-like extracellular serine protease inhibitor from *Phytophthora infestans* targets the tomato pathogenesis-related protease P69B, *J. Biol. Chem.* 279 (2004) 26370–26377.
- [48] P.S.M. Van Weymers, K. Baker, X. Chen, B. Harrower, D.E.L. Cooke, E.M. Gilroy, P.R.J. Birch, G.J.A. Thilliez, A.K. Lees, J.S. Lynott, M.R. Armstrong, G. McKenzie, G. J. Bryan, I. Hein, Utilizing “Omic” technologies to identify and prioritize novel sources of resistance to the oomycete pathogen *Phytophthora infestans* in potato germplasm collections, *Front. Plant Sci.* 7 (2016) 672.
- [49] R.R. Singh, Target enrichment approaches for next-generation sequencing applications in oncology, *Diagnostics* 12 (2022) 1539.
- [50] X. Niu, Gene regulatory machinery and proteomics of sexual reproduction in *Phytophthora infestans*, University of California, Ph.D. Thesis, Riverside, CA, USA, 2010.
- [51] M. Abrahamian, A.M. Ah-Fong, C. Davis, K. Andreeva, H.S. Judelson, Gene expression and silencing studies in *Phytophthora infestans* Reveal infection-specific nutrient transporters and a role for the nitrate reductase pathway in plant pathogenesis, *PLoS Pathog.* 12 (2016) e1006097.
- [52] D. Turnbull, H. Wang, S. Breen, M. Malec, S. Naqvi, L. Yang, L. Welsh, P. Hemsley, Z.D. Tian, F. Brunner, E.M. Gilroy, P.R.J. Birch, AVR2 targets BSL family members, which act as susceptibility factors to suppress host immunity, *Plant Physiol.* 180 (2019) 571–581.
- [53] P.M.J.A. van Poppel, J. Guo, P.J.I. van de Vondervoort, M.W. Jung, P.R. Birch, S.C. Whisson, F. Govers, The *Phytophthora infestans* avirulence gene *Avr4* encodes an RXLR-dEER effector, *Mol. Plant-Microbe Interact.* 21 (2008) 1460–1470.
- [54] S.K. Oh, C. Young, M. Lee, R. Oliva, T.O. Bozkurt, L.M. Cano, J. Win, J.L.B. Bos, H.Y. Liu, M. Van Damme, W. Morgan, D. Choi, E.A.G. Van Der Vossen, V.G.A.A. Vleeshouwers, S. Kamoun, 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 (2009) 2928–2947.
- [55] S.R.F. King, H. McLellan, P.C. Boevink, M.R. Armstrong, T. Bukharov, O. Sukarta, J. Win, S. Kamoun, P.R.J. Birch, M.J. Banfield, *Phytophthora infestans* RXLR effector PexRD2 interacts with host MAPKKKε to suppress plant immune signaling, *Plant Cell* 26 (2014) 1345–1359.
- [56] T.O. Bozkurt, S. Schornack, J. Win, T. Shindo, M. Ilyas, R. Oliva, L.M. Cano, A.M.E. Jones, E. Huitema, R.A.L. van der Hoorn, S. Kamoun, *Phytophthora infestans* effector AVRblb2 prevents secretion of a plant immune protease at the haustorial interface, *Proc. Natl. Acad. Sci. U. S. A.* 108 (2011) 20832–20837.