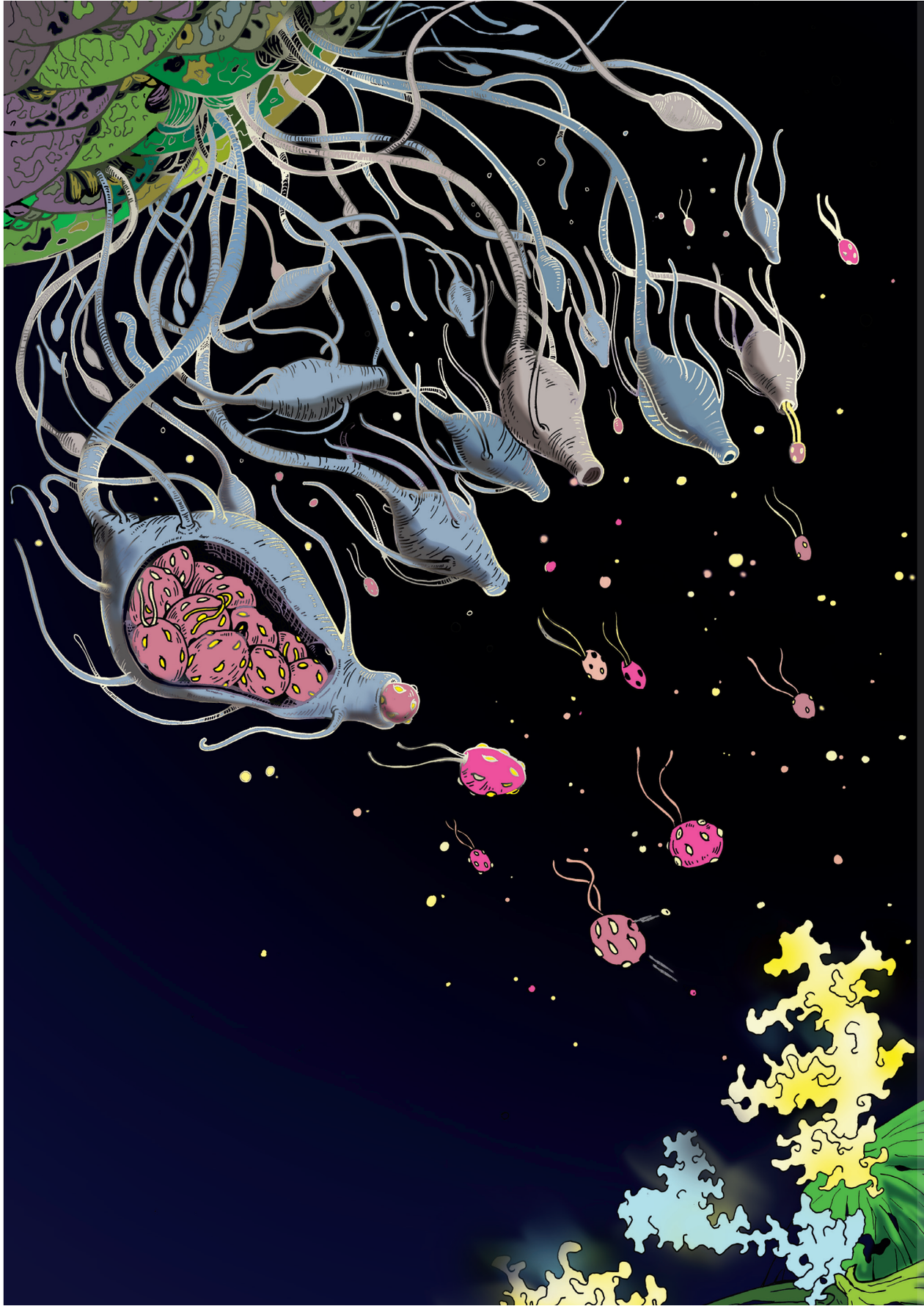
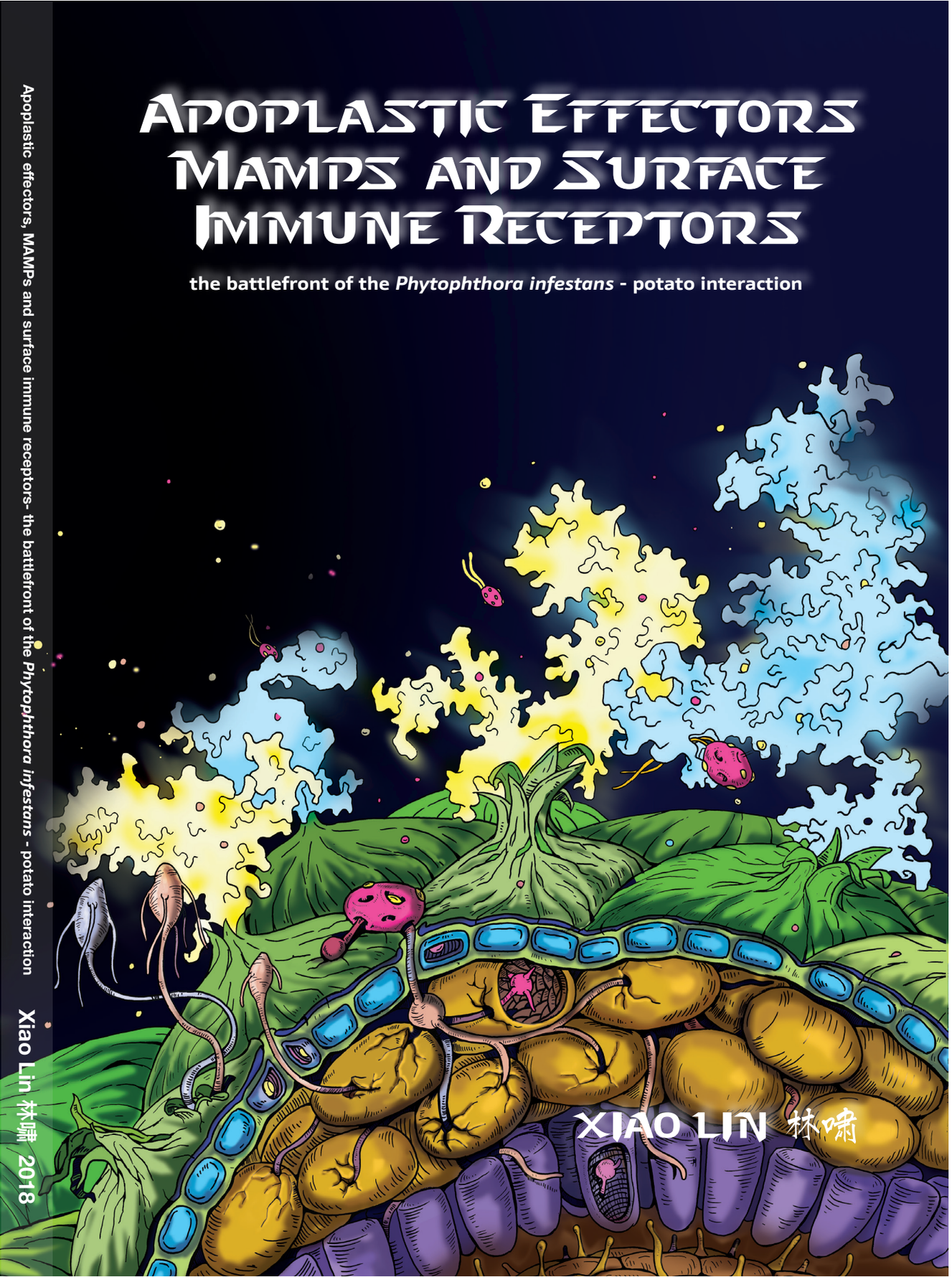
Apoplastic Effectors Mamps and Surface Immune Receptors

the battlefront of the *Phytophthora infestans* - potato interaction



Apoplastic effectors, MAMPs and surface immune receptors- the battlefront of the *Phytophthora infestans* - potato interaction

Xiao Lin 林啸 2018



Propositions

1. Durable disease resistance is unreachable, we just need "more durable" resistance.
   (this thesis)

2. The fertilization system of flowering plants might be evolved from plant immune system.
   (this thesis)

3. GMO is the most "organic" technology one can think of.

4. The breakthrough of modern science is purely build on money.

5. Wet lab works will be replaced by Artificial Intelligence (AI) and robots in the near future.

6. Basic bioinformatics and genomics need to be mastered by all biologists.

7. A PhD journey is similar to rock climbing, falling makes you stronger.

Propositions belonging to the thesis entitled "**Apoplastic effectors, MAMPs and surface immune receptors**- the battlefront of the *Phytophthora infestans* - potato interaction"

Xiao Lin

Wageningen, 31 October 2018

Apoplastic effectors, MAMPs and surface immune receptors

the battlefront of the *Phytophthora infestans* - potato interaction

Xiao Lin

Thesis Committee

Promotor
Prof. Dr Richard G. F. Visser
Professor of Plant Breeding
Wageningen University & Research

Co-promotor
Dr Vivianne G. A. A. Vleeshouwers
Assistant professor, Laboratory of Plant Breeding
Wageningen University & Research

Other members
Dr Edwin van der Vossen, Keygene N.V., Wageningen
Prof. Dr Francine P. M. Govers, Wageningen University & Research
Dr Stan Oome, Averis seeds B.V., Valthermond
Dr Frank L. W. Takken, University of Amsterdam

This research was conducted under the auspices of the Graduate School EPS (Experimental Plant Sciences).

Apoplastic effectors, MAMPs and surface immune receptors

the battlefront of the *Phytophthora infestans* - potato interaction

Xiao Lin

Thesis

submitted in fulfilment of the requirements for the degree of doctor at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A. P. J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Wednesday 31 October 2018
at 4 p.m. in the Aula

Xiao Lin

Apoplastic effectors, MAMPs and surface immune receptors - the battlefront of the *Phytophthora infestans* - potato interaction

198 pages

PhD thesis, Wageningen University, Wageningen, the Netherlands (2018)

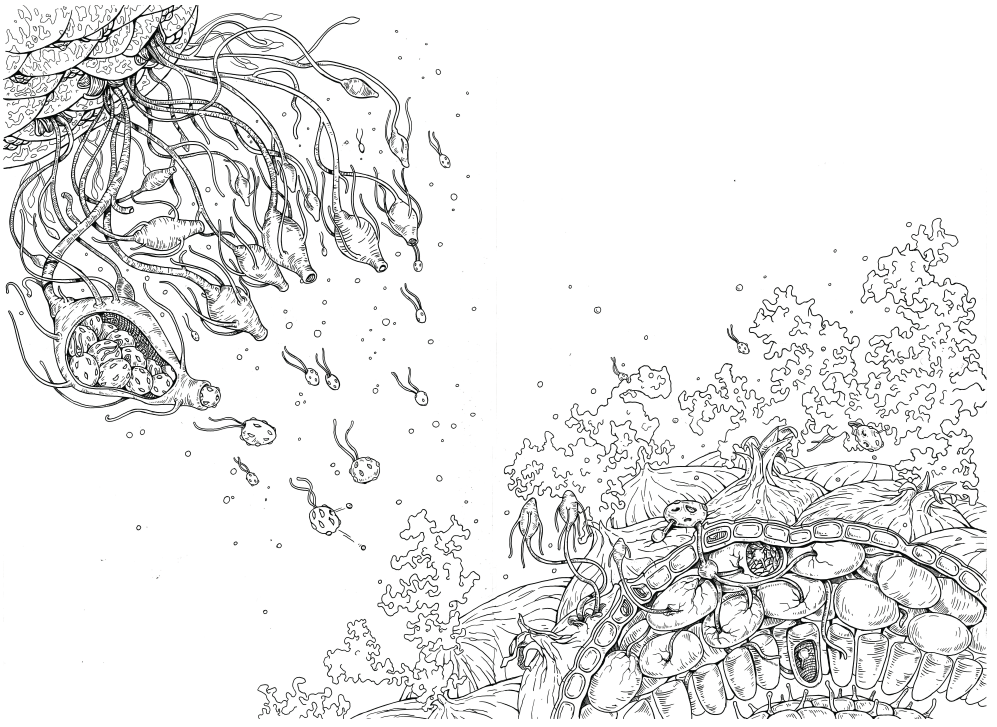
With references, with summaries in English and Chinese.

ISBN: 978-94-6343-325-9
DOI: 10.18174/456582

Table of Contents

Chapter 1

General Introduction



World population growth and crop loss due to plant pathogens

The world population is expected to reach 9-10 billion in 2050 (Alexandratos and Bruinsma, 2012). How to feed all these people will be a huge challenge, and the increase of crop yield will be crucial. However, many factors can cause yield loss. These include abiotic stresses like drought, salinity, cold, nutrient deficiencies influencing the yield, as well as many biotic factors like herbivores, pests, weeds and plant pathogens, such as bacteria, fungi, oomycetes, viruses and nematodes. More than 30% of the crops are lost due to plant pathogens or pests (Oerke, 2006). In traditional agriculture, chemicals are used to deal with those biotic stress factors. It is efficient, but it might cause environmental damage and food safety problems. Meanwhile, the pests and pathogens might start to build up resistance to some chemicals due to natural selection. Therefore, plant breeders are looking for resistance sources from wild relatives of the crops and introduce them into the modern cultivars by introgression breeding. The development of plant transformation technologies and genetic modification technologies (GM) since the 1980s, and additional technologies emerging in recent years like marker assisted selection (MAS), next and third generation sequencing and gene editing. They are expected to help mankind to deal with these plant pathogens and solve the food demand in 2050 (Boyd *et al.*, 2013).

Potato late blight and resistance breeding

Potato (*Solanum tuberosum*) is an important food crop worldwide. They are clonally propagated by the vegetative tubers, because of their highly heterozygous nature and inbreeding depression. Wild relatives of potato are tuber-bearing *Solanum* species of section *Petota* and include more than 200 species, originated from South, Central and North America (Spooner *et al.*, 2014). The ploidy level of these wild potatoes varies from diploid to hexaploid (Machida-Hirano, 2015). Cultivated tetraploid potato was domesticated from the diploid wild relatives in South America (Spooner *et al.*, 2005).

The most notorious disease in potato is late blight, which is caused by the devastating oomycete *Phytophthora infestans* (Mont.) de Bary **(Figure 1)**. This disease triggered the Irish famine in 1840, when more than a million people died due to starvation and another million people mostly emigrated to the USA (Bourke, 1989). Today, the late blight still hampers the worldwide potato production, "oomicides" (chemicals needed to control the oomycetes) need to

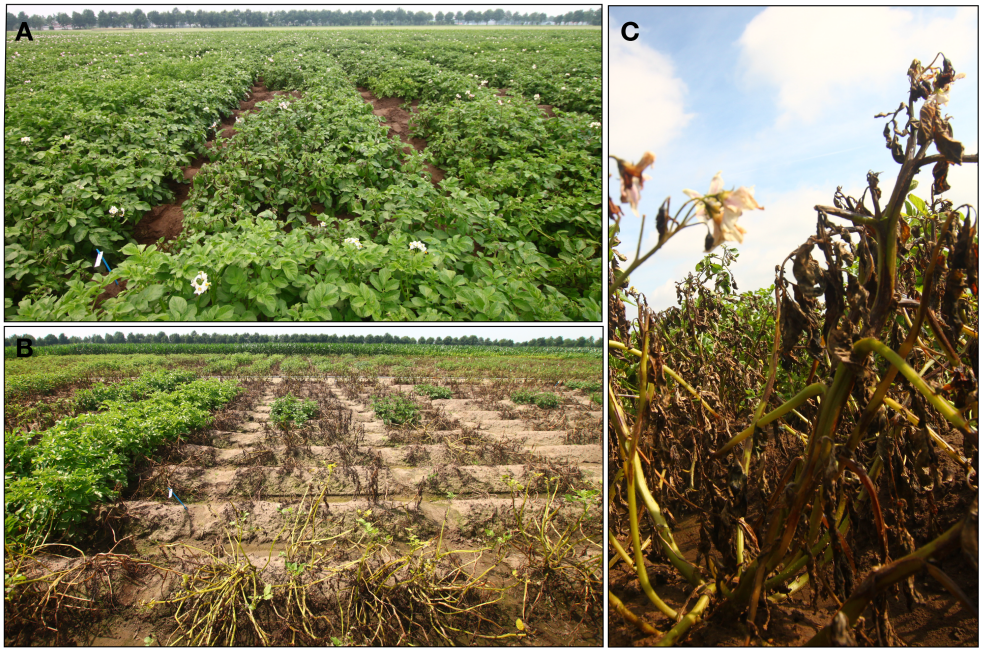


Figure 1. Field trial of potato late blight (Wageningen, 2013)

A. Day 10 after inoculation of the field trial; **B.** Day 31 after inoculation of the field trial; **C.** Close up of late blight symptoms

be sprayed many times to control the disease (Govers, 2001). In the beginning of the 20th century, a hexaploid (2n=72) wild potato *Solanum demissum* was found to be resistant to late blight (Salaman, 1937), since then, attempts have been made by breeders in the Netherlands, UK, Germany and Russia to transfer these resistance traits to cultivated potatoes by introgression breeding (Toxopeus, 1956). During the same period, the "gene-for-gene" hypothesis had been proposed by Flor (Flor, 1942), and it was found that the potato-*Phytophthora* interaction fits well with this model (Toxopeus, 1956). Eleven dominant *R* genes: *R1-R11* had been identified from *S. demissum* (Malcolmson and Black, 1966; Skidmore and Shattock, 1985). *R1* was the first *R* gene cloned from potato against *P. infestans* (Ballvora *et al.*, 2002), since then, many resistance genes against *P. infestans* (*Rpi*) genes from *S. demissum* and other wild *Solanum* species have been cloned, like *R3a*, *R8*, *Rpi-vnt1*, *Rpi-blb1*, *Rpi-blb2*, *Rpi-amr1* (Huang *et al.*, 2005; van der Vossen *et al.*, 2003; Vossen *et al.*, 2005; Foster *et al.*, 2009; Pel *et al.*, 2009; Song *et al.*, 2003; Witek *et al.*, 2016). Many of them have already been commercially used in modern cultivars, for example, the *Rpi-vnt1* in the 2nd generation of the Innate® potatoes from the Simplot company.

The *R* gene based immunity is typical effector triggered immunity (ETI), however, *P. infestans* can fast evolve to avoid the surveillance. For example, the *Avr* genes in the virulence *P. infestans* isolates can be absent, mutated or silenced (Vleeshouwers and Oliver, 2014), so late blight is called plant and (*R* gene) destroyer (Fry, 2008). Therefore, the *R* genes might be quickly defeated by *P. infestans* in the field. So far, most of the identified *R* genes have been overcome by different *P. infestans* isolates, and resistant varieties need an additionally *R* gene management system. How to achieve more durable resistance remains a challenge.

Except *R* genes, susceptibility (*S*) genes have also been applied in crops (van Schie and Takken, 2012), silencing the *S* genes in potato can lead to enhanced resistance to late blight (Sun *et al.*, 2016). This knowledge can now be combined efficiently with CRISPR-Cas9 technology to engineer foreign DNA free resistant crops (Langner *et al.*, 2018).

Plant immunity system

Unlike animals, plants are immobile and lack of mobile immunity cells. Apart from the barriers on the surface of plant cells like waxes, trichomes and stomata, plants evolved an innate immunity system to protect them against pathogens. This innate immunity system was summarized by a "zigzag" model (Jones and Dangl, 2006). In this model, plants perceive microbial- or pathogen-associated molecular patterns (MAMPs/PAMPs) by pattern recognition receptors (PRRs), and mount PAMP-triggered immunity (PTI) responses. To successfully colonize the host, the pathogens evolved cytoplasmic effectors that can inhibit the plant PTI, which is leading to effector-triggered susceptibility (ETS). For counter-acting this ETS, intercellular nucleotide-binding domain and leucine-rich repeat containing (NLR) receptors have evolved to recognize the cytoplasmic effectors, and mount a hypersensitive cell death response (HR) that delimits pathogen growth. This is called effector-triggered immunity (ETI) (Dodds and Rathjen, 2010; Jones and Dangl, 2006).

In previous reviews, the surface immune receptors are called PRRs (Dodds and Rathjen, 2010), however, with the increasing knowledge of the effector biology, the ligands of plant surface immune receptors can be both MAMPs/PAMPs or apoplastic effectors (Win *et al.*, 2013). In this thesis, to avoid ambiguity, we use the term 'surface immune receptor', regardless whether the ligands are MAMPs/PAMPs or apoplastic effectors.

Pathogen encoded MAMPs and MAMP receptors in plants

Up to now, many MAMPs/PAMPs from pathogens have been identified **(Table 1)**, like flg22 (flagellin), flgII-28 (flagellin), elf18 (EF-Tu), csp22 (cold-shock protein), peptidoglycans (PGN) and lipopolysaccharides (LPS) from bacteria (Felix *et al.*, 1999; Clarke *et al.*, 2013; Kunze, 2004; Böhm *et al.*, 2014; Oome *et al.*, 2014; Gust *et al.*, 2007; Dow *et al.*, 2000; Felix and Boller, 2003), chitin and β-glucan from fungi (Yamaguchi *et al.*, 2000; Böhm *et al.*, 2014), and nlp20 (NLP) from bacteria, fungi and oomycetes (Böhm *et al.*, 2014). In the meantime, some of the MAMPs/PAMPs receptors have been found as well **(Table 1)**. These include FLS2 (flg22), FLS3 (flgII-28), EFR (EF-Tu) , RLP23 (nlp20), LysM proteins (PGNs), LORE (LPS), CORE (csp22), LYsM RLK1/CERK1 (Chitin) (Gómez-Gómez and Boller, 2000; Hind *et al.*, 2016; Zipfel *et al.*, 2006; Albert *et al.*, 2015; B., Liu *et al.*, 2012; Ranf *et al.*, 2015; Wang *et al.*, 2016; Miya *et al.*, 2007; Wan *et al.*, 2008) **(Figure 2)**.

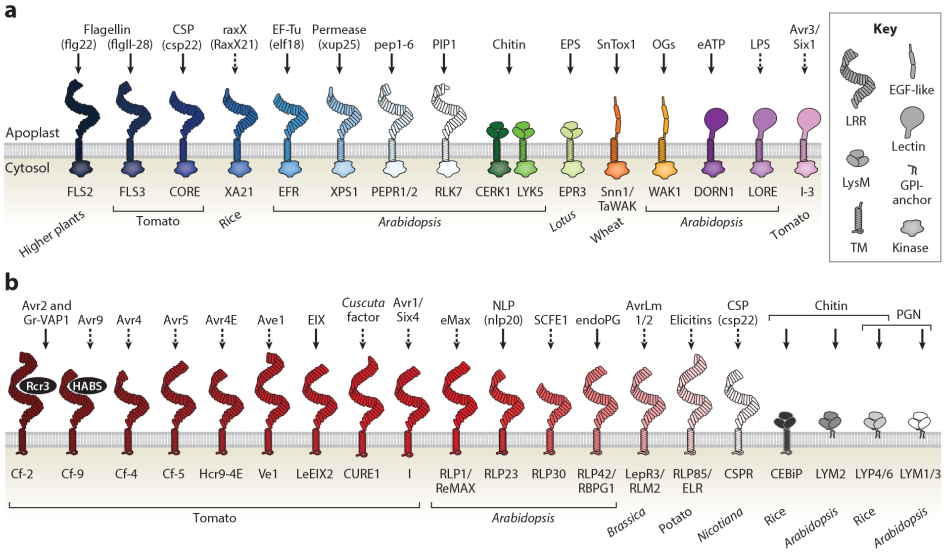


Figure 2. An overview of plant surface immune receptors and the corresponding pathogen MAMPs or apoplastic effectors (adapted from Boutrot and Zipfel, 2017)**.**

The surface immune receptors consist of cytoplastic kinase domains (RLKs only), transmembrane domain and an extracellular domain. The types of extracellular domains include leucine-rich repeat (LRR) domains, epidermal growth factor (EGF)-like domains, Lysin motif (LysM) domains, Lectin domains and glycosyl-phosphatidylinositol (GPI) domains **(a).** Cloned receptor-like kinases (RLKs). **(b).** Cloned receptor-like proteins (RLPs).

In oomycetes, MAMPs/PAMPs include INF1 (Elicitin), XEG1, Pep-13/25 (GP42), CBEL and NPP1 (Kamoun *et al.*, 1993; Ma *et al.*, 2015; Brunner *et al.*, 2002; Mateos *et al.*, 1997; Fellbrich *et al.*, 2002). Till now, only three surface immune receptors have been cloned that recognize oomycetes MAMPs/PAMPs: ELR was the first potato surface immune receptor cloned from wild *Solanum*, it recognizes the conserved elicitins from oomycetes (Du *et al.*, 2015), and RXEG1 and RLP23 were cloned from *N. benthamiana* and *Arabidopsis thaliana*, they can recognize XEG1 and NLPs (nlp20) respectively (Wang *et al.*, 2018; Albert *et al.*, 2015). In addition, a L-type LecRK gene *LecRK-I.9* from *Arabidopsis* enhances resistance to *P. infestans* in potato and *N. benthamiana* (Bouwmeester *et al.*, 2014). Whereas the other surface immune receptors remain unknown, like the Pep-13/25 and CBEL receptors.

Table 1. Examples of MAMPs/PAMPs and apoplastic effectors from fungi, bacteria and oomycetes and their surface immune receptors cloned from plants.

	MAMPs/ apoplastic effectors	Molecular patterns	Reference	Surface immune receptors	Cloned from	Reference
Bacterial	Flagellin	flg22	(Felix, 1999)	FLS2	*Arabidopsis*	(Gómez-Gómez, 2000)
		flgII-28	(Clarke, 2013)	FLS3	Tomato	(Hind, 2016)
	EF-Tu	elf18	(Kunze, 2004)	EFR	*Arabidopsis*	(Zipfel, 2006)
	NLP	nlp20	(Bohm, 2014; Oomy, 2014)	RLP23	*Arabidopsis*	(Albert, 2015)
	Peptidoglycans (PGNs)	PGNs	(Gust, 2007)	LysM proteins	Rice	(Liu, 2012)
	Lipopolysaccharides (LPSs)	LPSs	(Dow, 2000)	LORE	*Arabidopsis*	(Ranf, 2015)
	Cold shock protein	csp22	(Felix, 2003)	CORE	Tomato	(Wang, 2016)
	Cip1		(Shindo, 2016)	unknown		
	Chp7		(Lu, 2015)	unknown		
Fungi	Chitin		(Felix, 1993)	LysM RLK1/ CERK1	*Arabidopsis*	(Miya, 2007; Wan, 2008)
	β-glucan		(Yamaguchi, 2000)	unknown		
	NLP	nlp20	(Bohm, 2014; Oomy, 2014)	RLP23	*Arabidopsis*	(Albert, 2015)
	Avr2		(Luderer, 2002)	Cf-2	tomato	(Dixon, 1996)
	Avr4		(Joosten, 1994)	Cf-4	tomato	(Joosten, 1994)
	Avr9		(Jones, 1994)	Cf-9	tomato	(Jones, 1994)
	Ecp6		(Bolton, 2008)	unknown		
	Ave1		(deJonge, 2012)	Ve1	tomato	(Kawchuk, 2001)
Oomycete	INF1 (Elicitins)		(Kamoun, 1993)	ELR	potato	(Du, 2015)
	XEG1		(Ma, 2015)	RXEG1	*N. benthamiana*	(Wang, 2018)
	Transglutaminase (GP42)	Pep-13/25	(Brunner, 2002)	unknown		
	CBEL		(Mateos, 1997)	unknown		
	NPP1		(Mateos, 1997)			
	PcF		(Orsomando, 2001)	unknown		
	SCR74		(Liu, 2005)	unknown		
	EPI1 and EPI10		(Tian, 2005; Tian, 2004)	unknown		
	EPIC1 and EPIC2		(Tian, 2007)	unknown		

Pathogen apoplastic effectors and surface immune receptors in plants

Except the classical MAMPs/PAMPs, pathogens can secrete apoplastic effectors in order to manipulate the host to facilitate their colonization. For example, the tomato leaf mold agent *Cladosporium fulvum* is a biotrophic fungus, for which the host colonization is limited on the intercellular spaces. The molecular interaction between tomato and *C. fulvum* has been a classical model to study the plant surface immunity (de Wit, 2016). Many apoplastic effectors from *C. fulvum* have been characterized like AVR2, AVR4 and AVR9, as well as their matching surface immune receptors CF-2, CF-4 and CF-9 (Luderer *et al.*, 2002; Joosten *et al.*, 1994; Dixon *et al.*, 1996). Other extracellular proteins (ECP1-ECP6) were also found from *C. fulvum* (Laugé *et al.*, 2000; Bolton *et al.*, 2008), and the LysM motif containing effector Ecp6 was found to disturb chitin-triggered immunity in plants (de Jonge *et al.*, 2010).

In oomycetes, many apoplastic effectors have been characterized, like the small-cysteine rich (SCR) proteins PcF, SCR74 and SCR91 (Orsomando *et al.*, 2001; Liu *et al.*, 2005; Bos *et al.*, 2003); the enzyme inhibitors EPI1, EPI10, EPIC1 and EPIC2 (Tian *et al.*, 2005; Tian *et al.*, 2004; Tian *et al.*, 2007). It is noteworthy that, unlike the MAMPs, SCR74 proteins are highly diverse in different *P. infestans* isolates and they are under positive selection. However, the receptors for these oomycete apoplastic effectors have yet to be discovered.

Effectoromics, from cytoplasmic effectors to apoplastic effectors

In *Phytophthora*, most identified *Avr* genes contain an RXLR motif (Rehmany *et al.*, 2005). With the sequencing of *P. infestans* genome, 563 RXLR effectors have been predicted from the *P. infestans* genome (Haas *et al.*, 2009). This enabled "Effectoromics", a high-throughput functional screening method to detect *R* genes in different potato sources. The RXLR effectors were cloned into a plant expression system and screened for cell death responses by PVX agro-infection or agro-infiltration (Vleeshouwers *et al.*, 2008; Du *et al.*, 2015; Takken *et al.*, 2000). The Effectoromics methodology enables to rapidly identify the *R* genes from wild or cultivated *Solanum* species in complex genetic backgrounds (Rietman *et al.*, 2012), and it can also be used to identify novel *Avr* genes (Vleeshouwers *et al.*, 2011). However, previous Effectoromics studies in potato were mainly restricted to the RXLR effectors. Applying the Effectoromics strategy on the MAMPs/PAMPs and apoplastic effectors will help to identify the surface immune receptors in plants (Domazakis *et al.*, 2017).

Late blight resistance breeding in the genomics era

In 2011, a homozygous doubled-monoploid *S. tuberosum* group Phureja DM1-3 516 R44 (DM) was sequenced, and it has been widely used as the potato reference genome (Xu *et al.*, 2011). Recently, the genome of another diploid and inbred *Solanum chacoense* clone M6 was released, which is a wild relative of modern cultivated potato (Leisner *et al.*, 2018). To assess the genetic variation of the tuber-bearing *Solanum*, large-scale genomic re-sequencing was performed on a diversity panel of wild potatoes, landraces and cultivars. This study suggests that some cultivated traits, for example the potato's maturity locus in long-day adapted cultivars was introgressed from wild species like *S. microdontum* (Hardigan *et al.*, 2017).

The genomic sequencing of potato allows genome-wide studies of the *NLR* and *RLP/ RLK* genes (Andolfo *et al.*, 2014). Furthermore, resistance gene enrichment sequencing (RenSeq) was developed for fast mapping potato *Rpi* genes and re-annotating of the *NLR* genes (Jupe *et al.*, 2013). In combination with long-reads PacBio sequencing, an improved RenSeq can capture the full-length *NLR* repertoires from wild *Solanum* species and help to clone *Rpi* genes more quickly (Witek *et al.*, 2016). However, the first layer of defense mediated by the *RLP/ RLK* genes were ignored in previous RenSeq studies.

Objectives of this thesis

To fill the knowledge gap and deepen our knowledge about plant surface immunity, which is believed to be more durable and broad-spectrum, in **Chapter 2**, the PcF/SCR74 family from the oomycete is studied. Sequence analysis and phylogenetic analysis are used to study these apoplastic effectors. A large-scale Effectoromics screening on Solanaceous species is performed to explore the recognition specificity of these PcF/SCR74 effectors, and we try to explain our findings in the light of (co-)evolution with host plants. In **Chapter 3,** we propose a novel pipeline to fast identify and clone the plant surface immune receptors, by combining Effectoromics screening and RLP/RLK enrichment sequencing (RLP/KSeq) method. A previous identified INF1 response gene *ELR* is used as proof of concept and an undiscovered SCR74 receptor is studied. In **Chapter 4**, we expand the mapping population and fine-map the SCR74 receptor to a *G-LecRK* locus. A bacterial artificial chromosome (BAC) library from *Solanum microdontum* spp. *gigantophyllum* GIG362-6 that responds to SCR74 is generated. Three BAC clones spanning the mapping interval were isolated and sequenced. The candidate genes are functionally studied. In **Chapter 5**, we study the *Phytophthora* MAMP Pep-13/25. We perform a functional screening and identify several Pep-13/25 responsive wild

potatoes. A population that segregates to the Pep-25 response was generated, and we developed a BSA-RNAseq method to accelerate mapping of the Pep-25 receptor. Ultimately, the putative Pep-25 receptor is mapped to a small interval and the BAC clone within the mapping interval is isolated, sequenced and characterized. The findings in this thesis will expand our current knowledge about the effector biology of oomycete and plant surface immunity. The new methodology developed in this thesis will facilitate the identification of novel *R* genes in crops, and our work should lead to the identification and isolation of at least two novel surface immune receptors in potato.

References

Albert, I., Böhm, H., Albert, M., et al. (2015) An RLP23–SOBIR1–BAK1 complex mediates NLP-triggered immunity. *NPLANTS*, **1**, 15140.

Alexandratos, N. and Bruinsma, J. (2012) World agriculture towards 2030/2050: the 2012 revision.

Andolfo, G., Jupe, F., Witek, K., Etherington, G.J., Ercolano, M.R. and Jones, J.D.G. (2014) Defining the full tomato NB-LRR resistance gene repertoire using genomic and cDNA RenSeq. *BMC Plant Biology*, **14**, 1–12.

Ballvora, A., Ercolano, M.R., Weiß, J., Meksem, K., Bormann, C.A., Oberhagemann, P., Salamini, F. and Gebhardt, C. (2002) The *R1* gene for potato resistance to late blight (*Phytophthora infestans*) belongs to the leucine zipper/NBS/LRR class of plant resistance genes. *The Plant Journal*, 30, 361–371.

Bolton, M.D., van Esse, H.P., Vossen, J.H., et al. (2008) The novel *Cladosporium fulvum* lysin motif effector Ecp6 is a virulence factor with orthologues in other fungal species. *Molecular Microbiology*, **69**, 119–136.

Bos, J.I.B., Armstrong, M., Whisson, S.C., Torto, T.A., Ochwo, M., Birch, P.R.J. and Kamoun, S. (2003) Intraspecific comparative genomics to identify avirulence genes from *Phytophthora*. *New Phytol*, 159, 63–72.

Bourke, A. (1989) Potato blight in Europe in 1845: the scientific controversy. In *Phytophthora*, Symposium of the British Mycological Society for Plant Pathology and the society of Irish Plant Pathologists held at the Trinity College. Cambridge University Press.

Boutrot, F. and Zipfel, C. (2017) Function, discovery, and exploitation of plant pattern recognition receptors for broad-spectrum disease resistance. *Annu. Rev. Phytopathol.*, **55**, 257–286.

Bouwmeester, K., Han, M., Blanco Portales, R., Song, W., Weide, R., Guo, L.Y., van der Vossen, E.A.G. and Govers, F. (2014) The *Arabidopsis* lectin receptor kinase LecRK-I.9 enhances resistance to *Phytophthora infestans* in Solanaceous plants. *Plant Biotechnology J*, 12, 10–16.

Boyd, L.A., Ridout, C., O'Sullivan, D.M., Leach, J.E. and Leung, H. (2013) Plant–pathogen interactions: disease resistance in modern agriculture. *Trends in Genetics*, **29**, 233–240.

Böhm, H., Albert, I., Oome, S., Raaymakers, T.M., Van den Ackerveken, G. and Nürnberger, T. (2014) A conserved peptide pattern from a widespread microbial virulence factor triggers pattern-induced immunity in *Arabidopsis*. *PLoS Pathog*, **10**, e1004491.

Brunner, F., Rosahl, S., Lee, J., Rudd, J.J., Geiler, C., Kauppinen, S., Rasmussen, G., Scheel, D. and Nürnberger, T. (2002) Pep-13, a plant defense-inducing pathogen-associated pattern from *Phytophthora* transglutaminases. *The EMBO Journal*, **21**, 6681–6688.

Clarke, C.R., Chinchilla, D., Hind, S.R., et al. (2013) Allelic variation in two distinct *Pseudomonas syringae* flagellin epitopes modulates the strength of plant immune responses but not bacterial motility. *New Phytol*, **200**, 847–860.

de Jonge, R., van Esse, H.P., Kombrink, A., et al. (2010) Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. *Science*, **329**, 953–955.

de Wit, P.J.G.M. (2016) *Cladosporium fulvum* Effectors: Weapons in the arms race with tomato. *Annu. Rev. Phytopathol.*, **54**, 1–23.

Dixon, M.S., Jones, D.A., Keddie, J.S., Thomas, C.M., Harrison, K. and Jones, J.D. (1996) The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. *Cell*, **84**, 451–459.

Dodds, P.N. and Rathjen, J.P. (2010) Plant immunity: towards an integrated view of plant–pathogen interactions. *Nat Rev Genet*, **11**, 539.

Domazakis, E., Lin, X., Aguilera-Galvez, C., Wouters, D., Bijsterbosch, G., Wolters, P.J. and Vleeshouwers, V.G.A.A. (2017) Effectoromics-based identification of cell surface receptors in potato. In *Methods in Molecular Biology*. Humana Press, New York, NY, pp. 337–353.

Dow, M., Newman, M.-A. and Roepenack, von, E. (2000) The induction and modulation of plant defense responses by bacterial lipopolysaccharides. *Annu. Rev. Phytopathol.*, **38**, 241–261.

Du, J., Verzaux, E., Chaparro-Garcia, A., et al. (2015) Elicitin recognition confers enhanced resistance to *Phytophthora infestans* in potato. *NPLANTS*, 10.1038–nplants.2015.34.

Felix, G. and Boller, T. (2003) Molecular sensing of bacteria in plants. The highly conserved RNA-binding motif RNP-1 of bacterial cold shock proteins is recognized as an elicitor signal in tobacco. *Journal of Biological Chemistry*, **278**, 6201–6208.

Felix, G., Duran, J.D., Volko, S. and Boller, T. (1999) Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *The Plant Journal*, **18**, 265–276.

Fellbrich, G., Romanski, A., Varet, A., et al. (2002) NPP1, a *Phytophthora*-associated trigger of plant defense in parsley and *Arabidopsis*. *The Plant Journal*, **32**, 375–390.

Flor, H.H. (1942) Inheritance of pathogenicity in *Melampsora lini*. *Phytopathology*, **32**, 653–669.

Foster, S.J., Park, T.-H., Pel, M., Brigneti, G., Śliwka, J., Jagger, L., van der Vossen, E. and Jones, J.D.G. (2009) *Rpi-vnt1.1*, a *Tm-22* homolog from *Solanum venturii*, confers resistance to potato late blight. *MPMI*, **22**, 589–600.

Fry, W. (2008) Phytophthora infestans: the plant (and *R* gene) destroyer. Molecular Plant Pathology, **9**, 385–402.

Govers, F. (2001) Misclassification of pest as "fungus" puts vital research on wrong track. *Nature*, **411**, 1–1.

Gómez-Gómez, L. and Boller, T. (2000) FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. *Molecular Cell*, **5**, 1003–1011.

Gust, A.A., Biswas, R., Lenz, H.D., et al. (2007) Bacteria-derived peptidoglycans constitute pathogen-associated molecular patterns triggering innate immunity in *Arabidopsis*. *Journal of Biological Chemistry*, **282**, 32338–32348.

Haas, B.J., Kamoun, S., Zody, M.C., et al. (2009) Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature*, **461**, 393–398.

Hardigan, M.A., Laimbeer, F.P.E., Newton, L., et al. (2017) Genome diversity of tuber-bearing *Solanum* uncovers complex evolutionary history and targets of domestication in the cultivated potato. *Proceedings of the National Academy of Sciences*, **114**, E9999–E10008.

Hind, S.R., Strickler, S.R., Boyle, P.C., et al. (2016) Tomato receptor FLAGELLIN-SENSING 3 binds flgII-28 and activates the plant immune system. *NPLANTS*, **2**, 16128.

Huang, S., van der Vossen, E.A.G., Kuang, H., et al. (2005) Comparative genomics enabled the isolation of the *R3a* late blight resistance gene in potato. *The Plant Journal*, **42**, 251–261.

Jones, J.D.G. and Dangl, J.L. (2006) The plant immune system. *Nature*, **444**, 323–329.

Joosten, M.H.A.J., Cozijnsen, T.J. and de Wit, P.J.G.M. (1994) Host resistance to a fungal tomato pathogen lost by a single base-pair change in an avirulence gene. *Nature*, **367**, 384.

Jupe, F., Witek, K., Verweij, W., 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.

Kamoun, S., Klucher, K.M., Coffey, M.D. and Tyler, B.M. (1993) A gene encoding a host-specific elicitor protein of *Phytophthora parasitica*. *MPMI*, **6**, 573–581.

Kunze, G. (2004) The N terminus of bacterial elongation factor Tu elicits innate immunity in *Arabidopsis* plants. *The Plant Cell*, **16**, 3496–3507.

Langner, T., Kamoun, S. and Belhaj, K. (2018) CRISPR Crops: Plant Genome Editing Toward Disease Resistance. *Annu. Rev. Phytopathol.*, **56**, annurev–phyto–080417–050158–34.

Laugé, R., Goodwin, P.H., de Wit, P.J.G.M. and Joosten, M.H.A.J. (2000) Specific HR-associated recognition of secreted proteins from *Cladosporium fulvum* occurs in both host and non-host plants. *The Plant Journal*, **23**, 735–745.

Leisner, C.P., Hamilton, J.P., Crisovan, E., 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.

Liu, B., Li, J.-F., Ao, Y., et al. (2012) Lysin motif–containing proteins LYP4 and LYP6 play dual roles in peptidoglycan and chitin perception in rice innate immunity. *The Plant Cell*, **24**, 3406–3419.

Liu, Z., Bos, J.I.B., Armstrong, M., et al. (2005) Patterns of diversifying selection in the phytotoxin-like *scr74* gene family of *Phytophthora infestans*. *Mol Biol Evol*, **22**, 659–672.

Luderer, R., Takken, F.L.W., de Wit, P.J.G.M. and Joosten, M.H.A.J. (2002) *Cladosporium fulvum* overcomes *Cf-2*-mediated resistance by producing truncated AVR2 elicitor proteins. *Molecular Microbiology*, **45**, 875–884.

Ma, Z., Song, T., Zhu, L., et al. (2015) A *Phytophthora sojae* glycoside hydrolase 12 protein is a major virulence factor during soybean infection and is recognized as a PAMP. *The Plant Cell*, **27**, 2057–2072.

Machida-Hirano, R. (2015) Diversity of potato genetic resources. *Breeding Science*, **65**, 26–40.

Malcolmson, J.F. and Black, W. (1966) New *R* genes in *Solanum demissum* lindl. And their complementary races of *Phytophthora infestans* (Mont.) de bary. *Euphytica*, **15**, 199–203.

Mateos, F.V., Rickauer, M. and Esquerré-Tugayé, M.-T. (1997) Cloning and characterization of a cDNA encoding an elicitor of *Phytophthora parasitica* var. *nicotianae* that shows cellulose-binding and lectin-like activities. *MPMI*, **10**, 1045–1053.

Miya, A., Albert, P., Shinya, T., et al. (2007) CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in *Arabidopsis*. *Proceedings of the National Academy of Sciences*, **104**, 19613–19618.

Oerke, E.C. (2006) Crop losses to pests. The Journal of Agricultural Science, **144**, 31–43.

Oome, S., Raaymakers, T.M., Cabral, A., Samwel, S., Böhm, H., Albert, I., Nürnberger, T. and 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*, **111**, 16955–16960.

Orsomando, G., Lorenzi, M., Raffaelli, N., Dalla Rizza, M., Mezzetti, B. and Ruggieri, S. (2001) Phytotoxic protein PcF, purification, characterization, and cDNA sequencing of a novel hydroxyproline-containing factor secreted by the strawberry pathogen *Phytophthora cactorum*. *Journal of Biological Chemistry*, **276**, 21578–21584.

Pel, M.A., Foster, S.J., Park, T.-H., et al. (2009) Mapping and cloning of late blight resistance genes from *Solanum venturii* using an interspecific candidate gene approach. *MPMI*.

Ranf, S., Gisch, N., Schäffer, M., et al. (2015) A lectin S-domain receptor kinase mediates lipopolysaccharide sensing in *Arabidopsis thaliana*. *Nat Immunol*, **16**, 426–433.

Rehmany, A.P., Gordon, A., Rose, L.E., et al. (2005) Differential recognition of highly divergent downy mildew avirulence gene alleles by *RPP1* resistance genes from two *Arabidopsis* lines. *The Plant Cell*, **17**, 1839–1850.

Rietman, H., Bijsterbosch, G., Cano, L.M., Lee, H.-R., Vossen, J.H., Jacobsen, E., Visser, R.G.F., Kamoun, S. and Vleeshouwers, V.G.A.A. (2012) Qualitative and quantitative late blight resistance in the potato cultivar Sarpo Mira is determined by the perception of five distinct RXLR effectors. *MPMI*, **25**, 910–919.

Salaman, R. (1937) Potato variety production: a new departure. *Gardeners Chronicle*, **102**, 326–327.

Skidmore, D.I. and Shattock, R.C. (1985) Monitoring of virulence in populations of *phytophthora infestans* using a Schwarzbach spore trap. *J Phytopathol*, **113**, 141–149.

Song, J., Bradeen, J.M., Naess, S.K., et al. (2003) Gene *RB* cloned from *Solanum bulbocastanum* confers broad spectrum resistance to potato late blight. *Proceedings of the National Academy of Sciences*, **100**, 9128–9133.

Spooner, D.M., Ghislain, M., Simon, R., Jansky, S.H. and Gavrilenko, T. (2014) Systematics, diversity, genetics, and evolution of wild and cultivated potatoes. *Bot. Rev.*, **80**, 283–383.

Spooner, D.M., McLean, K., Ramsay, G., Waugh, R. and Bryan, G.J. (2005) A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. *Proceedings of the National Academy of Sciences*, **102**, 14694–14699.

Sun, K., Wolters, A.-M.A., Vossen, J.H., Rouwet, M.E., Loonen, A.E.H.M., Jacobsen, E., Visser, R.G.F. and Bai, Y. (2016) Silencing of six susceptibility genes results in potato late blight resistance. *Transgenic Res*, **25**, 731–742.

Tian, M., Benedetti, B. and Kamoun, S. (2005) A second kazal-like protease inhibitor from *Phytophthora infestans* inhibits and interacts with the apoplastic pathogenesis-related protease P69B of tomato. *PLANT PHYSIOLOGY*, **138**, 1785–1793.

Tian, M., Huitema, E., da Cunha, L., Torto-Alalibo, T. and Kamoun, S. (2004) A kazal-like extracellular serine protease inhibitor from *Phytophthora infestans* targets the tomato pathogenesis-related protease P69B. *Journal of Biological Chemistry*, **279**, 26370–26377.

Tian, M., Win, J., Song, J., van der Hoorn, R., van der Knaap, E. and Kamoun, S. (2007) A *Phytophthora infestans* cystatin-like protein targets a novel tomato papain-like apoplastic protease. *PLANT PHYSIOLOGY*, **143**, 364–377.

Toxopeus, H.J. (1956) Reflections on the origin of new physiologic races in *Phytophthora infestans* and the breeding for resistance in potatoes. *Euphytica*, **5**, 221–356.

van der Vossen, E., Sikkema, A., Hekkert, B.T.L., et al. (2003) An ancient *R* gene from the wild potato species *Solanum bulbocastanum* confers broad-spectrum resistance to *Phytophthora infestans* in cultivated potato and tomato. *The Plant Journal*, 36, 867–882.

van Schie, C.C.N. and Takken, F.L.W. (2012) Susceptibility genes 101: how to be a good host. *Annu. Rev. Phytopathol.*, **52**, 140626173629003.

Vleeshouwers, V.G.A.A. and Oliver, R.P. (2014) Effectors as tools in disease resistance breeding against biotrophic, hemibiotrophic, and necrotrophic plant pathogens. *MPMI*, **27**, 196–206.

Vleeshouwers, V.G.A.A., Raffaele, S., Vossen, J.H., et al. (2011) Understanding and exploiting late blight resistance in the age of effectors. *Annu. Rev. Phytopathol.*, **49**, 507–531.

Vossen, E.A.G., Gros, J., Sikkema, A., Muskens, M., Wouters, D., Wolters, P., Pereira, A. and Allefs, S. (2005) The *Rpi-blb2* gene from S*olanum bulbocastanum* is an *Mi-1* gene homolog conferring broad-spectrum late blight resistance in potato. *The Plant Journal*, 44, 208–222.

Wan, J., Zhang, X.-C., Neece, D., Ramonell, K.M., Clough, S., Kim, S.-Y., Stacey, M.G. and Stacey, G. (2008) A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in Arabidopsis. *The Plant Cell*, **20**, 471–481.

Wang, L., Albert, M., Einig, E., Fürst, U., Krust, D. and Felix, G. (2016) The pattern-recognition receptor CORE of Solanaceae detects bacterial cold-shock protein. *NPLANTS*, 2, 1

Wang, Y., Xu, Y., Sun, Y., et al. (2018) Leucine-rich repeat receptor-like gene screen reveals that *Nicotiana* RXEG1 regulates glycoside hydrolase 12 MAMP detection. *Nat Commun*, 1–12.

Win, J., Chaparro-Garcia, A., Belhaj, K., et al. (2013) Effector biology of Plant-associated organisms: concepts and perspectives. *Cold Spring Harbor Symposia on Quantitative Biology*, **77**, 235–247.

Witek, K., Jupe, F., Witek, A.I., Baker, D., Clark, M.D. and Jones, J.D.G. (2016) Accelerated cloning of a potato late blight–resistance gene using RenSeq and SMRT sequencing. *Nature Biotechnology*, **34**, 656–660.

Xu, X., Pan, S., Cheng, S., et al. (2011) Genome sequence and analysis of the tuber crop potato. *Nature*, **475**, 189–195.

Yamaguchi, T., Yamada, A., Hong, N., Ogawa, T., Ishii, T. and Shibuya, N. (2000) Differences in the recognition of glucan elicitor signals between rice and soybean: β-Glucan fragments from the rice blast disease fungus *Pyricularia oryzae* that elicit phytoalexin biosynthesis in suspension-cultured rice cells. *The Plant Cell*, **12**, 817–826.

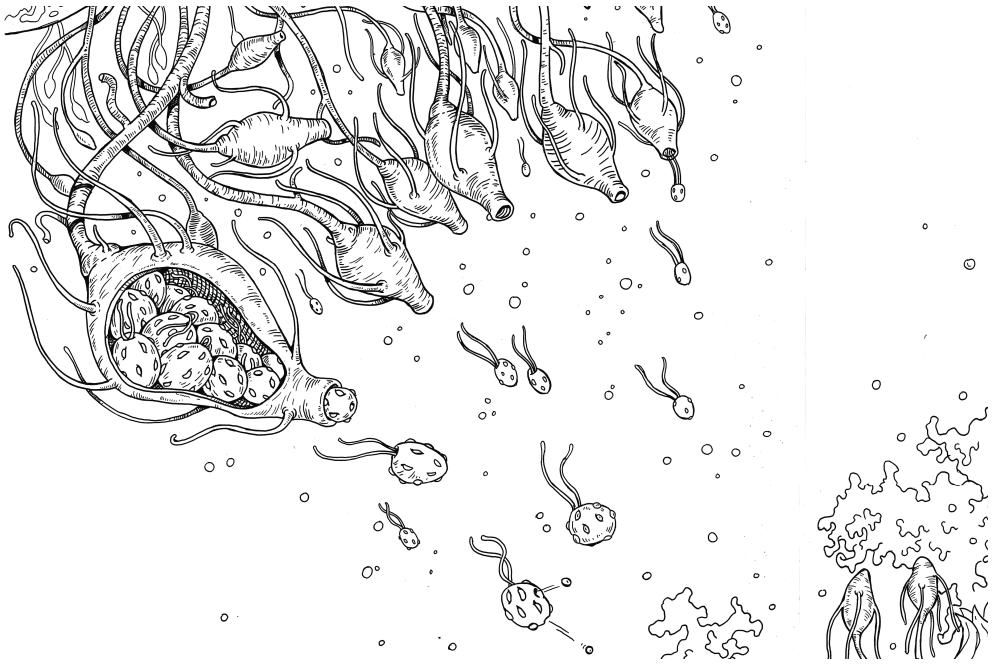
Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D.G., Boller, T. and Felix, G. (2006) Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts *Agrobacterium*-mediated transformation. *Cell*, **125**, 749–760.

Chapter 2

Divergent evolution of PcF-like and SCR74 effectors in oomycetes and their recognition in Solanaceae plants

Xiao Lin, Shumei Wang, Laura de Rond*, Roland Wouters*, Nicoletta Bertolin*, Joe Win, Emmanouil Domazakis, Mulusew Kassa Bitew, Suomeng Dong, Richard G.F. Visser, Paul Birch, Sophien Kamoun and Vivianne G. A. A. Vleeshouwers

***Authors contributed equally to this work**



Manuscript in preparing

Abstract

In the first phase of colonizing the host plant, pathogens secrete effectors into the apoplast. These apoplastic proteins often represent small-cysteine rich (SCR) proteins, and can be perceived by surface immune receptors of the hosts to trigger immune responses. Some apoplastic effectors have features of microbe-associated molecular patterns (MAMPs), but the function and evolution of most of the SCR proteins remain unclear. Here we report on the evolution of a family of oomycete apoplastic effector proteins, that includes PcF from *Phytophthora cactorum* and the SCR74 family from *Phytophthora infestans*. We found that *PcF* is highly conserved in *P. cactorum* isolates and under balancing selection. The *PcF* orthologs from different *Phytophthora* species share a co-linear genomic architecture and PcF is widely recognized by various Solanaceae species and even strawberry. In contrast, the SCR74 family from *P. infestans* is highly diverse, and *scr74* genes reside in a gene-sparse and transposon-rich region of the *P. infestans* genome. SCR74 variants are specifically recognized by diverse genotypes of *Solanum* section Petota, and single amino acid changes of SCR74 show altered recognition specificity. Our results demonstrate that a family of oomycete effectors appears to contain typical MAMPs as well as effectors under diversifying selection. This suggests that some members may have retained their ancestral function, whereas other members have started co-evolving with host genes. This finding provides new insights in the evolution of effectors in relation to their host plants.

Introduction

The plant apoplast is the battlefront of the plant-pathogen interaction. Plants deploy antimicrobial proteins, proteases, protease inhibitors and surface immune receptors to perceive pathogen proteins and protect themselves against disease (Jashni *et al.*, 2015). Conversely, pathogens secrete an arsenal of effector proteins, including proteases and protease inhibitors for facilitating their infection and manipulating the plant immune system. In oomycetes, a wide diversity of apoplastic proteins have been characterized, for example XEG1 (Ma *et al.*, 2015), elicitins (Kamoun *et al.*, 1997), necrosis- and ethylene-inducing proteins (Nep1) proteins (NLPs), Nep1-like proteins NPP1 and PsojNIP (Fellbrich *et al.*, 2002; Qutob *et al.*, 2002; Oome *et al.*, 2014), serine and cysteine enzyme inhibitors EPI1, EPI10, EPIC1 and EPIC2B (Tian *et al.*, 2004; Tian *et al.*, 2005; Tian *et al.*, 2007), and small cysteine-rich (SCR) proteins PcF, SCR74 and SCR91 (Orsomando *et al.*, 2001; Bos *et al.*, 2003; Liu e*t al.*, 2005). Some of the secreted oomycete proteins represent typical microbe-associated molecular patterns (MAMPs), which are generally conserved, widespread, abundant and essential molecules from pathogens that contribute to the general fitness of pathogens (Pel and Pieterse, 2013). In contrast, other apoplastic proteins are species- or strain-specific fast evolving molecules that help to manipulate the host immune system to facilitate their infection, according to the concept of effectors (Kamoun, 2007). However, with the growing evidence, the dichotomy of MAMPs and effectors is getting obscured (Thomma *et al.*, 2011; Cook *et al.*, 2014).

PcF (*Phytophthora cactorum-fragaria*) is a 5.6-kDa small cysteine-rich protein of 73-amino-acids, which forms 3 disulfide bridges by 6 cysteines on the mature protein. It was isolated from culture filtrate of *Phytophthora cactorum*, which triggers a defense-related response on strawberry and tomato (Orsomando *et al.*, 2001; Orsomando *et al.*, 2003). Up to date, only two *PcF* homologs from *P. cactorum* isolates have been isolated, with only one single nucleotide polymorphism (SNP) on the signal peptide (Orsomando *et al.*, 2001; Chen *et al.*, 2016). The NMR structure of PcF was solved and a novel helix-loop-helix motif was identified. The structure has a good fold match with the plant pollen allergenic protein Ole-e-6, therefore, a molecular mimicry hypothesis was proposed (Nicastro *et al.*, 2009). More PcF domain-containing (Pfam: 09461) SCR proteins have been identified afterwards. For example, SCR74 and SCR91 were identified from expressed sequence tags (EST) of multiple *P. infestans* strains (Bos *et al.*, 2003; Liu *et al.*, 2005). *Scr74* belongs to a highly polymorphic gene family in *P. infestans*. So far, 21 *Scr74* genes have been sequenced from different *P. infestans* strains. The sequences are highly diverse and they were shown to be under strong positive selection pressure. The expression of *Scr74* is up-regulated after infection of tomato

(Liu *et al.*, 2005). These findings suggest that SCR74 variants are important effectors in *P. infestans* and we hypothesize that they may have co-evolved with host factors, notably receptor genes. SCR96 is also from *P. cactorum* and related to PcF, but without a typical PcF domain, and triggers cell death responses in some Solanaceae species, including *Nicotiana benthamiana* and tomato (Chen *et al.*, 2016).

In this study, we retrieved all known members of the PcF/SCR family from public databases. We studied the phylogeny, sequence polymorphism and genomic architecture of the PcF/SCR family to determine whether they are conserved or diverse. Using mutagenesis, we determined the role of cysteine residues in the SCR74 protein. Sub-cellular localization was used to visualize the secretion of SCR74-mRFP in stable *P. infestans* transformants. Disease tests were used to evaluate whether the SCR74 protein contributes to the virulence of *P. infestans* in *N. benthamiana* and potato. To study the recognition of PcF/SCR proteins at plant family level, we performed a large scale effectoromics screening on Solanaceae species, including potato, tomato, eggplant, pepper and *Nicotiana* genotypes. Our finding provides a new insight in how apoplastic effectors can co-evolve with their host.

Results

PcF/SCR effectors from oomycetes

To study the PcF/SCR family from oomycetes, 57 PcF domain (PF09461) containing proteins were obtained from InterPro, by searching for the keyword "PcF", or Pfam domain ID PF09461. The PcF/SCR proteins were only detected in oomycetes, including *Hyaloperonospora arabidopsidis* (2), *Phytophthora cactorum* (2), *Phytophthora capsici* (1), *Phytophthora parasitica* (16), *Phytophthora ramorum* (1), *Phytophthora sojae* (4) and *Phytophthora infestans* (31). 11 redundant PcF-like proteins were removed. The remaining 46 PcF/SCR proteins were renamed by the species abbreviation and the number of amino acids of the full-length protein **(Supplementary table S1)**.

The PcF domains of the 46 PcF/SCR proteins were subjected to sequence alignment by MAFFT and a NJ tree was generated **(Figure 1)**. Due to reticulate sequence exchange events that might have happened in this family (Liu *et al.*, 2005), a network analysis was also made to reflect the phylogeny of this family **(Figure 2)**. SCR96 from *P. cactorum* was included as an outgroup.

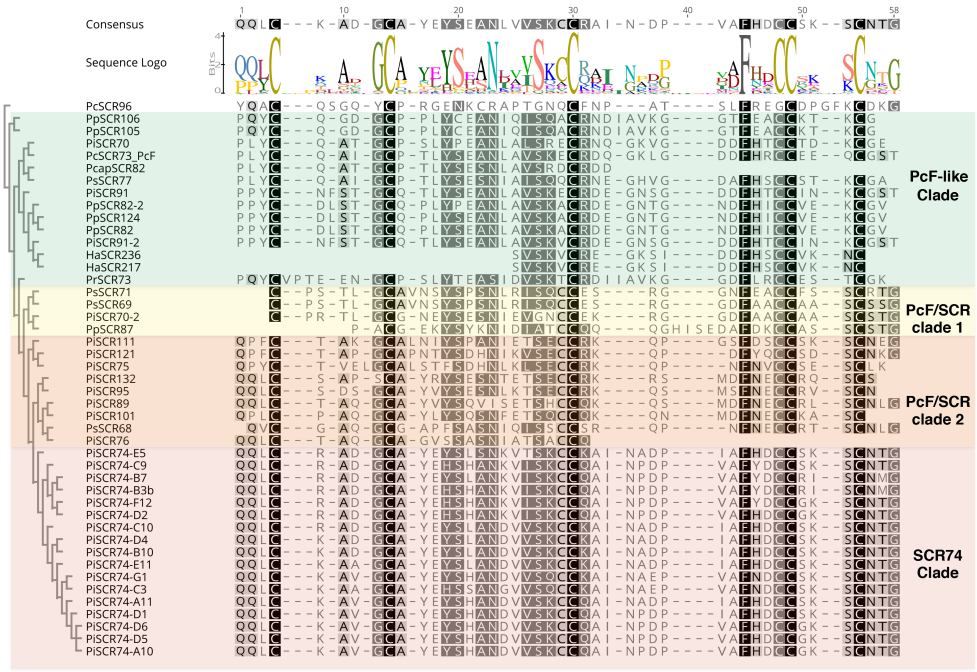


Figure 1. Alignment of the PcF domain (PF09461) of 46 PcF/SCR proteins.

The 46 PcF/SCR proteins are classified into 4 clades, i.e. a PcF-like clade, PcF/SCR clade1, PcF/SCR clade2 and SCR74 clade. The majority of the consensus amino acids and the sequence logo are shown on the top. In addition to the six conserved cysteines, other conserved motifs include YSx(A/S)N and (V/I)SK. A highly variable motif between position 31-43 aa in the PcF domain seems diagnostic for the four clades.

Based on the alignment and network analysis, the PcF/SCR proteins were classified into 4 clades, i.e. a PcF clade, a SCR74 clade and PcF/SCR clade 1 and 2, respectively. **(Figure 1 and Figure 2)**. The two PcF/SCR proteins from *H. arabidopsidis* only have a partial PcF domain, but we also classified them into the PcF-like clade. Except SCR96, all the full length PcF/SCR proteins contain 6-8 highly conserved cysteines. Proteins of the SCR74 clade and PcF/SCR clade 1 and 2, have an extra cysteine residue at amino acid position 28, which differentiates them from the members of the PcF-like clade. Besides the cysteine residues, the PcF/SCR protein family shares some other conserved motifs, like Y/HSxS/ANXXI/VSQ/K motif of 18-27 aa. Furthermore a highly variable region from amino acid position 31 to 43 can also diagnostic for the four clades, for example, the SCR74 clade shares an AINA/PDPV/IA motif in this variable region, which is more diverse compared to the other clades **(Figure 1)**.

Our data also shows that the PcF-like clade contains proteins from many oomycete species. PcF/SCR Clade 1 only contains proteins from *P. infestans*, *P. sojae* and *P. parasitica*, and PcF/SCR Clade 2 only from *P. infestans* and *P. sojae*. Noteworthy, the SCR74 clade consists only of SCR74 variants from *P. infestans*. This finding suggests that SCR74 proteins are specifically evolving in *P. infestans*.

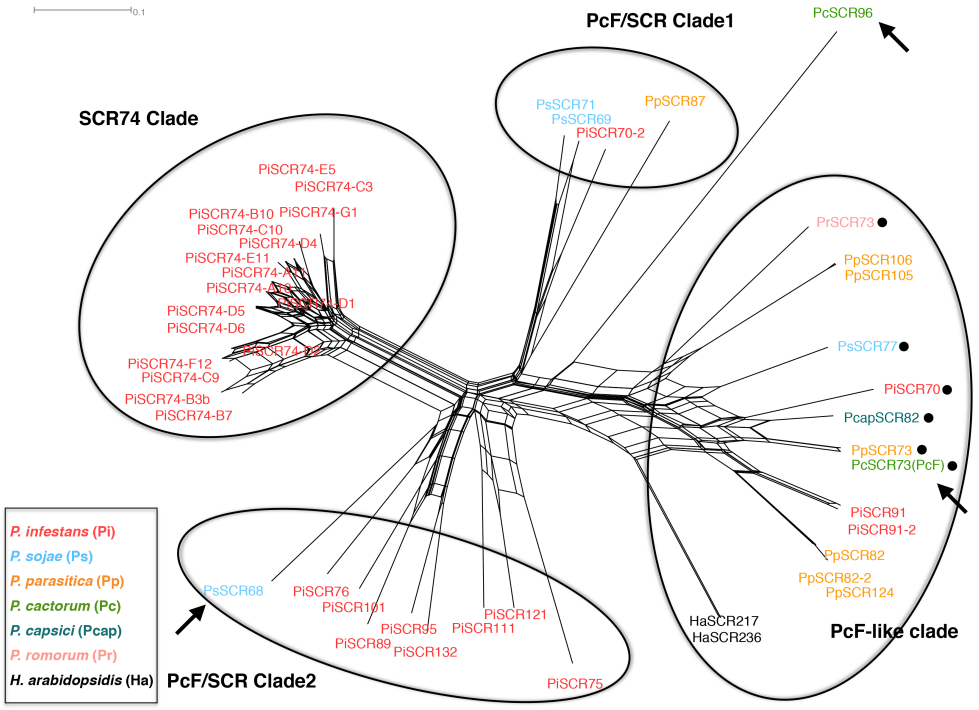


Figure 2. Phylogenetic network of PcF/SCR proteins.

PcF domain of these proteins were aligned by MAFFT v7.309, the phylogeny network was made by SplitTree4, computed by the NeighborNet method and the tree was displayed using the EqualAngle algorithm. The proportion of different positions between two sequences was computed by UncorrectedP method. 46 PcF/SCR proteins are shown from 6 *Phytophthora* species and *Hyaloperonospora arabidopsidis*. They are marked by different colors that are representative of the species (indicated in the box at the left corner). The four clades are marked by circles. PcF orthologs are marked with black dots (see **Figure 4**). 13 variants from the SCR74 clade, and 3 PcF/SCR variants (PsSCR68, PcSCR96 and PcF, marked by arrows) were used in effectoromics screening **(Figure 7)**.

PcF is conserved in P. cactorum and PcF loci are conserved in oomycetes

To study if *PcF* genes are similar to *Scr74* family, *PcF* genes from 9 *P. cactorum* isolates, originating from USA or Europe, were amplified by PCR, cloned into pGEM-T Easy vector and sequenced **(Material and Methods)**. The sequence alignments indicate that *PcF* genes are highly conserved in *P. cactorum* isolates from different geographic locations **(Figure 3A)**.

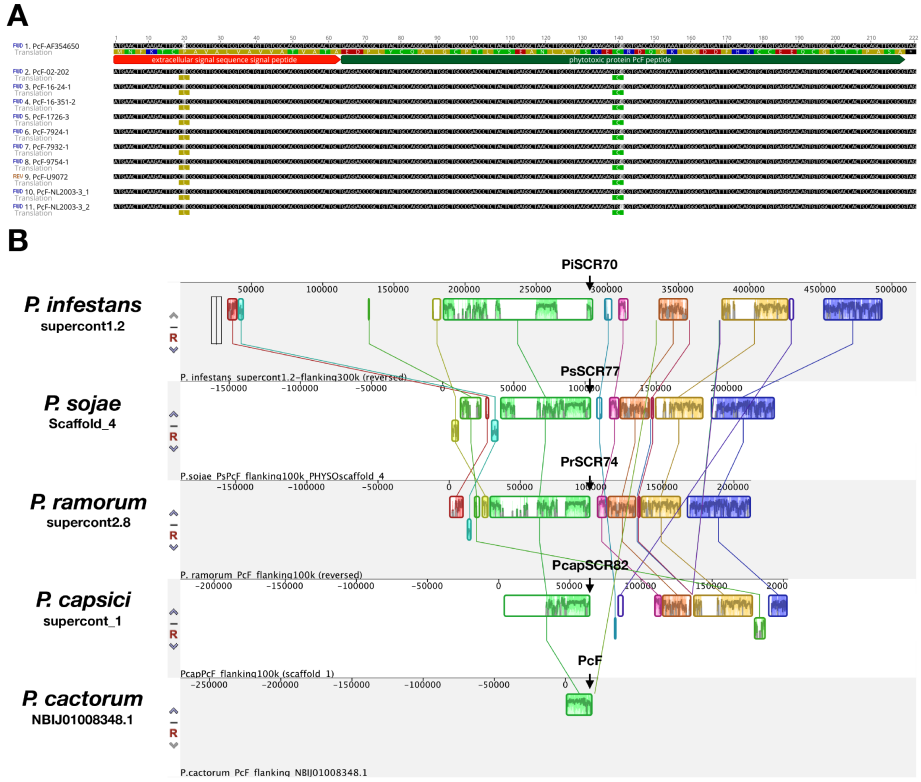


Figure 3. *PcF* loci are conserved in different oomycetes and *PcF* genes are conserved in different *P. cactorum* isolates.

A. Alignment of *PcF* genes of nine *P. cactorum* isolates from USA and Europe.
B. 500kb *PcF* flanking sequences from *P. infestans* and 200kb *PcF* flanking sequences from *P. sojae*, *P. ramorum*, *P. capsici* and a short contig from *P. cactorum* were aligned by Mauve. Regions of significant synteny are displayed as colored locally collinear blocks (LCBs) based on Mauve's progressive algorithm. The LCBs are connected by colored lines among the species. The *PcF* orthologs (*PiSCR70*, *PsSCR77*, *PrSCR74*, *PcapSCR82* and *PcF*) from the 5 *Phytophthora* species are shown with black arrows.

Only one nonsynonymous mutation was found in the predicted signal peptide of PcF-AF354650, and another synonymous mutation was found in the effector domain of NL2003-3. The amino acid sequence of the effector domain was fully conserved for all identified PcF homologs. Our results indicate that, unlike the *Scr74* family, *PcF* genes are highly conserved and under balancing selection in *P. cactorum* strains from different locations (USA and Europe), which fits the characteristics of MAMPs/PAMPs.

To further study whether *PcF* loci are conserved in diverse *Phytophthora* species, we extracted the *PcF* loci and the flanking 250kb region from the genome of *P. infestans*, 100kb flanking sequence from *P. sojae*, *P. ramorum*, *P. capsici* and a short contig containing *PcF* from *P. cactorum*. Mauve (v2.4.0) was used to align the *PcF* loci **(Figure 3B)**. This result shows a colinear structure of *PcF* loci in *Phytophthora* species, which is also in line with characteristics of MAMPs/PAMPs. Considering that these *Phytophthora* species cover the breadth of diversity of the genus, we postulate that the *PcF* gene is an ancient and fairly conserved gene in *Phytophthora*.

SCR74 is highly diverse in P. infestans and located in repeat-rich regions of the genome

Unlike PcF, SCR74 proteins were reported to be highly diverse and under strong positive selection pressure, based on 21 *scr74* variants from 8 *P. infestans* strains (Liu *et al.*, 2005). With the increased amount of NGS data, we re-evaluated the sequence diversity of 51 *P. infestans* isolates present in the public sequence databases and 2 *P. infestans* isolates sequenced in this study **(Supplementary Table S2)**. Our observation supports previous findings, that: 1) *Scr74* genes are present in all sequenced *P. infestans* isolates; 2) the sequence of *scr74* genes are highly diverse and some amino acids are under strong selection pressure; 3) the cysteine residues are conserved in all tested SCR74 proteins **(Figure 4)**.

To test whether the cysteine residues are important for the PcF/SCR proteins, we synthesized two SCR74-B3b mutants: SCR74-synB3b-27A and -47A, the cysteines were replaced by alanine in position 27 and 47, respectively. The codons were optimized for *in planta* expression, and we also included a synthesized SCR74-synB3b as positive control. We cloned them into PVX vector PGWC-PVX and transformed them to *Agrobacterium* strain GV3101. The constructs were tested on SCR74-B3b responding genotype *Solanum microdontum spp. gigantophyllum* (GIG362-6) by PVX agro-infection. The wild-type SCR74-B3b, pGR106-CRN2 and empty pGR106 vector were used as positive and negative controls. Our result

shows that both wild type and synthesized SCR74-B3b triggered cell death response in GIG362-6, but the two mutants caused no symptoms **(Supplementary Figure S1)**. This observation indicates that these cysteines are important in these PcF/SCR proteins for recognition by the plant, likely because of lower stability or misfolding of the mutated SCR74 protein.

To study the genomic architecture of *scr74* genes in *P. infestans* reference genome, we extracted three *scr74* containing supercontigs 1.36,1.73 and 1.4 **(Figure 5A)** from the *P. infestans* reference genome. There are 3 *scr74* homologs, including a pseudogene in supercontig 1.4. By comparing the flanking region of these *scr74* loci, we found the *scr74* genes and the flanking regions (≈2kb) from supercontigs 1.73 and 1.4 with a high level of identity. This observation indicates that there was a translocation event of the *scr74* loci, which might be driven by gypsy transposons surrounding these *scr74* genes.

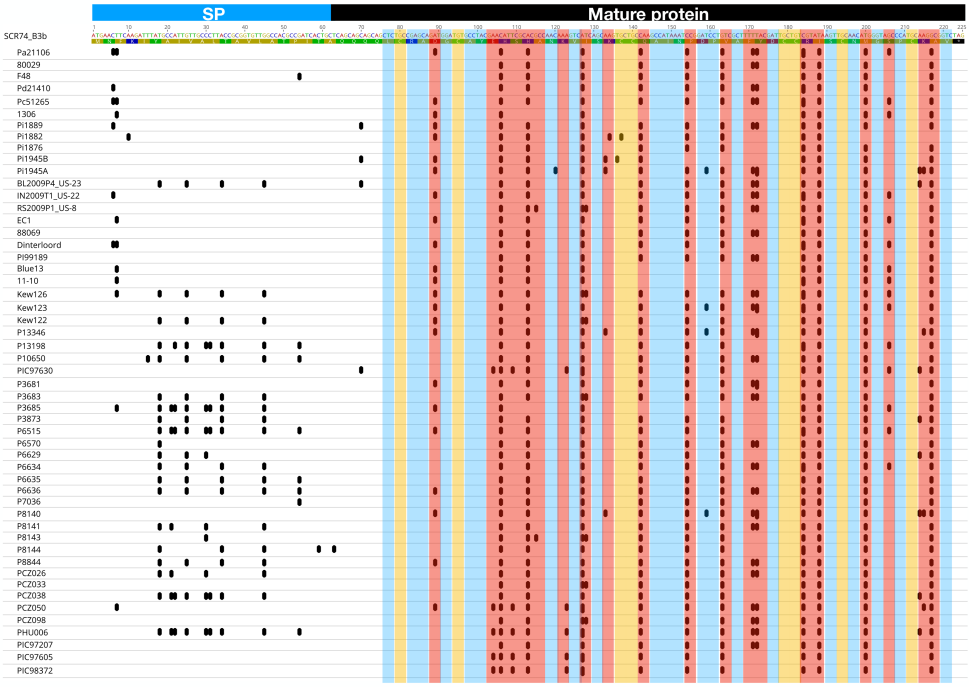


Figure 4. Polymorphism of SCR74 genes from 53 sequenced *P. infestans* isolates.

The sequencing reads from 53 *P. infestans* isolates were mapped to SCR74-B3b sequence, the SNPs are shown as black dots. The protein consists of a 21 amino-acid signal peptide (SP, blue bar), and a 53 amino-acids mature protein (black bar). The cysteine residues are highlighted in yellow, they are conserved in most of the variants. Other conserved amino-acids with no change or only synonymous mutation are highlighted in blue. The highly diverse amino-acids with nonsynonymous mutations are highlighted in red.

Based on the "two speed genome" model, rapidly evolving effectors tend to be located in gene-sparse repeat-rich regions (GSR) of the pathogen genome (Dong *et al.*, 2015). In order to test if *scr74* genes are located in GSR, 2 *scr74* homologs (PITG_14645 and PITG_18592) from the *P. infestans* reference genome (T30-4) were mapped to the flanking intergenic region heatmap, as in Haas et al (Haas *et al.*, 2009). As expected, the 2 *scr74* genes are located in gene sparse regions **(Figure 5B)**.

In summary, different clades of the PcF/SCR74 family have opposite evolutionary patterns. Our results support the previous finding that the SCR74 clade is highly diverse and under strong positive selection pressure. We also found that the *scr74* genes are located in the fast evolving GSR parts of the genome surrounded by gypsy transposons, which may enable accelerated evolution of *scr74* genes. Thus, the SCR74 clade seems to contain fast-evolving apoplastic effectors in *P. infestans*, whereas the PcF clade seems more conserved and similar to MAMPs/PAMPs.

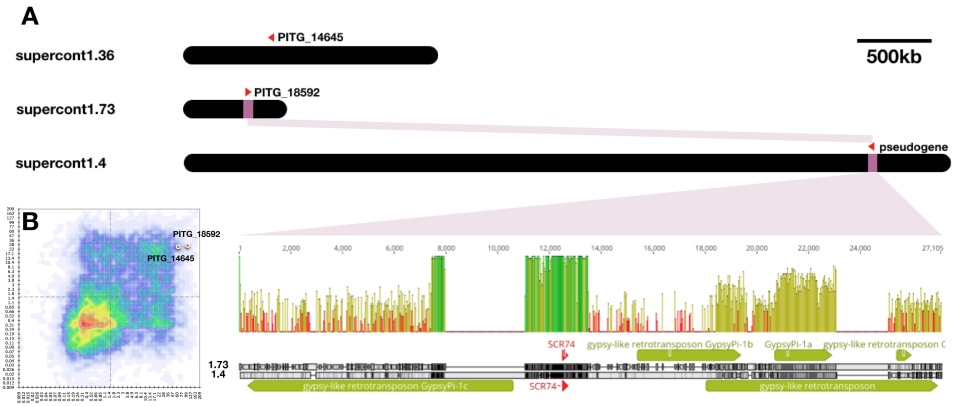


Figure 5. Physical position of SCR74 homologs in the genome of *P. infestans* strain T30-4

A. 3 *scr74* homologs (PITG_14645, PITG_18592 and a pseudogene are from supercontigs 1.36, 1.73 and 1.4 respectively, marked by red arrows. 15kb regions from supercontig 1.73 and supercontig 1.4 were extracted for alignment. The *scr74* genes and the flanking 3kb show co-linearity in these two supercontigs. *Scr74* homologs are surrounded by gypsy retrotransposons.

B. The distance between flanking genes were plotted in a heatmap, the Y axis represents the 5' intergenic distances and X axis the 3' intergenic distances. The gene density is shown by different colors. Two *scr74* homologs (PITG_14645 and PITG_18592) are located on fast evolving gene sparse regions (GSR).

SCR74-B3b is secreted into infected host apoplastic space at P. infestans haustoria

To investigate the site of secretion of SCR74-B3b *in planta*, *P. infestans* transformants were generated with constitutively expressed free GFP in the cytoplasm, and stably expressed either SCR74-B3b-monomeric red fluorescent protein (mRFP) or SCR74-B3b-27A-mRFP fusion proteins under control of the constitutive Ham34 promoter. The stability of fusion proteins was tested using immunoblotting of mycelium (M) and culture filtrate (CF) fraction. Western blots show that both fusion proteins were secreted into CF **(Figure 6)**. GFP as a cellular indicator was detected only in

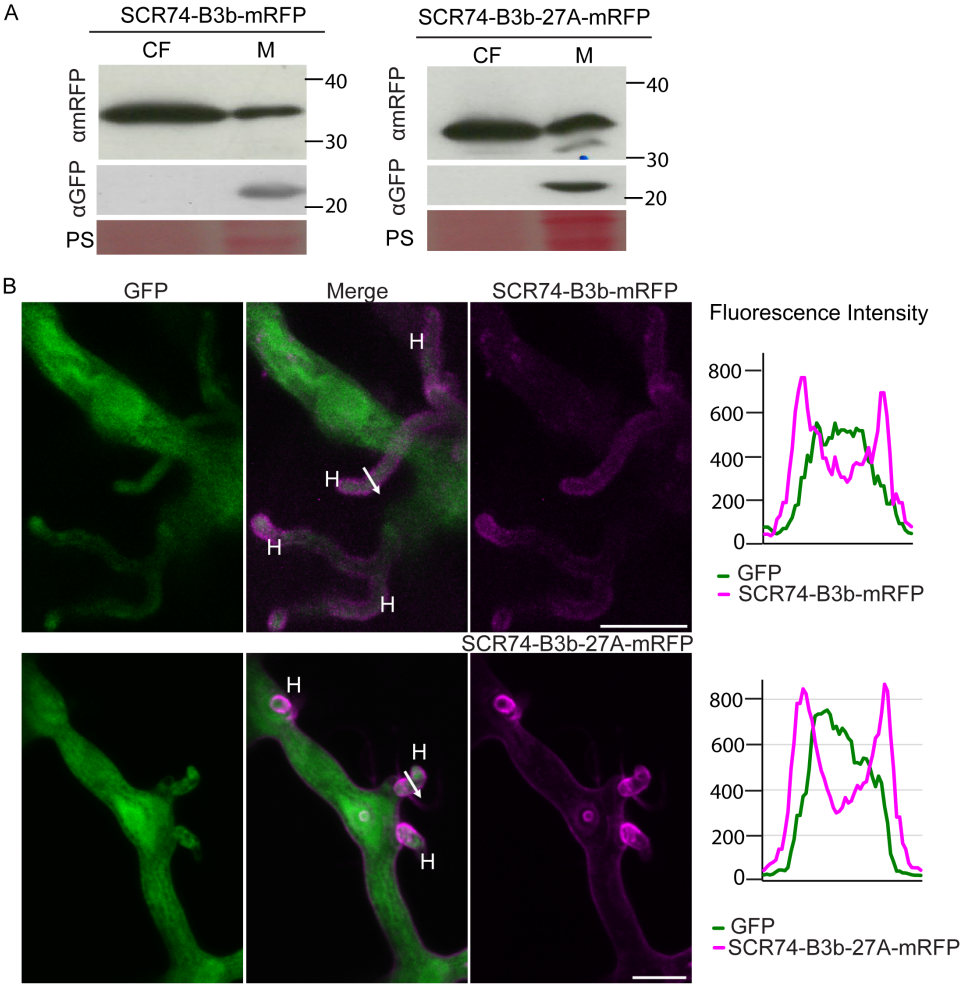


Figure 6. ***Phytophthora infestans*** **apoplastic effector SCR74-B3b and cysteine mutant SCR74-B3b-27A are secreted from haustoria.**

A. The expression of SCR74-B3b-mRFP and SCR74-B3b-27A-mRFP were confirmed in the mycelium (M) and in the culture filtrate (CF) using immunoblotting with αmRFP antibody, αGFP primary antibody was used to detect intercellular GFP protein to show there was no leakage in the CF with cellular proteins. Ponceau stain (PS) was used as protein loading. Protein size markers are indicated in kDa.
B. Confocal projections reveal that both fusion proteins of SCR74-B3b-mRFP and SCR74-B3b-27A-mRFP are secreted from haustoria (H) in infected tissues by *P. infestans* transformants expressing SCR74-B3b-mRFP and SCR74-B3b-27A-mRFP, respectively.

mycelium, suggesting that there was no detectable contamination in the CF **(Figure 6A)**. The transformed *P. infestans* strains were used to infect *N. benthamiana* leaves. Confocal microscopy revealed that both SCR74-B3b-mRFP and SCR74-B3b-27A-mRFP fusion proteins clearly accumulated at haustoria and in the host apoplastic space **(Figure 6B)**. This indicates that haustoria are the main secretion sites for this apoplastic effector.

Functional screening of PcF, SCR68, SCR96 and SCR74 in Nicotiana, potato, tomato pepper and eggplant

To bridge the gap between the sequence analysis and the function of these PcF/SCR proteins, we performed an effectoromics screening in Solanaceae plants. Tested plants included 245 genotypes of wild Solanaceae species, i.e. 206 wild *Solanum* section *Petota*, 23 tomato, 7 eggplant, 10 pepper and 8 *Nicotiana* genotypes. The effectors PcF and SCR96 from *P. cactorum*, SCR68 from *P. sojae* and 13 SCR74 variants from *P. infestans* were selected for the functional screening **(Figure 2, Supplementary Table S3)**. Effectors were synthesized or cloned into PVX vector pGWC-PVX or pGR106, and transformed into *Agrobacterium tumefaciens* strain GV3101 for transient expression **(Materials and Methods)**. The *Agrobacterium* clones carrying single *PcF/SCR* genes were infected by toothpick inoculation. pGR106-CRN2 was used as positive control and the empty pGR106 vector was used as negative control. At least 6 leaves from 3 plants were PVX-agroinfected per genotype. The symptoms were scored 12-14 days after infection, in the range of 0 to 10, reflecting no visible response up to strong response in all replicates, respectively. After removing genotypes that showed cell death to pGR106 treatment, or failed to show cell death to pGR106-CRN2 **(Materials and Methods)**, in the end, 4 *Nicotiana*, 1 pepper, 2 eggplant, 16 tomato and 136 potato genotypes were presented for their response to the effectors **(Figure 7 and Supplementary Table S4)**.

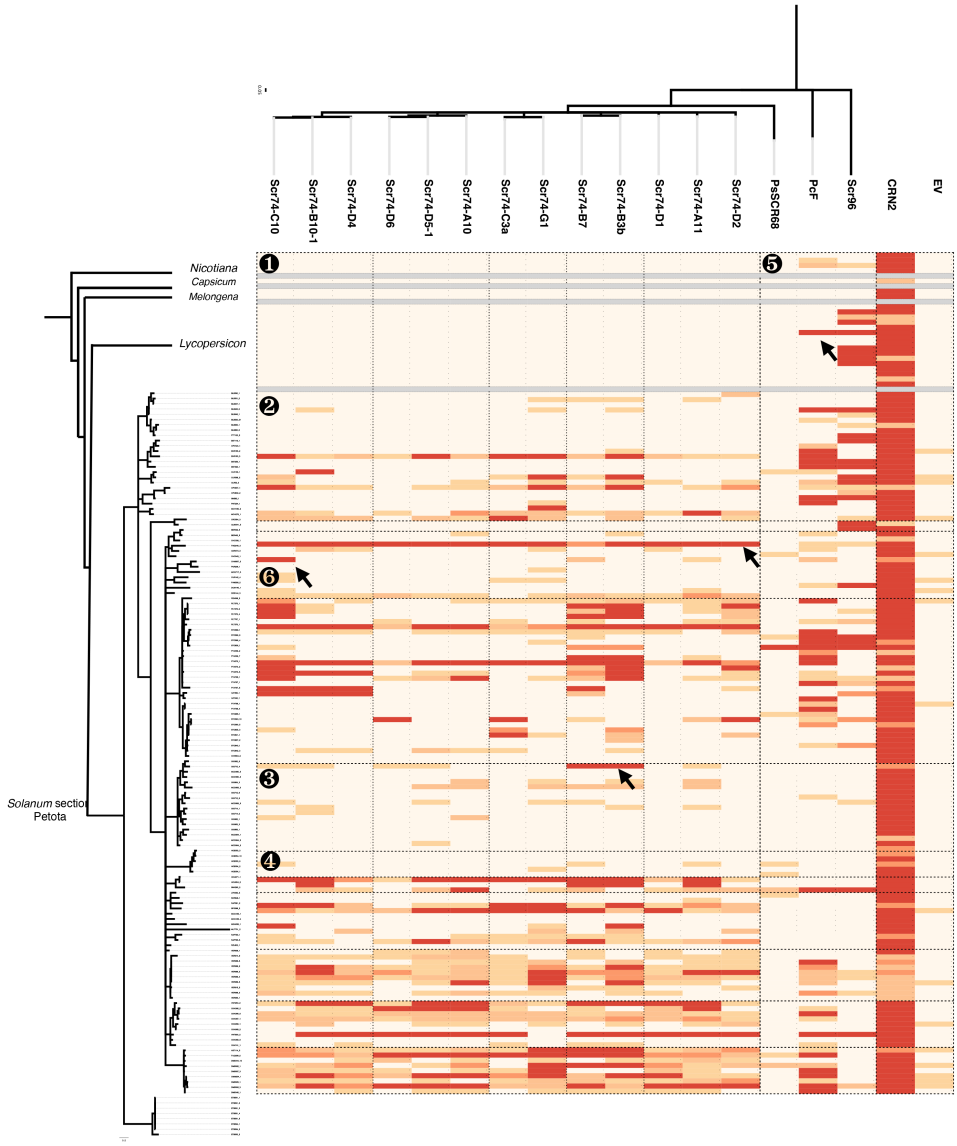


Figure 7. Recognition pattern of PcF, SCR68, SCR96 and 13 SCR74 variants in wild Solanaceae species.

The Bayesian tree of *Solanum* section *Petota* was generated based on previously produced AFLP data (Jacobs *et al.*, 2008). The phylogeny of other Solanaceae species is an illustration based on classical taxonomy (Bohs and Olmstead, 1997). For the PcF/SCR effectors, a NJ tree was made based on the PcF domain, SCR96 was used as outgroup. The intensity of cell-death response after PVX agro-infection of apoplastic effectors in leaves is represented by a heat map that ranges from dark red (strong response, score >8), dark orange (score 7-8), light orange (score 5-6) to beige (score 0-4). CRN2 and empty pGR106 vector were used as positive and negative control, respectively. The grey blocks are the spacer. The numbers and arrows are examples of specific responses, which are explained below.

Our data shows that cell death response to PcF and SCR96 from *P. cactorum* is widely present in Solanaceae species. Recognition was found in various species in *Solanum* section *Petota* **(Figure 7 ❺)**, as well as tomato and some *Nicotiana* accessions **(Figure 7 ❺**, black arrow**)**. Especially SCR96 triggers strong cell death in many tomato accessions. Unlike PcF and SCR96, SCR74 recognition is restricted to *Solanum* section *Petota*, and no response was noted in any other Solanaceae **(Figure 7)**. The pattern of responses does not seem to show any correlation to clade, species, or geographic origin, but the noted responses are highly specific. For example, in **Figure 7 ❸**, most genotypes from *Solanum microdontum* (MCD) and *Solanum microdontum* spp. *gigantophyllum* (GIG) do not recognize any SCR74 variants, except GIG362-6, which can specifically recognize SCR74-B3b and SCR74-B7. In contrast, some genotypes show response to all tested SCR74 variants, like *Solanum chacoense* CHC338-1 **(Figure 7 ❻** black arrow on the top, **Supplementary Figure S2)**, CHC338-1 also shows weak response to PcF and SCR96, but not to SCR68. SCR68 is not recognized by most Solanaceae species, we only found one specific response in *S. stoloniferum* genotype STO389-4 **(Figure 7 ❺)**.

Collectively, our results indicate highly diverse recognition specificities to these PcF/SCR effectors in *Solanum* plants, and the recognition to the SCR74 clade seems to be restricted to wild potatoes. To further explore the specificity of SCR74 recognition in wild potato, the phylogenic relation of the diverse SCR74 members was related to the heat map of responses in potatoes. Some of the SCR74 variants are in a same cluster, like SCR74-C10, -B10-1 and -D4; SCR74-D6, -D5-1 and -A10; SCR74-C3a and -G1; SCR74-B7 and -B3b; SCR74-D1, -A11 and -D2 **(Figure 7)**. First of all, for all individual SCR74 variants, at least one responding wild potato was identified. In most cases, the close SCR74 variants share a similar recognition pattern. For example, SCR74-C10, -B10-1 and -D4 that are highly similar **(Figure 1)** are specifically recognized by the same set of genotypes, e.g. CHC338-1, PLT789-6, PTA369-7, PTA370-2 **(Figure 7❶)**. Another example is SCR74-B3b and SCR74-B7, which only have one polymorphic amino acid **(Figure 1)**, share specific recognition of another set of *Solanum* genotypes that include PLT378-1, PTA369-2, PTA369-7, GIG362-6, HOU271-1, HOU654-2, VNT250-2, VNT896-2, AST114-5, DMS299-2, DMS582-1 **(Figure 7)**.

To check if one single amino acid can lead to an altered recognition specificity, SCR74-D4, -B4 and -C10 were used as example. These effectors are highly similar, with only one amino acid (L or F) difference on the signal peptide between SCR74-D4 and SCR74-B4, or one (K or R) at position 28 of the mature protein between SCR74-C10 and SCR74-B4 **(Figure 8A)**.

The structure of the mature protein of SCR74 was homology modeled and visualized based on the solution structure of PcF (2BIC) by SWISS-MODEL (Bienert *et al.*, 2017). The polymorphic amino acid resides on the loop of the protein **(Figure 8B)**. The effectoromics screening **(Figure 7❻)**, PTA767-1 and PLT378-2 showed a distinct recognition pattern to SCR74-D4, -B4 and -C10, and results were confirmed with an independent PVX agro-infection experiment for this subset of treatments. PTA767-1 shows a clear cell death response to all three SCR74 variants, but PLT378-2 can only respond to SCR74-C10 **(Figure 8C)**. These data show that one amino-acid change of the SCR74 variants may cause altered recognition in potato genotypes, which supports the hypothesis that SCR74 may have co-evolved with the plant receptor.

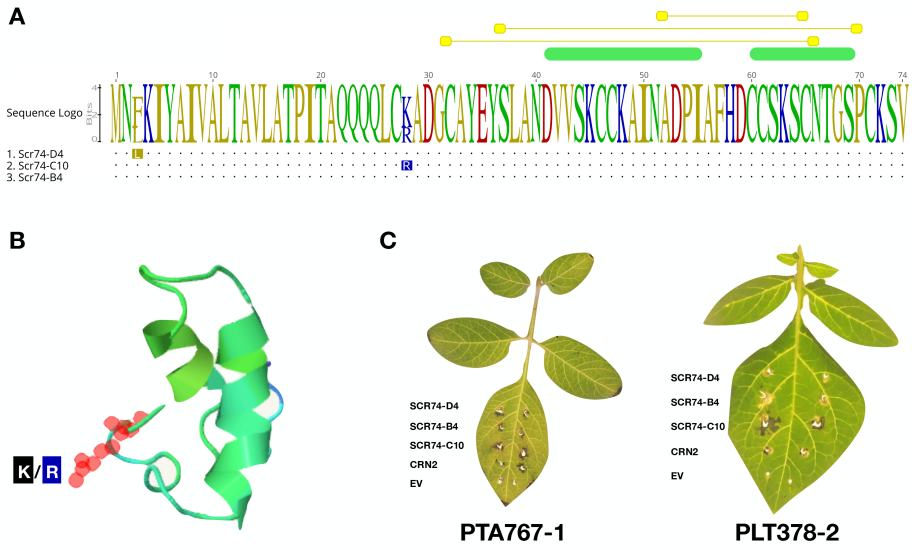


Figure 8. Single amino acid change of SCR74 leads to altered recognition specificity.

A. Protein alignment of SCR74-D4, -B4 and -C10. The S-S bridges are indicated by yellow bars and the predicted α-helix are shown by green bars.
B. Predicted structure of the mature SCR74 protein. The polymorphic amino acid between SCR74-D4, -B4 and -C10 at position 28 is shown.
C. PTA767-1 recognizes SCR74-D4, -B4 and -C10, whereas PLT378-2 can only recognize SCR74-C10.

Late blight resistance does not co-segregate with the SCR74 response

To test if the putative SCR74 receptor contributes to late blight resistance, we performed disease tests on a population that segregates to SCR74-B3b response. The F1 population 7026 generated by crossing GIG362-6 which responds to SCR74-B3b with the non-responsiveness genotype MCD360-1. The population was found to segregate to SCR74-B3b response in a 1: 1 ratio **(Chapter 3)**. To identify the *P. infestans* isolates carrying SCR74-B3b variant, we cloned and sequenced the SCR74 variants from 12 *P. infestans* isolates, and we found 4 *P. infestans* isolates IPO-C, PIC99177, PIC99183 and UK3928-A that carry SCR74-B3b **(Supplementary Table S5)**. Therefore, these isolates are suitable for testing if the SCR74 recognition is correlated with late blight resistance.

In this study, IPO-C was inoculated on the parents and 40 F1 progenies from population 7026. Of these, 20 respond to SCR74-B3b and 20 do not. Our results showed that the lesion size of GIG362-6 is slightly smaller than MCD360-1 **(Figure 9A)**, but the difference is not statistically significant (Tukey's multiple comparisons test, data not shown). The resistance to IPO-C in the F1 population segregated quantitatively **(Figure 9A)**. To test whether the SCR74-responsive progenies are more resistant, we grouped the data based on SCR74 responsiveness. The SCR74-B3b responsive group showed on average a slightly larger lesion size than the SCR74-B3b non-responsive group, which was statistic significant **(Figure 9B)**. However, in an independently repeated experiment there was no significant difference in lesion size between the groups (data not shown). These results do not support that SCR74-B3b responsiveness would correlate with the late blight resistance to *P. infestans* strain IPO-C. However, since the SCR74-B3b responsive and non-responsive parents also did not show a significant difference in the lesion size, additional *P. infestans* strains like PIC99177 and PIC99183 should be tested in the future.

To test if the SCR74-B3b will increase the virulence of *P. infestans* by transient expression, we expressed SCR74-B3b by agro infiltration and inoculated *P. infestans* zoospores on leaves of *N. benthamiana* and potato cultivar Bintje on day after inoculation. However, our result indicated that SCR74-B3b did not enhance the virulence of *P. infestans* (data not shown). We also tested the SCR74-B3b stable transformed *P. infestans* isolate UK3928A, by using a SCR74-B3b-27C1 transformed isolate and the wild type isolate as control. It was observed that isolate UK3928A transformed with SCR74-B3b did not increase the virulence in our disease test (data not shown).

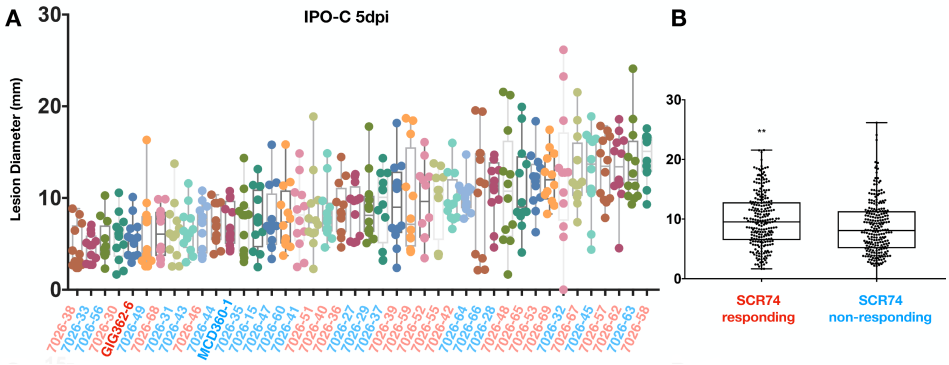


Figure 9. Disease test of plants from population 7026, obtained by crossing GIG362-6 with MCD360-1, after inoculation with *P. infestans* isolate IPO-C.

A. SCR74-responsive genotype GIG362-6 (dark red), non-responding genotype MCD360-1 (dark blue), 20 SCR74 responsive progenies (light red) and 20 SCR74 non-responsive progenies (light blue) were tested by detached leaf assay. 12 inoculation points were measured for each genotype. The lesion size is plotted by box chart, all the data points are shown, the whiskers of boxes indicate the minimum and maximum point. The genotypes are sorted based on the mean lesion size.
B. The lesion size of responsive (R) progenies are compared to the non-responsive (NR) progenies by T- test. The whiskers of boxes indicate the minimum and maximum point. The double asterisk indicate P<0.01.

In summary, current data do not support the hypothesis that SCR74 effectors increase the virulence of the pathogen. Also, SCR74 recognition was not found to correlate with enhanced late blight resistance.

Discussion

Here, we studied the evolution and function of the PcF/SCR protein family, which is only present in the oomycete plant pathogen *Phytophthora* and *Hyaloperonospora* species. We performed sequence analysis, phylogenetic analysis, mutagenesis, subcelullar localization and effectoromics-based *in-planta* functional screenings in Solanaceae plants. Sequence analysis of all the family members point to six highly conserved cysteines and a number of highly polymorphic amino acids. The disulfide bridges formed by the cysteines are considered to be important for the stabilization of these proteins (Orsomando *et al.*, 2011). Indeed, we found that SCR74-B3b cysteine mutants failed to induce cell-death response on the SCR74-B3b responsive plant GIG362-6, which supports the significance of the cysteines for protein stabilization in the harsh environment in the apoplast.

Based on further sequence and network analysis, PcF was found to fit most characteristics of MAMPs/PAMPs. *PcF* orthologs are wide-spread among *Phytophthora* species, have conserved genomic architectures and gene sequences, and they are under balancing selection. Altogether, this suggests a conserved function of PcF for the pathogen's fitness. In contrast, SCR74 is rather a fast evolving effector, exclusively present and expanded in *P. infestans.* By retrieving SCR74 sequencing reads from the sequence read archive (SRA), we found that SCR74 homologs are present in all sequenced *P. infestans* strains, and many amino acids are under strong positive selection pressure, in line with previous studies (Liu *et al.*, 2005). By plotting the flanking intergenic distance of SCR74 genes in the *P. infestans* reference genome FIR heat map, we found that *scr74* homologs are located in the gene sparse regions (GSR), similar to the RXLR effectors. Based on the "two speed genome" model (Dong *et al.*, 2015), the GSR is full of transposons and able to facilitate accelerated evolution (Raffaele and Kamoun, 2012; Dong *et al.*, 2015; Oliva *et al.*, 2015). Additionally, after analysis of the flanking sequences of *scr74* loci, we found a translocation event of this gene in one scaffold of *P. infestans* reference genomes. This translocation might be driven by the gypsy transposable elements surrounding these *scr74* genes, that can enable fast evolution. In sum, the apoplastic effector SCR74 seems to be a fast-evolving effector and displaying features that are typical of AVR proteins, whereas PcF seems to have retained a function resembling a MAMP/PAMP.

Our effectoromics screens further support this hypothesis, PcF is recognized by a wide diversity of plant species including many wild potatoes, tomatoes as well as strawberry (Orsomando *et al.*, 2001). On the other hand, SCR74 recognition is restricted in *Solanum* section *Petota*, and the diversity of recognition pattern suggests an arm-race with the corresponding host receptors. These data suggest that there are multiple different receptors present in the diverse genotypes, which can recognize different SCR74 variants. In addition, we found that even single amino acid changes lead to altered recognition. All data suggests that SCR74 is under strong positive selection pressure, which could be explained by a tight co-evolution with the host.

Taken together, our findings show that an apoplastic effector family may contain members that have characteristics of MAMPs/PAMPs as well as members that resemble fast-evolving effectors. The effectors might evolved from the conserved molecules upon co-evolution with their host, leading to blurring of the boundary between MAMPs/PAMPs and effectors (Thomma *et al.*, 2011) and fitting with the invasion model of plant immunity (Cook *et al.*, 2015) **(Figure 10)**. Cloning and functional characterizing the PcF receptor and multiple SCR74 receptors will help us to have a better understanding of this system.

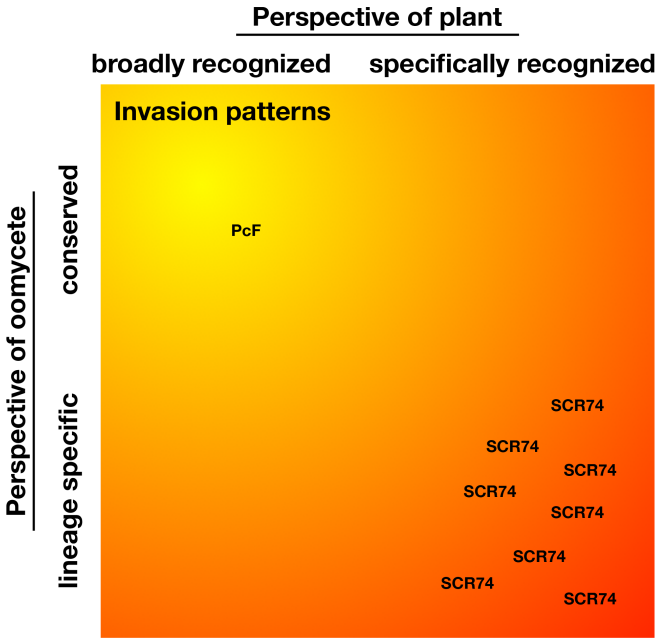


Figure 10. The invasion model for the PcF/SCR effectors. (Modified from Cook *et al.*, 2015)

Unlike the Zig-zag model (Jones and Dangl, 2006), which defined the MAMP and effectors strictly, the invasion model describe the plant immune system as a continuum. From yellow to red: conserved molecules from pathogens and broadly recognized in plants to lineage specific effectors and specifically recognized in some plant species. PcF lands on the conserved region and SCR74 variants lands on the fast-evolving and lineage specific region.

So far, the biological function of the PcF/SCR effectors remain unclear. In addition to cell death, PcF can induce the expression of pathogenesis-related (*PR*) genes and defence-related enzyme PAL21 in tomato (Orsomando *et al.*, 2001; Orsomando *et al.*, 2003). The structure of PcF was found to have similarity with a pollen allergenic protein Ole e 6, and therefore, a molecular mimic function was hypothesized (Nicastro *et al.*, 2009). Intriguingly, we also found that SCR74 shares many common features with a small cysteine-rich protein SCR/SP11 in the self-incompatibility (SI) system in *Brassica* species, which is acting as ligand of the pistil receptor SRK. The interaction between SCR/SP11 and SRK will culminate in a defense response, which is leading to self-incompatibility (Kachroo *et al.*, 2001). Similar to *scr74* genes in *P. infestans*, SCR/SP11 are highly diverse and under strong positive selection pressure in plants. To see if the SCR74 can enhance the virulence of *P. infestans*, like some RXLR effectors such as AVR2 (Turnbull *et al.*, 2017), several attempts were made in this study. However, our pilot experiments do not support this feature for SCR74, but it would be interesting to explore this further.

Materials and Methods

Phylogenetic analysis of PcF/SCR proteins

All the 46 PcF domain containing proteins (InterPro: IPR018570) were obtained from InterPro (https://www.ebi.ac.uk/interpro/). The protein sequences were aligned by MAFFT v7.309 (Katoh and Standley, 2013) and Geneious R10 (http://www.geneious.com) (Kearse *et al.*, 2012). The redundant sequences were removed manually based on the alignment outputs, the Neighbor-Joining tree was performed by Geneious R10, using Jukes-Cantor model. The phylogeny network was made by SplitTree4 (Huson and Bryant, 2006), computed by the NeighborNet method and the tree was displayed using the EqualAngle algorithm. The proportion of different positions between two sequences were computed by UncorrectedP method.

Genome data and sequence analysis

The oomycete genomes were obtained from EnsemblProtists (http://protists.ensembl.org/) or JGI genome portal (https://genome.jgi.doe.gov), including *P. infestans* (ASM14294v1) (Haas *et al.*, 2009), *P. sojae* (*P. sojae* V3.0), *P. ramorum* (ASM14973v1) (Tyler *et al.*, 2006), *P. capsici* (LT1534 v11.0) (Lamour *et al.*, 2012). The draft genome of *P. cactorum* strain LV007 was obtained from GenBank (NBIJ01000000) (Grenville-Briggs *et al.*, 2017). All genomes, genomic annotations, predicted genes, cDNAs and proteins were imported into Geneious R10 for local Blast (Blast+ v2.6.0) or genome visualization. The genome alignment was performed by Mauve with progressive Mauve algorithm, with match seed weight 15 and minimum LCB score 30,000 (Darling *et al.*, 2004).

57 publicly available sequencing data of *P. infestans* were used for studying the genetic diversity of SCR74 in *P. infestans* **(Supplementary Table S2)**. These data include some historical samples like Pi1845A, Pi1845B, Pi1882, Pi1876, Pi1889 (Martin *et al.*, 2013) and some hybrids *P. andiana* EC3394, PaX and P13803 (Martin *et al.*, 2016). SCR74-B3b (GenBank: AY723717.1) was used as query for obtaining the reads based on each sequence read archive (SRA) database from NCBI. The SAM format outputs were imported into Geneious R10, SCR74-B3b was used as reference sequence. Then the Find Variation/SNPs function from Geneious R10 was used for detecting the SNPs from each individual isolates.

Isolates and DNA isolation of Phytophthora cactorum

Phytophthora cactorum isolates **(Supplementary Table S6)** from USA or Europe were used in this study. The isolates were routinely cultured on V8 medium and stored at 20˚ C in the dark. One liter of V8 medium contains 200 mL of V8-vegetable juice, 800 mL tap water, 20 g of agarose (n°3) and 3 g of CaCO3. Genomic DNA was extracted from *Phytophthora cactorum* mycelia by DNeasy Plant Mini Kit (QIAGEN). The *PcF* genes were amplified by primer PcF-F: ATGAACTTCAAGACTTGCC and PcF-R: AAACAGTCTACGCGGAAG with Phusion PCR high-fidelity DNA polymerase. After A-tailing by DreamTaq DNA polymerase, the *PcF* genes were ligated into pGEM-T Easy Vector system (Promega) and transformed to *E. coli* DH5α competent cells. The plasmid of positive clones were isolated by QIAprep Spin Miniprep Kit (QIAprep® QIAGEN), then sent for sequencing (GATC). The sequence analysis and alignment were performed by Geneious R10.

Plant materials

Seeds of 24 tomato accessions belonging to 12 species and 1 hybrid (*Lycopersicum x cheesmaniae*), 10 pepper accessions of 9 species and 7 eggplant accessions of 7 species were obtained from the Centre for Genetic Resources, Wageningen, the Netherlands (CGN). The potato genotypes were from the *in vitro Solanum* collection of Plant Breeding at Wageningen Univeristy & Reserch. The potatoes were propagated *in vitro*, they were grown for 2 weeks in a climate room at 22˚C before moving to the greenhouse. They were further grown for 4-6 weeks in the greenhouse before the PVX agroinfection.

PVX agro-infection

The 13 SCR74 homologs (Liu *et al.*, 2005) were cloned into PVX vector pGR106, PcF, PsSCR68 and PcSCR96 were synthesized by Genewiz (https://www.genewiz.com/) and cloned into a modified gateway compatible pGWC-PVX vector. The effectors were then transformed into *Agrobacterium tumefaciens* strain GV3101. Before PVX agro-infection, the *Agrobacterium* carrying the target effectors were grown 2 days in liquid LB medium, then 2 days in solid LB medium. The *Agrobacterium* were collected by sterilized spreading rod and toothpicks were used for the agro-infection (Du *et al.*, 2014).

Mutagenesis

Two SCR74-B3b cysteine mutations (SCR74-B3b-27A and SCR74-B3b-47A) together with the wild type SCR74-B3b were codon optimized and synthesized by Genewiz, and cloned into pGWC-PVX. They were transformed into *Agrobacterium* and used for the PVX agro-infection.

Immunoblotting

Phytophthora infestans transformants expressing either SCR74-B3b-mRFP or SCR74-B3b-27C-mRFP were cultured in amended lima bean (ALB) liquid medium (Bruck *et al.*, 1981). *In vitro* grown mycelium was harvested by centrifugation at 2 dpi, then transferred into 1 ml of ALB liquid medium. The culture filtrate (CF) was retained separately after 24 h incubation, four times the sample volume of cold (-20˚C) acetone (Thermo Fisher Scientific, Loughborough, UK) was used to precipitate proteins overnight. Protein was precipitated at 10,000 g for 10 min (Wang *et al.*, 2017). *P. infestans* mycelium and resuspended CF were mixed in 2x SDS PAGE loading buffer (100 mM Tris-HCl (pH 6.8), 200 mM DTT, 4 % (w/v) SDS, 0.2 %(w/v) Bromophenol Blue, 20 % (v/v) glycerol). Samples (10 µl) were loaded onto a 10 % Bis-Tris SDS PAGE gel. The gel was run with 1X MES SDS running buffer (Invitrogen) for 30min at 80 V, then at 110 V for another 1 h. Gel blotting and membrane blocking were carried out as described by (McLellan *et al.*, 2013). amRFP and aGFP primary antibodies (Sigma-Aldrich) were used at 1 : 4000 and 1 : 2000 dilution, respectively. Secondary antibodies anti-rat immunoglobulin G (IgG) horseradish peroxidase (HRP) or anti-rabbit IgG HRP (Sigma-Aldrich) were used at 1: 5000 dilutions. Protein bands imaged with Amersham Hyperfilm™ ECL, were processed by Xograph imaging system

Confocal imaging

Nicotiana benthamiana leaf tissues were mounted on slides and imaged on Nikon A1R confocal microscope. GFP was imaged with 488 nm excitation and emissions collected between 500 and 530 nm, respectively. mRFP fluorescent proteins were excited with 561 nm light and fluorophores emission detected between 600 and 630 nm. Pinhole was set to 1 Airy unit for the longest wavelength of fluorophore. Cells expressing low levels of fluorescence were imaged to minimize over-expression artefacts. Projections were collected from leaf tissue infected by *P. infestans* transformants. The leaf tissue was not heavily inoculated for minimizing auto-fluorescence from the cell damage and death. Images were processed with

propriety confocal software. Figures were analyzed with Adobe Photoshop and Adobe Illustrator.

Disease test

A detached leaf test was performed as previously described (Vleeshouwers *et al.*, 1999), and lesion sizes were measured by digital caliper. The data were analyzed and visualized by Prism7 for the T-test, ANOVA and post-hoc Tukey's test.

Acknowledgements:

This work was supported by a NWO-VIDI grant 12378, China Scholarship Council (CSC). We thank Dr. Natalia A. Peres and Marcus Marin from University of Florida (USA) for kindly providing us the *P. cactorum* isolates from USA, Thijs van Dijk for providing the *P. cactorum* isolates from Europe, Juan Antonio Garcia for the pGWC-PVX vector, Isolde Bertram-Pereira for culturing *Solanum* plants, Henk Smid and Harm Wiegersma for help in the greenhouse and Prof. Dr. Evert Jacobsen for reviewing the manuscript.

References:

Bienert S, Waterhouse A, de Beer TAP, Tauriello G, Studer G, Bordoli L, Schwede T (2017) The SWISS-MODEL Repository—new features and functionality. Nucleic Acids Research **45**: D313–D319

Bohs L, Olmstead RG (1997) Phylogenetic Relationships in *Solanum* (Solanaceae) Based on ndhF Sequences. Systematic Botany **22**: 5

Bos JIB, Armstrong M, Whisson SC, Torto TA, Ochwo M, Birch PRJ, Kamoun S (2003) Intraspecific comparative genomics to identify avirulence genes from *Phytophthora*. New Phytol **159**: 63–72

Bruck RI, Fry WE, Apple AE, Mundt C (1981) Effect of protectant fungicides on the developmental stages of Phytophthora infestans in potato foliage. Phytopathology

Chen X-R, Li Y-P, Li Q-Y, Xing Y-P, Liu B-B, Tong Y-H, Xu J-Y (2016) SCR96, a small cysteine-rich secretory protein of Phytophthora cactorum, can trigger cell death in the Solanaceae and is important for pathogenicity and oxidative stress tolerance. Molecular Plant Pathology **17**: 577–587

Cook DE, Mesarich CH, Thomma BPHJ (2015) Understanding plant immunity as a surveillance system to detect invasion. Annu Rev Phytopathol **53**: 541–563

Darling ACE, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Research **14**: 1394–1403

Dong S, Raffaele S, Kamoun S (2015) The two-speed genomes of filamentous pathogens: waltz with plants. Current Opinion in Genetics & Development **35**: 57–65

Du J, Rietman H, Vleeshouwers VGAA (2014) Agroinfiltration and PVX agroinfection in potato and *Nicotiana benthamiana*. JoVE e50971–e50971

Fellbrich G, Romanski A, Varet A, Blume B, Brunner F, Engelhardt S, Felix G, Kemmerling B, Krzymowska M, Nürnberger T (2002) NPP1, a *Phytophthora*-associated trigger of plant defense in parsley and *Arabidopsis*. The Plant Journal **32**: 375–390

Grenville-Briggs LJ, Kushwaha SK, Cleary MR, Witzell J, Savenkov EI, Whisson SC, Chawade A, Vetukuri RR (2017) Draft genome of the oomycete pathogen *Phytophthora cactorum* strain LV007 isolated from European beech (*Fagus sylvatica*). Genomics Data **12**: 155–156

Haas BJ, Kamoun S, Zody MC, Jiang RHY, Handsaker RE, Cano LM, Grabherr M, Kodira CD, Raffaele S, Torto-Alalibo T, et al (2009) Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature **461**: 393–398

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Mol Biol Evol **23**: 254–267

Jacobs MM, van den Berg RG, Vleeshouwers VG, Visser M, Mank R, Sengers M, Hoekstra R, Ben Vosman (2008) AFLP analysis reveals a lack of phylogenetic structure within *Solanum* section Petota. BMC Evol Biol **8**: 1

Jashni MK, Mehrabi R, Collemare J, Mesarich CH, de Wit PJGM (2015) The battle in the apoplast: further insights into the roles of proteases and their inhibitors in plant–pathogen interactions. Frontiers in Plant Science. doi: 10.3389/fpls.2015.00584

Jones JDG, Dangl JL (2006) The plant immune system. Nature **444**: 323–329

Kachroo A, Schopfer CR, Nasrallah ME, Nasrallah JB (2001) Allele-specific receptor-ligand interactions in *Brassica* self-incompatibility. Science **293**: 1824–1826

Kamoun S (2007) Groovy times: filamentous pathogen effectors revealed. Current Opinion in Plant Biology **10**: 358–365

Kamoun S, van West P, de Jong AJ, de Groot KE, Vleeshouwers VG, Govers F (1997) A gene encoding a protein elicitor of *Phytophthora infestans* is down-regulated during infection of potato. MPMI **10**: 13–20

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol **30**: 772–780

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

Lamour KH, Mudge J, Gobena D, Hurtado-Gonzales OP, Schmutz J, Kuo A, Miller NA, Rice BJ, Raffaele S, Cano LM, et al (2012) Genome Sequencing and Mapping Reveal Loss of Heterozygosity as a Mechanism for Rapid Adaptation in the Vegetable Pathogen *Phytophthora capsici*. Molecular plant-microbe interactions. doi: 10.1094/MPMI-02-12-0028-R

Liu Z, 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 *scr74* gene family of *Phytophthora infestans*. Mol Biol Evol **22**: 659–672

Ma Z, Song T, Zhu L, Ye W, Wang Y, Shao Y, Dong S, Zhang Z, Dou D, Zheng X, et al (2015) A *Phytophthora sojae* glycoside hydrolase 12 protein is a major virulence factor during soybean infection and is recognized as a PAMP. The Plant Cell **27**: 2057–2072

Martin MD, Cappellini E, Samaniego JA, Zepeda ML, Campos PF, Seguin-Orlando A, Wales N, Orlando L, HO SYW, Dietrich FS, et al (2013) Reconstructing genome evolution in historic samples of the Irish potato famine pathogen. Nat Commun **4**: 2172

Martin MD, Vieira FG, HO SYW, Wales N, Schubert M, Seguin-Orlando A, Ristaino JB, Gilbert MTP (2016) Genomic Characterization of a South American *Phytophthora* Hybrid Mandates Reassessment of the Geographic Origins of *Phytophthora infestans*. Mol Biol Evol **33**: 478–491

McLellan H, Boevink PC, Armstrong MR, Pritchard L, Gomez S, Morales J, Whisson SC, Beynon JL, Birch PRJ (2013) An RxLR Effector from *Phytophthora infestans* Prevents Re-localisation of Two Plant NAC Transcription Factors from the Endoplasmic Reticulum to the Nucleus. PLoS Pathog **9**: e1003670

Nicastro G, Orsomando G, Ferrari E, Manconi L, Desario F, Amici A, Naso A, Carpaneto A, Pertinhez TA, Ruggieri S, et al (2009) Solution structure of the phytotoxic protein

PcF: The first characterized member of the *Phytophthora* PcF toxin family. Protein Sci **18**: 1786–1791

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 **111**: 16955–16960

Orsomando G, Brunetti L, Pucci K, Ruggeri B, Ruggieri S (2011) Comparative structural and functional characterization of putative protein effectors belonging to the PcF toxin family from *Phytophthora* spp. Protein Sci **20**: 2047–2059

Orsomando G, Lorenzi M, Ferrari E, de Chiara C, Spisni A, Ruggieri S (2003) PcF protein from *Phytophthora cactorum* and its recombinant homologue elicit phenylalanine ammonia lyase activation in tomato. CMLS, Cell Mol Life Sci **60**: 1470–1476

Orsomando G, Lorenzi M, Raffaelli N, Dalla Rizza M, Mezzetti B, Ruggieri S (2001) Phytotoxic protein PcF, purification, characterization, and cDNA sequencing of a novel hydroxyproline-containing factor secreted by the strawberry pathogen *Phytophthora cactorum*. Journal of Biological Chemistry **276**: 21578–21584

Pel MJC, Pieterse CMJ (2013) Microbial recognition and evasion of host immunity. Journal of Experimental Botany **64**: 1237–1248

Qutob D, Kamoun S, Gijzen M (2002) Expression of a *Phytophthora sojae* necrosis-inducing protein occurs during transition from biotrophy to necrotrophy. The Plant Journal **32**: 361–373

Thomma BPHJ, Nürnberger T, Joosten MHAJ (2011) Of PAMPs and Effectors: the blurred PTI-ETI dichotomy. The Plant Cell **23**: 4–15

Tian M, Benedetti B, Kamoun S (2005) A second kazal-like protease inhibitor from *Phytophthora infestans* inhibits and interacts with the apoplastic pathogenesis-related protease P69B of tomato. PLANT PHYSIOLOGY **138**: 1785–1793

Tian M, Huitema E, da Cunha L, Torto-Alalibo T, Kamoun S (2004) A kazal-like extracellular serine protease inhibitor from *Phytophthora infestans* targets the tomato pathogenesis-related protease P69B. Journal of Biological Chemistry **279**: 26370–26377

Tian M, Win J, Song J, van der Hoorn R, van der Knaap E, Kamoun S (2007) A *Phytophthora infestans* cystatin-like protein targets a novel tomato papain-like apoplastic protease. PLANT PHYSIOLOGY **143**: 364–377

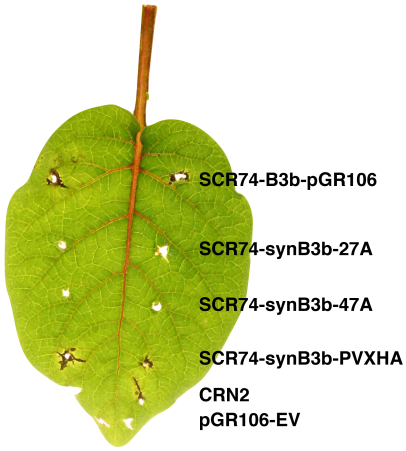
Turnbull D, Yang L, Naqvi S, Breen S, Welsh L, Stephens J, Morris J, Boevink PC, Hedley PE, Zhan J, et al (2017) RXLR Effector AVR2 Up-Regulates a Brassinosteroid-Responsive bHLH Transcription Factor to Suppress Immunity. PLANT PHYSIOLOGY **174**: 356–369

Tyler BM, Tripathy S, Zhang X, Dehal P, Jiang RHY, Aerts A, Arredondo FD, Baxter L, Bensasson D, Beynon JL, et al (2006) *Phytophthora* Genome Sequences Uncover Evolutionary Origins and Mechanisms of Pathogenesis. Science **313**: 1261–1266

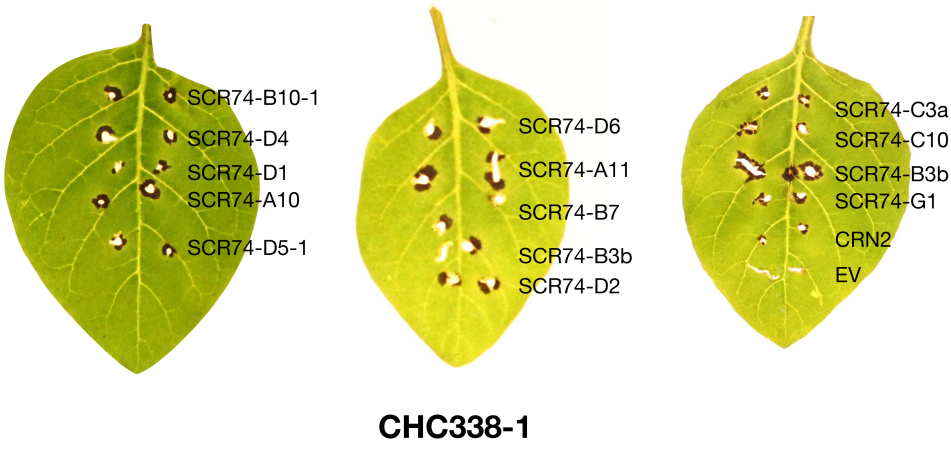
Vleeshouwers VGAA, van Dooijeweert W, Keizer LCP, Sijpkes L, Govers F, Colon LT (1999) A Laboratory Assay for Phytophthora infestans Resistance in Various Solanum Species Reflects the Field Situation. Eur J Plant Pathol **105**: 241–250

Wang S, Boevink PC, Welsh L, Zhang R, Whisson SC, Birch PRJ (2017) Delivery of cytoplasmic and apoplastic effectors from *Phytophthora infestans* haustoria by distinct secretion pathways. New Phytol **42**: 385

Supplementary data



Supplementary Figure S1. SCR74-B3b cysteine mutations do not trigger cell death response on GIG362-6 by PVX agro-infection.



Supplementary Figure S2. CHC338-1 respond to all the tested SCR74 variants by PVX agro-infection.

Supplementary Table S1. List of PcF/SCR effectors from oomycete

Lab-ID	UniProtKB Protein Accession	Sequence Length	Species
HaSCR217	M4BIM8	217	*Hyaloperonospora arabidopsidis* (strain Emoy2) (Downy mildew agent)
HaSCR236	F6MEX8	236	*Hyaloperonospora arabidopsidis*
PcaSCR82	A0A0M3Q8U4	82	*Phytophthora capsici*
PcSCR73	Q94FS7	73	*Phytophthora cactorum*
PiSCR101	D0NEI9	101	*Phytophthora infestans* (strain T30-4) (Potato late blight fungus)
PiSCR111	D0NDE4	111	*Phytophthora infestans* (strain T30-4) (Potato late blight fungus)
PiSCR121	D0ND98	121	*Phytophthora infestans* (strain T30-4) (Potato late blight fungus)
PiSCR132	D0NEI3	132	*Phytophthora infestans* (strain T30-4) (Potato late blight fungus)
PiSCR147	D0NQS1	147	*Phytophthora infestans* (strain T30-4) (Potato late blight fungus)
PiSCR70	D0MV82	70	*Phytophthora infestans* (strain T30-4) (Potato late blight fungus)
PiSCR70	D0NU29	70	*Phytophthora infestans* (strain T30-4) (Potato late blight fungus)
PiSCR75	D0N5R9	75	*Phytophthora infestans* (strain T30-4) (Potato late blight fungus)
PiSCR76	D0NEI4	76	*Phytophthora infestans* (strain T30-4) (Potato late blight fungus)
PiSCR89	D0NEI6	89	*Phytophthora infestans* (strain T30-4) (Potato late blight fungus)
PiSCR91	Q2M413	91	*Phytophthora infestans* (strain T30-4) (Potato late blight fungus)
PiSCR91-2	Q2M443	91	*Phytophthora infestans* (strain T30-4) (Potato late blight fungus)
PiSCR95	D0NPI3	95	*Phytophthora infestans* (strain T30-4) (Potato late blight fungus)
PpSCR105	W2NBA3	105	*Phytophthora parasitica* (Potato buckeye rot agent)
PpSCR106	W2IWX0	106	*Phytophthora parasitica* (Potato buckeye rot agent)
PpSCR124	W2K3T3	124	*Phytophthora parasitica* (Potato buckeye rot agent)
PpSCR76	W2RAA9	76	*Phytophthora parasitica* (strain INRA-310)
PpSCR82	W2QWJ9	82	*Phytophthora parasitica* (strain INRA-310)
PpSCR82-2	A0A080Z4E3	82	*Phytophthora parasitica* P1976
PpSCR87	W2J3R9	87	*Phytophthora parasitica* (Potato buckeye rot agent)
PrSCR73	H3G5T2	73	*Phytophthora ramorum* (Sudden oak death agent)
PsSCR71	G4ZGN4	71	*Phytophthora sojae* (strain P6497) (Soybean stem and root rot agent)
PsSCR68	G4ZGN3	68	*Phytophthora sojae* (strain P6497) (Soybean stem and root rot agent)
PsSCR69	G4ZGN1	69	*Phytophthora sojae* (strain P6497) (Soybean stem and root rot agent)
PsSCR71	G4ZGN2	71	*Phytophthora sojae* (strain P6497) (Soybean stem and root rot agent)
SCR74-E5	Q646U3	74	*Phytophthora infestans* (Potato late blight fungus)
SCR74-C10	Q646U4	74	*Phytophthora infestans* (Potato late blight fungus)
SCR74-C3	Q646U5	74	*Phytophthora infestans* (Potato late blight fungus)
SCR74-G1	Q646U7	74	*Phytophthora infestans* (Potato late blight fungus)
SCR74-F12	Q646U8	74	*Phytophthora infestans* (Potato late blight fungus)
SCR74-D2	Q646U9	74	*Phytophthora infestans* (Potato late blight fungus)
SCR74-Cu6	Q646V1	74	*Phytophthora infestans* (Potato late blight fungus)
SCR74-B7	Q646V3	74	*Phytophthora infestans* (Potato late blight fungus)
SCR74-C9	Q646V4	74	*Phytophthora infestans* (Potato late blight fungus)
SCR74-A11	Q646V5	74	*Phytophthora infestans* (Potato late blight fungus)
SCR74-D6	Q646V6	74	*Phytophthora infestans* (Potato late blight fungus)
SCR74-D5	Q646V8	74	*Phytophthora infestans* (Potato late blight fungus)
SCR74-A10	Q646W0	74	*Phytophthora infestans* (Potato late blight fungus)
SCR74-D1	Q646W1	74	*Phytophthora infestans* (Potato late blight fungus)
SCR74-E11	Q646W2	74	*Phytophthora infestans* (Potato late blight fungus)
SCR74-D4	Q646W3	74	*Phytophthora infestans* (Potato late blight fungus)
SCR74-B10	Q646W6	74	*Phytophthora infestans* (Potato late blight fungus)

Supplementary Table S2. List of *P. infestans* sequencing data from public database

Project number	Isolate	Type of data	SRA ID
PRJNA299503	EC-1	RNA-Seq	SRX1382496
PRJNA275926	Kew126	RNA-Seq	SRX889647
PRJNA275926	Kew123	RNA-Seq	SRX889645
PRJNA275926	Kew122	RNA-Seq	SRX889643
PRJNA275926	*P. andina* EC3394	RNA-Seq	SRX889641
PRJNA275926	*P. andina* PaX	RNA-Seq	SRX889640
PRJNA275926	*P. andina* P13803	RNA-Seq	SRX889639
PRJNA275926	PIC98372	RNA-Seq	SRX889638
PRJNA275926	P6636	RNA-Seq	SRX889637
PRJNA275926	P3873	RNA-Seq	SRX889636
PRJNA275926	P13346	RNA-Seq	SRX889635
PRJNA275926	P13198	RNA-Seq	SRX889634
PRJNA275926	P10650	RNA-Seq	SRX889633
PRJNA275926	PIC97630	RNA-Seq	SRX889632
PRJNA275926	PIC97605	RNA-Seq	SRX889631
PRJNA275926	PIC97207	RNA-Seq	SRX889630
PRJNA275926	PCZ098	RNA-Seq	SRX889629
PRJNA275926	PCZ050	RNA-Seq	SRX889628
PRJNA275926	PCZ033	RNA-Seq	SRX889627
PRJNA275926	PCZ026	RNA-Seq	SRX889626
PRJNA275926	PCO038	RNA-Seq	SRX889625
PRJNA275926	PHU006	RNA-Seq	SRX889624
PRJNA275926	P6570	RNA-Seq	SRX889623
PRJNA275926	P6515	RNA-Seq	SRX889622
PRJNA275926	P8144	RNA-Seq	SRX889621
PRJNA275926	P8143	RNA-Seq	SRX889620
PRJNA275926	P8141	RNA-Seq	SRX889619
PRJNA275926	P8140	RNA-Seq	SRX889618
PRJNA275926	P7036	RNA-Seq	SRX889617
PRJNA275926	P6635	RNA-Seq	SRX889616
PRJNA275926	P6634	RNA-Seq	SRX889615
PRJNA275926	P6629	RNA-Seq	SRX889614
PRJNA275926	P3685	RNA-Seq	SRX889613
PRJNA275926	P3683	RNA-Seq	SRX889612
PRJNA275926	P3681	RNA-Seq	SRX889611
PRJNA275926	P8844	RNA-Seq	SRX889610
PRJNA415282	80029	RNAseq	SRX3330543
PRJNA415282	F48	RNAseq	SRX3330544
PRJNA415282	Pd21410	RNAseq	SRX3330541
PRJNA415282	Pc51265	RNAseq	SRX3330542
PRJNA415282	Pa21106	RNAseq	SRX3330545
PRJNA361417	1306	RNAseq	SRX2495351
PRJNA353657	XA-4	RNAseq	SRX2355392
PRJNA329480	11/10	RNAseq	SRR3990754
PRJNA323952	*Pi* GBS Plate 1	WGS	SRX2169699
PRJNA323952	*Pi* GBS Plate 2	WGS	SRX2169700
PRJNA323952	*Pi* GBS Plate 3	WGS	SRX2169701
PRJEB4015	Pi1889	WGS	ERX272905
PRJEB4015	Pi1882	WGS	ERX272923
PRJEB4015	Pi1876	WGS	ERX272917
PRJEB4015	Pi1845B	WGS	ERX272917
PRJEB4015	Pi1845A	WGS	ERX272909
PRJEB4015	BL2009P4_US-23	WGS	ERX272899
PRJEB4015	IN2009T1_US-22	WGS	ERX272898
PRJEB4015	RS2009P1_US-8	WGS	ERX272897
PRJEB20998	PIC99189	WGS	ERR1990236
PRJEB20998	88069	WGS	ERR1990235

Supplementary Table S3. Effectors used for Effectoromics screening

	Protein sequence	Vector	Agrobacterium strain	Notes
PcF	MNFKTCLAVALVAVVATVATAEDPLYCQAIGCPTLYSEANLA VSKECRDQGKLGDDFHRCCEEQCGSTTPASA	pGWC-PVX	GV3101	Synthesized by GenWiz
SCR96	MKFSTVFTAAAVAVVCLLQPSVAEEQASVHLRVHTVEQSN NGAICYQACQSGQYCPRGENKCRAPTGNQCFNPATSLFR EGCDPGFKCDKGKCVYK	pGWC-PVX	GV3101	Synthesized by GenWiz
SCR68	MNLKIFVFVSTVLATSLVSVTAQSQVCGAQGCGAPFSASNI QISSCCSRQPNFNECCRTSCNLGSPCQ	pGWC-PVX	GV3101	Synthesized by GenWiz
SCR74-C10	MNFKIYAIVALTAVLATPITAQQQQLCRADGCAYEYSLANDV VSKCCKAINADPIAFHDCCSKSCNTGSPCKSV	pGR106	GV3101	(Liu, 2005)
SCR74-B10-1	MNFKIYAIVALTAVLATPITAQQQQLCKADGCAYEYSLANDV VSKCCKAINADPIAFHDCCSKSCNTGSPCKSV	pGR106	GV3101	(Liu, 2005)
SCR74-D4	MNLKIYAIVALTAVLATPITAQQQQLCKADGCAYEYSLANDV VSKCCKAINADPIAFHDCCSKSCNTGSPCKSV	pGR106	GV3101	(Liu, 2005)
SCR74-D6	MNLKIYAIVVLTAVLATPITAQQQQLCKAVGCAYEYSHANDV VSKCCQAINPDPVAFHDCCGKSCNTGIPCKSV	pGR106	GV3101	(Liu, 2005)
SCR74-D5-1	MNLKIYAIVALTAVLATPITAQQQQLCKAVGCAYEYSHANDV VSKCCQAINPDPVAFHDCCGKSCNTGIPCKSV	pGR106	GV3101	(Liu, 2005)
SCR74-A10	MNFKIYAIVALTAVLATPITAQQQQLCKAVGCAYEYSHANDV VSKCCQAINPDPVAFHDCCGKSCNTGIPCKSV	pGR106	GV3101	(Liu, 2005)
SCR74-C3a	MNFKIYAIVALTAVLATPITAQQQQLCKAAGCAYEHSSANGV VSQCCKAINAEPVAFNDCCSKSCNTGSPCRSV	pGR106	GV3101	(Liu, 2005)
SCR74-G1	MNLKIYAIVALTAVLATPITAQQQQLCKAVGCAYEYSHANDV VSQCCKAINAEPVAFNDCCSKSCNTGSPCKSV	pGR106	GV3101	(Liu, 2005)
SCR74-B7	MNFKIYAIVALTAVLATPITAQQQQLCRADGCAYEYSLANKVI SKCCQAINPDPVAFYDCCRISCNMGSPCKAV	pGR106	GV3101	(Liu, 2005)
SCR74-B3b	MNFKIYAIVALTAVLATPITAQQQQLCRADGCAYEHSHANKV ISKCCQAINPDPVAFYDCCRISCNMGSPCKAV	pGR106	GV3101	(Liu, 2005)
SCR74-D1	MNFKIYAIVALTAVLATPITAQQQQLCKAVGCAYEYSHANDV VSKCCQAINPDPIAFHDCCSKSCNTGSPCKSV	pGR106	GV3101	(Liu, 2005)
SCR71-A11	MNLKIYAIVALTAVLATPITAQQQQLCKAVGCAYEYSHANDV VSKCCQAINADPIAFHDCCSKSCNTGSPCKSV	pGR106	GV3101	(Liu, 2005)
SCR74-D2	MNFKIYAIVALTAVLATPITAQQQQLCRADGCAYEHSHANKV ISKCCQAINPDPVAFHDCCGKSCNTGIPCKSV	pGR106	GV3101	(Liu, 2005)

Supplementary Table S4. Effectoromics screening on Solanaceae genotypes

For convenience, this supplementary table is deposited in 4TU.Centre:
https://data.4tu.nl/repository/uuid:9032af40-6cd0-409d-be0d-76fe0f4ed576

Supplementary Table S5. SCR74 homologs from different isolates

***P. infestans* isolates**	**SCR74 variants**													
	C3a	**B11**	**C4**	**C10**	**E6**	**C4**	**B10**	**B3b**	**D6**	**D5**	**E5**	**H1**	**C3a**	**D6**
90128							x							
89148-09						x								
EC1	x													
H3OPO4		x												
IPO-0		x	x	x										
IPO-C			x					x						
Katshaar									x	x				
PIC99177					x			x						
PIC99183								x		x				
PIC99189			x								x			
UK3928-A					x			x		x		x	x	
UK7824					x									

Supplementary Table S6. List of *P. cactorum* isolates. The geographic origin, year of collection, sampled tissue and disease are indicated.

Isolate	Place	Year	Tissue	Disease
NL 2003-3	Netherlands	n.a.	n.a.	n.a.
10531-3	Netherlands	n.a.	n.a.	n.a.
PD_04764040	Norway	2013	n.a.	n.a.
7932	Florida	2017	Crown	*P. cactorum* rot
1726	Florida	2017	Crown	*P. cactorum* rot
16-24	Florida	2016	Crown	*P. cactorum* rot
16-353	Florida	2016	Crown	*P. cactorum* rot
16-351	Florida	2016	Crown	*P. cactorum* rot
7924	Florida	2017	Crown	*P. cactorum* rot
02-202	Florida	2002	Crown	*P. cactorum* rot
9754	Florida	1997	Fruit	Leather rot

Chapter 3

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

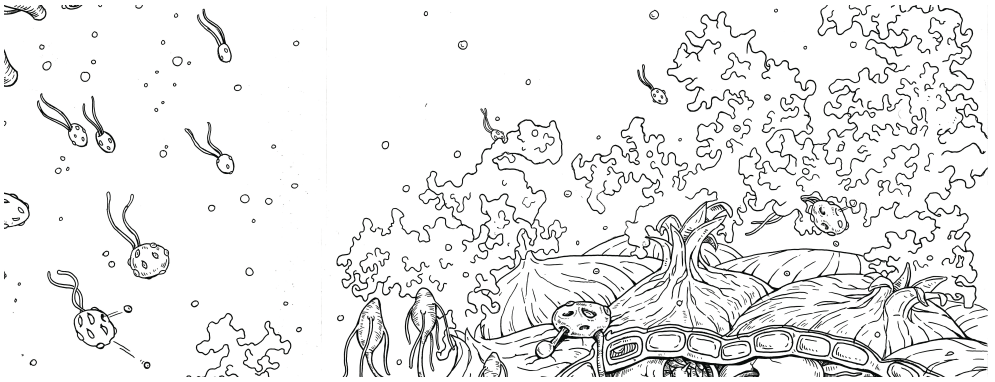
Xiao Lin[1], Katie Baker[2], Miles Armstrong[2], Doret Wouters[1], Richard G. F. Visser[1], Pieter J. Wolters[1], Ingo Hein[2,3*], Vivianne G. A. A. Vleeshouwers[1*]

[1]Plant Breeding, Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB, Wageningen, The Netherlands.

[2]Cell and Molecular Sciences, The James Hutton Institute, DD2 5DA, Dundee, UK

[3]School of Life Sciences, Division of Plant Sciences, University of Dundee at the James Hutton Institute, Dundee DD2 5DA, UK

*Co-corresponding authors



This chapter has been submitted to New Phytologist

Abstract

The identification of plant immune receptors in crop plants is a challenging and time-consuming process. In recent years, resistance (*R*) gene enrichment sequencing (RenSeq) has been developed to accelerate the mapping and cloning of nucleotide-binding domain and leucine-rich repeat containing (*NLR*) genes. However, the first layer of defense that is mediated by receptor-like proteins (RLPs) or receptor-like kinases (RLKs) has remained elusive. Here, we developed a new methodology for the identification of *RLP* and *RLK*-like genes in plants. We combine effectoromics with RLP/K enrichment sequencing (RLP/KSeq). An RLP/RLK bait library was designed based on the potato DM genome. INF1 and SCR74 of the Irish famine pathogen *Phytophthora infestans* were functionally screened on segregating populations of wild *Solanum* species. *Solanum microdontum* accessions responsive and non-responsive to INF1 and SCR74, respectively, together with the bulked segregating population were sequenced after RLP/RLK gene enrichment. This resulted in the accelerated mapping of the immune receptors of SCR74 on potato linkage group 9 and confirmed the position of previously cloned INF1 receptor ELR on linkage group 12. These results demonstrate that the combination of the RLP/KSeq with effectoromics provides a new and effective tool for the accelerated identification of cell surface immune receptors in plants.

Introduction

To protect themselves against pathogens, plants have evolved two layers of defense (Jones and Dangl, 2006). The first layer of defense 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, microbial/pathogen-associated molecular patterns (MAMPs/PAMPs) from plant pathogens and danger-associated plant breakdown products (DAMPs). The second layer of defense is mounted upon recognition of cytoplasmic effectors by internal immune receptors that typically encoded by resistance (*R*) genes of the nucleotide-binding domain and leucine-rich repeat (NLR) class. Stacking and pyramiding *R* genes and surface immune receptors 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.*, 2011; 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 provide a broader spectrum of resistance and is based on RLP/RLK-mediated recognition of MAMPs/PAMPs or apoplastic effectors. Examples of MAMPs are flagellin and elicitins, of bacteria and oomycetes, respectively (Felix *et al.*, 1999; Derevnina *et al.*, 2016). INF1 is a well-studied elicitin from *Phytophthora* that triggers defense 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 the small cysteine-rich protein SCR74 of *P. infestans* (Liu *et al.*, 2005). Cloning and characterizing additional plant surface immune receptors, such as the receptor of SCR74, will deepen 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 re-sequencing (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., *R* genes), target enrichment sequencing significantly reduces the complexity of the genome prior to sequencing (Hodges *et al.*, 2007). In 2013, *R* gene enrichment sequencing (RenSeq) was described to aid the re-annotation and mapping of *NLR* genes in potato. All *NLR* genes from the potato reference genome DM, v4.03 (doubled monoploid *S. 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 *S. berthaultii, S. ruiz-ceballosii* and *S. verrucosum* respectively (Jupe *et al.*, 2013; Chen *et al.*, 2018). When combined with single-molecule real-time (SMRT) PacBio sequencing, an advanced RenSeq resulted in true sequence representation of full-length *NLR* genes, and the late blight *R* gene *Rpi-amr3i* from *S. americanum* was cloned directly (Witek e*t 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). 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 RenSeq, 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 – *P. infestans* pathosystem as an example. We developed a pipeline **(Figure 1)** that consists of two steps: 1) effectoromics: screening wild *Solanum* species to identify plants that recognize apoplastic effectors, 2) RLP/RLK enrichment sequencing (RLP/KSeq) to accelerate the genetic mapping of the underlying immune receptors through bulked segregant analysis (BSA).

Results

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

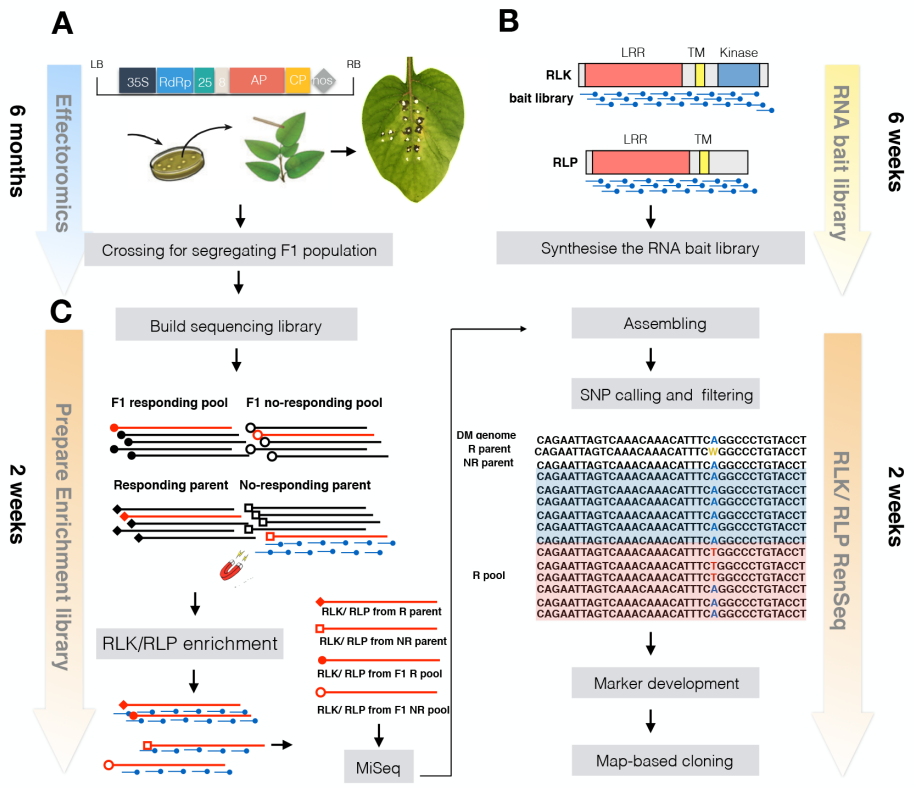


Figure 1. Overview of the effectoromics and RLP/KSeq pipeline for fast identification and mapping of surface immune receptors.

A. Predicted P. infestans apoplastic effectors were 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 dpi for occurrence of cell death phenotype. Responsive and non-responsive genotypes are crossed for creating a segregating F1 populations.

B. Prediction of the *RLP* and *RLK* genes from the reference genome, then design and synthesis of the RLP/ RLK bait library in the target species.

C. The F1 population was screened for segregation and pooled, based on their response pattern. Building the sequencing library for enrichment of RLP and RLK genes by the bait library. Sequencing of enriched samples and the reads were mapped to the reference genome, performing SNP calling and filtering for potential linkage markers. The candidate markers are tested on the segregating population by SNP genotyping technologies like LightScanner.

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

Figure 2. *Solanum* species showed specific response to INF1 and SCR74 after PVX agro-infection.

The *Solanum* genotypes show response to either INF1, SCR74 or both are indicated. The empty pGR106 vector and pGR106-CRN2 were included as negative and positive controls, respectively. The response was scored from 0-10. Scores are presented as heat plot, as no (0-2, blank), weak (3-4 yellow, 5-6 orange), strong (7-10, red) responses. Experiments were repeated at least 3 times. The ploidy level and endosperm balance number (EBN) are shown. The country are shown by abbreviation: MEX=Mexico, BOL=Bolivia, PER=Peru, ARG=Argentina, ECU=Ecuador.

To explore the recognition spectra of apoplastic effectors of *Phytophthora infestans*, transient effectoromics screens with INF1 elicitin and SCR74 were performed on a wide range of wild *Solanum* genotypes **(Figure 1a)**. In total, 105 *Solanum* genotypes were screened for response to INF1 and SCR74 by PVX agro-infection. 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 non-responsive plants, is presented in **Supplementary Table S1**. A set of 34 *Solanum* genotypes showed specific cell death responses to INF1, and/or SCR74 two weeks after agro-infection **(Figure 2)**. These responsive plants belong to 17 different wild *Solanum* species and vary in ploidy levels as well as endosperm balance numbers (EBN), and originate from different geographic origins **(Figure 2)**. In most cases, the specific effector responses were clear and highly reproducible (i.e., clear cell death phenotypes scores >7, n>10). In some cases, we observed more variability (cell death phenotypes scores ranging from 4 to 6, n>10), but these variations were also observed for the positive control, CRN2 which suggest that these plant accessions were less amenable to the PVX-based transient expression system. As expected, response to INF1 elicitin was confirmed in *Solanum microdontum* genotype MCD360-1 **(Figure 2)**, which is the source of the elicitin receptor ELR (Du *et al.*, 2015). In addition, other *Solanum* genotypes were also found to respond to INF1 **(Supplementary Table S1)**. Similarly, SCR74 was recognized in various plants including *Solanum microdontum* ssp. *gigantophyllum* GIG362-6 **(Figure 2)**. In conclusion, responses to INF1 and SCR74 are widely distributed 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 Solanum 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 **(Figure 1a)**. We crossed the INF1- but not SCR74-responsive genotype MCD360-1 with the SCR74- but not INF1-responsive genotype GIG362-6 and generated the F1 population 7026 **(Figure 3)**. From this population, 55 progenies were tested for responses to INF1 and SCR74 by PVX agroinfection. The population segregated for responses to INF1, with 31 responsive versus 24 non-responsive offspring clones, which is close to a 1:1 segregation ($\chi 2$=3.311, P=0.068). Segregation for responses to SCR74 was also observed at a near 1:1 ratio ($\chi 2$=2.778 P= 0.095), as 25 responsive versus 30 non-responsive offspring genotypes were identified. Importantly, the responses to SCR74 were independent of the responses to

INF1. Both segregation ratios are consistent with single dominant genes that mediate the responses to INF1 and SCR74, respectively.

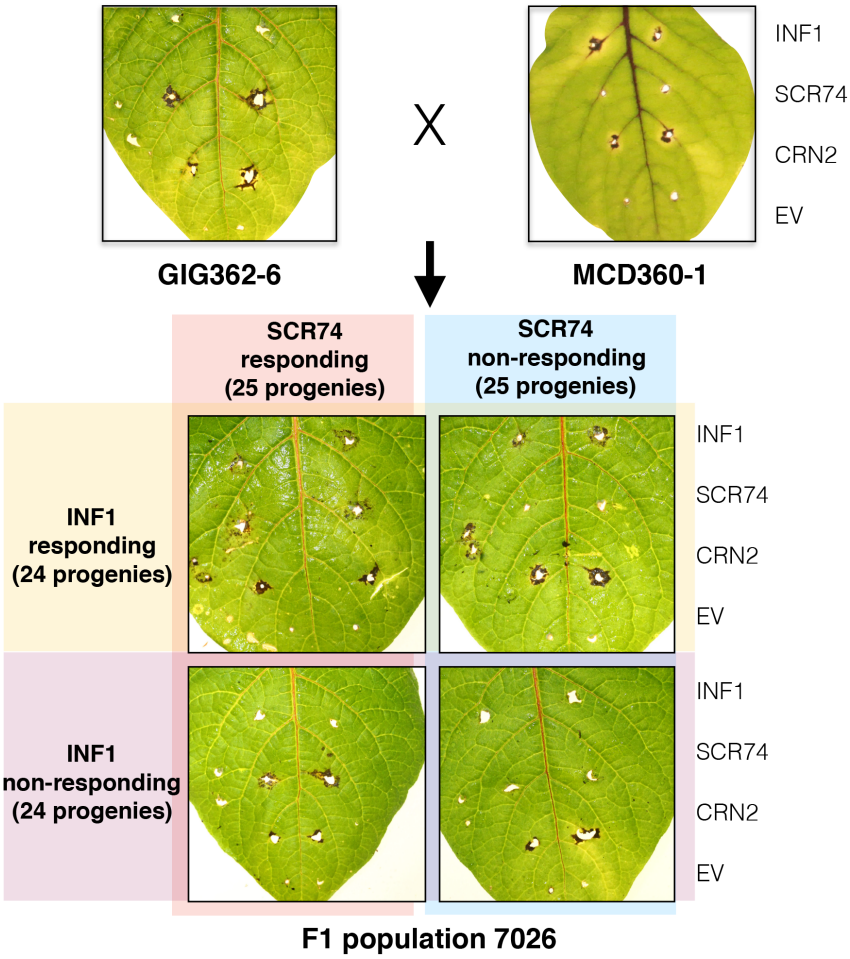


Figure 3. Independent segregation of responses to SCR74 and INF1 in F1 population 7026 of *S. microdontum*.

S. microdontum ssp. *gigantophyllum* GIG362-6 (SCR74 responsive) was crossed with MCD360-1 (INF1 responsive) and 54 progeny plants were PVX agroinfected with pGR106-INF1 and pGR106-SCR74. The empty pGR106 vector and pGR106-CRN2 were included as negative and positive controls, respectively. Four classes of response to the effectors are represented. Photographs display representatives of the phenotypes for each class, at 14 days post infection (dpi).

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 (Jupe *et al.*, 2013) **(Figure 1b)**. Since the INF1 receptor ELR was originally cloned from MCD360-1, we used this genotype and the segregating progeny as a positive control throughout this study (Du *et al.*, 2015).

To design a comprehensive bait library for Solanaceae RLPs and RLKs, we combined 301 *RLK* and 404 *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 349 *LRR-RLK* genes and 450 *LRR-RLP* genes from the potato reference genome DM, v4.03 (Supplementary Dataset S1 and S2). The localization of these predicted *LRR-RLK* and *LRR-RLP* genes on the 12 potato chromosomes is visualized in **Supplementary Figure S1**. Additionally, 15 known *RLP/ RLK* genes from Solanaceae species were included (Supplementary Dataset S3) alongside the *RLP/RLK* homologs predicted from 42 potato cultivars (unpublished data). In total, our reference set for the RLP/KSeq-specific bait library design featured 359 individual *LRR-RLP* and 455 distinct *LRR-RLK* genes.

A customized target enrichment RNA bait library with 2x coverage comprising 55,970 120-mer biotinylated RNA oligo probes was synthesized (SureSelectXT, Agilent Technologies Inc., Santa Clara, CA, USA) to represent 4.499 Mb targeted genome space (Supplementary Dataset 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 **(Figure 1a and 1b)**.

Bulk segregant analysis 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 non-responsive, plus two parents **(Figure 1c)**. In this case, since we multiplex for two target genes, we composed four bulked pools. These comprised response to INF1 (INF-R), no response to INF1 (INF1-NR), response to SCR74 (SCR74-R), no response to SCR74 (SCR74-NR), consisting of 24, 24, 25 and 25 progeny individuals, respectively **(Figure 3, Supplementary Figure S2)**. DNA was isolated from each progeny and then pooled. DNA from the parents GIG362-6 and MCD360-1 were included in the enrichment as well **(Materials and methods)**. The quality of the enrichment was checked by subjecting samples

of enriched and non-enriched DNA to qPCR utilizing two primer pairs specifically designed for ELR. The results indicate significant enrichment for both amplicons with approximately 70 and 83-fold (ΔCT of 6.133 and 6.382) **(Supplementary Table S2)**.

Mapping reads to the reference genome and SNP calling

Table 1. RLP/KSeq reads mapped to DM genome v4.03 or RLP/ RLK genes at 5%, 10% and 20% mismatch rates

Sample		Raw reads	Mismatch%	Mapped reads					
				total	% Mapped	On target	% On target	Coverage to Andolfo RLK/RLP genes (x)	Coverage to new and known RLK/RLP genes (x)
Parent	GIG362-6	6946620	5	3812996	54.89	2677269	70.21	123	102.38
			10	5280676	76.02	3668681	69.47	156.87	139.7
			20	5983818	86.14	4084436	68.26	170.31	155.33
	MCD360-1	6797696	5	3696826	54.38	2581490	69.83	125.81	103.38
			10	5209584	76.64	3617967	69.45	164.62	145
			20	5972014	87.85	4086740	68.43	180.67	163.75
INF1	Non-responsive bulk	6619040	5	3399574	51.36	2335512	68.7	109.11	89.31
			10	4667260	70.51	3174685	68.02	138.63	121.32
			20	5295008	80	3533629	66.74	150.22	134.76
	Responsive bulk	5192274	5	2753730	53.04	1893944	68.78	89.82	74.28
			10	3858046	74.3	2627109	68.09	116.76	103.18
			20	4414266	85.02	2952107	66.88	127.71	115.91
SCR74	Non-responsive bulk	4104188	5	2033698	49.55	1414158	69.54	67.34	55.21
			10	2833640	69.04	1952277	68.9	87.21	76.41
			20	2052094	79.1	2197327	67.69	95.6	86.02
	Responsive bulk	5267716	5	2789252	52.95	1889301	67.74	90.72	74.78
			10	3884278	73.74	2604040	67.04	116.87	103.03
			20	4428630	84.07	2914742	65.82	127.26	115.11

The *RLP/RLK* genes enriched DNA library from the bulks and parents were sequenced with Illumina 2x300 bp chemistry on a MiSeq platform **(Figure 1c)**. The number of raw reads that passed quality control ranged from 4,404,188 to 6,946,620 in different samples/pools **(Table 1)**. High-quality paired-end reads were mapped to the potato reference genome (DM v4.03) using Bowtie2. To compensate for differences between the potato reference DM and *S. microdontum*, three mismatch rates, 5%, 10% and 20%, were used for the reads mapping **(Materials and methods, Table 1)**. The mapping rates ranged from 50% to 85%, with reads on target accounting for 66% to 70%, depending on the mismatch rate **(Table 1)**. The resulting coverage of known *RLP/ RLK* genes was calculated and ranged from 67.34x to 180.67x. To enable the identification of informative SNPs whilst ensuring sufficient accuracy, a 10% 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 (**Materials and methods, Supplementary Dataset S5**; Chen *et al.*, 2018).

Proof of concept, response to INF1 mapped to ELR locus

As a proof of concept, we first tested if the RLP/KSeq approach could correctly identify the known mapping position for *ELR*, the gene conferring INF1 response. The SNPs from INF1 non-responsive and responsive bulks were called and filtered for the expected ratios of heterozygosity as described by Chen *et al.*, 2018. For a dominant gene segregating in a diploid population, the allele frequencies were set at 0-10% or 90-100% for the INF1 non-responsive bulk and 40-60% for INF1 responsive bulk. The SNPs were compared to the parental SNPs, and only the accordant SNPs were maintained as informative SNPs **(Materials and methods, Supplementary Table S3**).

Allowing for a 10% mismatch rate of RLP/KSeq enriched *S. microdontum* reads against the DM reference genome, 269 SNPs passed the filter criteria in the bulks and 10,697 SNPs in the parents. Among those, 157 SNPs were shared in both bulks and parents **(Supplementary Table S3)**, 12 of them cannot be assigned to any chromosome (Chr0). The number of informative genic SNPs per 1 Mb interval was placed on the 12 chromosomes of potato. With the exception of 1 significant SNP on chromosome 6, the remaining 144 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 **(Figure 4a)**. The majority of SNPs were localized in two *RLP/RLK* loci that correspond to 25 polymorphic genes, 3 of these 25 genes could not be assigned to any chromosome, 1 gene is on chromosome 6, and 21 of 24 genes are on chromosome 12 **(Figure 4b, Supplementary Table S3)**. Amongst these genes identified with significant SNPs linked to the INF1 receptor, PGSC0003DMG401002888 is one of the closest homolog of ELR in DM and shares 86.40% pairwise identity **(Figure 4c)**. 20 SNPs (orange and pink) were called by VarScan, but only 3 SNP (pink) were filtered based on our criteria **(Figure 4c, Supplementary Table S3)**. These results agree well with the known position of *ELR* near the top of chromosome 12 (Du *et al.*, 2015), and provide the proof of concept for RLP/KSeq as a powerful tool for efficient mapping of plant surface immune receptors.

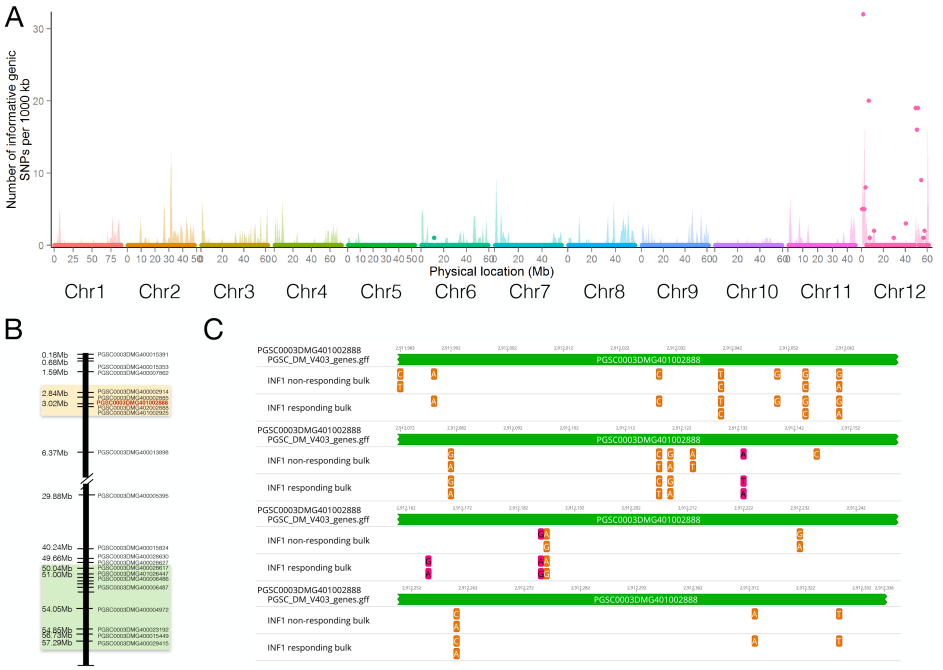


Figure 4. *ELR* was mapped to chromosome 12 by RLP/KSeq

A. The background color and the dots represent the number of annotated RLP/RLKs and the number of SNP per 1 Mb bin, respectively. X-axis, physical location on individual chromosomes; Y-axis, number of SNPs per 1 Mb bin. The peak in chromosome 12 indicates various SNPs that are linked with ELR, which confers response to INF1.
B. Name and position of linkage SNPs containing genes on chromosome 12 of the reference genome. PGSC0003DMG401002888 (highlighted by red) is the closest homolog of *ELR* in DM.
C. Many SNPs (orange) on PGSC0003DMG401002888 were called by VarScan, but only 3 SNPs (pink) are filtered based on our criterion, these 3 SNPs located on 2,912,133 bp, 2,912,166 bp and 2,912,186 bp of chromosome 12, based on the reference genome, which are T->A, G->A and A->G in the INF1 responsive pool and the allele frequency are 52.3%, 46.67% and 49.13% respectively.

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

To map the gene that confers response to SCR74, the same SNP filtering approach was performed as shown for ELR above. The SNPs that meet 0-10% or 90-100% allele frequency in the SCR74 non-responsive bulk and 40-60% allele frequency in the SCR74 responsive bulk were identified and corroborated in the parental material. This resulted in the identification of 20 informative SNPs of which 15 could be placed on the short arm of chromosome 9. The SNPs correspond to 10 polymorphic genes **(Figure 5a. Supplementary Table S3)**.

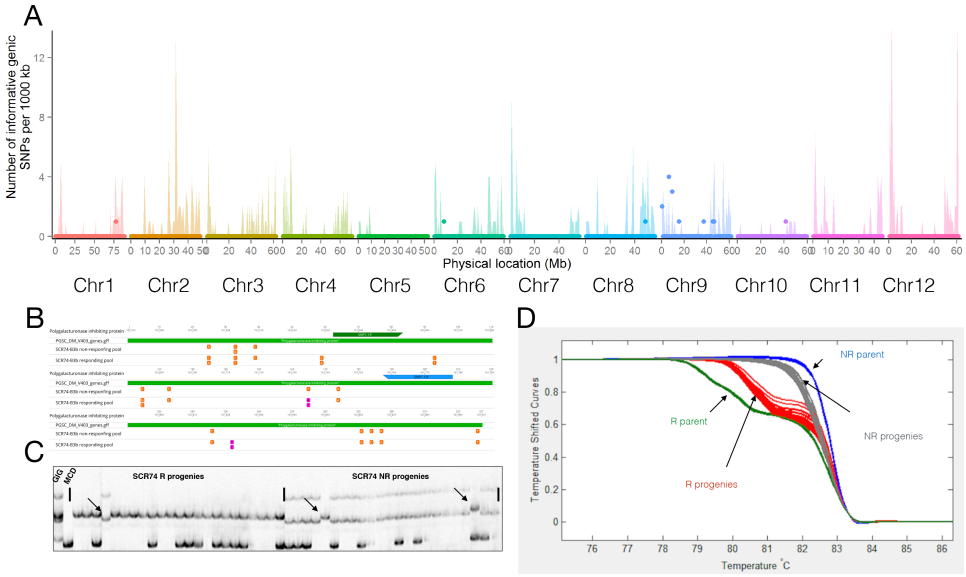


Figure 5. The gene conferring response to SCR74 was mapped to chromosome 9 by RLP/KSeq

A. The background color and the dots represent the number of annotated RLP/RLK genes and the number of SNPs per 1 Mb bin, respectively. X-axis, physical location on individual chromosomes; Y-axis, number of SNPs per 1Mb bin. The peak in chromosome 9 indicate the linkage SNPs with the SCR74 response gene.
B. A polygalacturonase inhibiting protein (PGIP, PGSC0003DMG400006492) contains 2 SNPs (marked in pink) which shows a nearly 1:1 frequency in the SCR74-B3b responsive pool, is linked with the SCR74 receptor. These SNPs are A->T and T->C, respectively. Two primers flanking the first SNP were designed (green arrow) for genotyping. The other polymorphic SNPs between the genome or two pools were marked in orange.
C. A SSR marker STM1051 on chromosome 9 is linked with the SCR74 response, the mapping parents GIG362-6, MCD360-1, responsive progenies and non-responsive progenies were tested by STM1051, and 3 recombination events (arrow) were found.
D. The SNP1.1 marker obtained by RLP/KSeq was tested on the mapping parents and progenies. The melting curve was detected by LightScanner for genotyping.

The identified SNPs were converted to high resolution melting (HRM) markers and used for recombinant screening. RLP/KSeq-snp1.1 (A->T), which corresponds to PGSC0003DMG400008492, a polygalacturonase inhibiting protein (PGIP), located to position 0.16Mb on Chromosome 9, was used to identify recombinants **(Figure 5b)**. This SNP displayed a 59% frequency in the responsive bulk and 0% or 100% frequency in the non-responsive bulk and was tested in the mapping parents and F1 population (56 progenies) via HRM **(Figure 5d)**. In this population, three recombination events between the RLP/KSeq-

snp1.1 and the SCR74 response were detected, which indicates that the SCR74 receptor is likely located on chromosome 9.

To further confirm our RLP/KSeq methods, 78 SSR markers dispersed over all 12 potato chromosomes were tested on 56 offspring F1 progeny of population 7026 **(Materials and methods**, **Supplementary Table S4)**. The strongest correlation between SSR marker genotype and the responsive phenotype to SCR74 was achieved for SSR marker STM1051. This marker resides on position 6.15 Mb of DM chromosome 9 and confirms the RLP/KSeq analysis for the SCR74 response **(Figure 5c)**. Consequently, the SCR74 response gene was mapped rapidly to a 10.7 cM region on potato chromosome 9 between RLP/KSeq-snp1.1 and STM1051 **(Supplementary Figure S3)**. This region spans a 5.99 Mb physical distance on chromosome 9 based on the DM genome.

Discussion

In this paper, we present a workflow that combines RLP/KSeq with effectoromics of apoplastic effectors, to rapidly map plant surface immune receptors **(Figure 1)**. We used potato and *P. infestans* as a model system. We screened for wild potato species that mount specific cell death response to the apoplastic effectors INF1 and/or SCR74 of *P. infestans*. *S. 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 then subjected pools of genomic DNA derived from responding versus non-responding genotypes to a BSA RLP/K enrichment sequencing, using the bait libraries to enrich for genomic DNA covering these types of genes. This approach quickly led to identification of SNPs that are used as molecular markers to genetically map the genes encoding the putative RLP/RLK genes.

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 like 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; Arora *et al.*, 2018), however, they all target *NLR* genes. RLP/KSeq can complement the RenSeq toolbox to target a broader range of

plant immune receptors, including RLPs/RLKs that also function as effective immune receptors (Boutrot and Zipfel, 2017).

Effectoromics has proven a high throughput approach to identify the plants carrying *R* genes as well as surface immune receptors (Vleeshouwers *et al.*, 2011; Du *et al.*, 2015; Domazakis *et al.*, 2017). The accurate specificity and robustness of effector responses enables identifying multiple receptors in a single segregating population **(Figure 3)**. Another advantage of combining the enrichment sequencing with effectoromics is that targeted libraries can be used for surface immune receptors or *NLRs*, depending on the matching effector. Effectoromics was pioneered for the potato – late blight pathosystem, and has been successfully applied in various other Solanaceae, such as *N. 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 and Bos, 2011). This demonstrates the wide application of the effectoromics strategy for pathogens with well-characterized genomes.

To summarize, our newly developed pipeline enables the rapid identification of plants carrying novel immune receptors and mapping them. This strategy is complementing the current RenSeq toolbox and will help to understand the first layer of the plant immune system, to ultimately achieve more durable disease resistance in plants.

Materials and methods

Plant material

Solanum genotypes used in this study are listed in **Table 1** and **Table S1** (Vleeshouwers *et al.*, 2011). These plants were maintained *in vitro* on MS20 medium at 25 °C, as described in (Du *et al.*, 2014). Top shoots of plants were cut and clonally propagated *in vitro* 2 weeks before transfer to the soil in a climate controlled greenhouse compartment at 22 °C/18 °C day/night regime under long day conditions. The F1 population 7026 was generated by crossing *Solanum microdontum* subsp. *gigantophyllum* (GIG362-6) with *Solanum microdontum* (MCD360-1). The plants were grown in a crossing greenhouse 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 weeks, the ripe berries were removed from the plants. The seeds were collected and cleaned by water and dried on filter paper. The seeds were sown on MS20

medium or were soaked on filter paper after 3-4 months dormancy. Gibberellic acid (GA3) was used for breaking dormancy if necessary. 56 progenies of 7026 population were tested in the greenhouse by PVX infection, and each experiment was repeated at least 3 times.

Cloning of effectors for PVX agroinfection

Inf1 (XM_002900382.1) and *Scr74-B3b* (AY723717.1) were cloned into potato virus X (PVX) vector pGR106, and electro-transformed to *A. tumefaciens* strain GV3101. Recombinant *A. tumefaciens* strain GV3101 carrying the effector constructs were grown for 2 days in LB medium at 28 °C with kanamycin (50 μg/ml). 1 ml of *Agrobacterium* culture was plated out onto LBA plates supplemented with kanamycin (50 μg/ml) and incubated at 28 °C for 2 more days. The *Agrobacterium* culture was collected from the Petri dishes with a plate spreader and used to inoculate 3-4 week-old plants through toothpick inoculation (Takken *et al.*, 2000; Du *et al.*, 2014). 2 spots were inoculated per leaf for each construct and 3 leaves were used per plant. 3 replicated plants were used for each genotype. 2 weeks after infection, cell death responses were scored on a range from 0 (no response) to 10 (strong response).

Design of customized RLP/ RLK enrichment library

The RLP/RLK genes were predicted from the DM genome by HMM, BLASTp, and InterPro, from both PGSC and iTAG annotation (the FASTA file of predicted *RLK* and *RLP* genes from DM are in **(Supplementary Dataset S1 and S2)**. All the homologs of the 349 RLKs and 450 RLPs from 42 potato cultivars were extracted from the BAM file, including 14658 RLKs and 18900 RLPs (data not shown). All *RLP* and *RLK* genes were split into 120 bp fragments with 2 times coverage, and the duplicated oligonucleotides were removed. The unique RNA oligonucleotides were synthesized for customized Agilent SureSelect Target Enrichment kit (Agilent Technologies Inc., Santa Clara, CA, USA). In total, 55970 probes were synthesized with the size of 4.499 Mb **(Supplementary dataset S4)**.

Preparation of sequencing library and target capture

Genomic DNA was isolated from GIG362-6, MCD360-1 and the F1 progenies using the DNeasy Plant Mini Kit (Qiagen, Valencia, CA). Equal amounts of DNA were pooled from 25 responsive and 25 non-responsive progenies, for INF1 response and SCR74 response, respectively. The pooled DNA was fragmented by using a Covaris sonicator and then purified using AMPure XP Beads (Beckman Coulter, Inc.). The DNA concentrations were measured using a Qubit fluorometer (Thermofisher, Dubuque, IA, USA). The NEBNext Ultra™ DNA Library Prep Kit (New England Biolabs) was used for adaptor ligation and to add indexes. The Bioanalyzer with a Bioanalyzer high sensitivity DNA chip was used for detecting the size of DNA after fragmentation. The pooled DNA was enriched for RLPs and RLKs with the customized Agilent SureSelect Target Enrichment kit (Agilent Technologies Inc., Santa Clara, CA, USA), following a hybridization period of 24 hours. Post enrichment PCR was performed, and the products were quantified by Qubit in order to check the enrichment efficiency. A qPCR analysis was performed by amplifying two fragments from ELR from the enriched samples, as well as the sample before enrichment. The primers used in this study are in **Supplementary Table S5**. Paired-end sequencing was performed for the 8 samples on the Illumina MiSeq platform using one flowcell.

Read mapping and SNP calling

Paired-end Illumina MiSeq reads were first checked with FastQC (v0.10.0; http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and then quality and adapter trimmed with fastq-mcf (v1.04.676; https://expressionanalysis.github.io/ea-utils/) to a minimum base quality of 20. The trimmed reads were then mapped to the DM (v4.03) reference genome using Bowtie2 (v2.2.1) (Langmead and Salzberg, 2012) in very-sensitive end-to-end mode. Discordant and mixed mappings were disabled and the maximum insert was set to 1000 bp. Three score-min parameters were used in different mapping runs: "L,-1.2,-1.2", "L,-0.6,-0.6" and "L,-0.3,-0.3", approximately equal to 20%, 10% and 5% mismatch rate respectively. The BAM files for the bulks were sorted, merged and indexed using SAMtools (v0.1.18; (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 VarScan mpileup2snp (v2.3.7; (Koboldt *et al.*, 2012)) with --strand-filter 0 and --output-vcf 1 for variant calling.

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 (bait homology and alignment length >=80%). Intersecting these regions (+/- 1000 bp) against the mapped reads using BEDTools gave the number of on-target reads. Mean read coverage to RLP/RLK genes was calculated from the previously generated BAM files using BEDTools coverage.

SNP filtering

SNPs were filtered using custom Java code (Supplementary Dataset S5) to retain informative SNPs present in both bulks and parents. SNPs were filtered based on expected allele ratios for responsive/non-responsive (responsive: Rr; non-responsive: rr). 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 non-responsive and 40-60% alternate allele for responsive. BEDTools intersect (v2.20.1; (Quinlan and 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 (Andolfo *et al.*, 2013). 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 SSR markers

The BAM file and VCF file for the filtered informative SNPs was visualized using Geneious R10 (Kearse *et al.*, 2012) (http://www.geneious.com). The primers were designed in Geneious R10, the PCR product should contain the informative SNP and have a size between 80-150 bp. Primers flanking the informative SNPs were manually selected on the conserved sequences of both parents, R and NR bulks. The HRM markers were tested on the parents and the F1 progenies with the following protocol for a 10 µL reaction mixture: (1 µL template (20 ng gDNA), 1 µL dNTP (5mM), 0.25 µl forward primer and 0.25 µl reverse primer (10 mM), 1 µL LCGreen® Plus+ (BioFire), 2 µL 5x Phire Buffer, 0.06 µL Phire taq, 4.44 µL MQ water). 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: 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 40s. The LightScanner® System (Biofire) was then used for measuring

and analyzing the melting curve. The primers used in this study are listed in Supplementary **Table S5**. 78 SSR markers as described in (Milbourne *et al.*, 1998) were used in this study **(Supplementary Table S4)**.

Accession numbers:

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

Acknowledgements:

This work was supported by a NWO-VIDI grant 12378 (X.L. D.W., V.G.A.A.V.), Short Term Scientific Mission (STSM) of COST Action SUSTAIN FA1208 (X.L.), China Scholarship Council (CSC) (X.L), the Rural & Environment Science & Analytical Services Division of the Scottish Government (I.H.) and the Biotechnology and Biological Sciences Research Council (BBSRC) through awards BB/L008025/1 (I.H.) and BB/K018299/1 (I.H.) We thank Dr. Hendrik Rietman for performing effectoromics screens, Isolde Bertram-Pereira for culturing *Solanum* plants, Henk Smid and Harm Wiegersma for help in the greenhouse. We thank Prof. Dr. Evert Jacobsen for reviewing the manuscript.

References

Andolfo G, Sanseverino W, Rombauts S, Van de Peer Y, Bradeen JM, Carputo 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

Arora S, Steuernagel B, Chandramohan S, Long Y, Matny O, Johnson R, Enk J, Periyannan S, Hatta MAM, Athiyannan N, Cheema J, Yu G, Kangara N, Ghosh S, Szabo LJ, Poland J, Bariana H, Jones JDG, Bentley AR, Ayliffe M, Olson E, Xu SS, Steffenson BJ, Lagudah E, Wulff BBH (2018) Resistance gene discovery and cloning by sequence capture and association genetics. bioRxiv

Austin RS, Vidaurre D, Stamatiou G, Breit R, Provart NJ, Bonetta D, Zhang JF, Fung P, Gong YC, Wang PW, McCourt P, Guttman DS (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. Annu Rev Phytopathol **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, Lees AK, Jones JD, Bryan GJ, Hein I (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. Theor Appl Genet **https://doi.org/10.1007/s00122-018-3078-6**

Clark MJ, Chen R, Lam HY, Karczewski KJ, Chen R, Euskirchen G, Butte AJ, Snyder M (2011) Performance comparison of exome DNA sequencing technologies. Nat Biotechnol **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, Aguilera-Galvez C, Cruz-Mireles N, Vleeshouwers VGAA, Kamoun S (2016) Nine things to know about elicitins. 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 Mol Biol **1578:** 337-353

Du J, Rietman H, Vleeshouwers VG (2014) Agroinfiltration and PVX agroinfection in potato and *Nicotiana benthamiana*. J Vis Exp**:** e50971

Du J, Verzaux E, Chaparro-Garcia A, Bijsterbosch G, Keizer LC, Zhou J, Liebrand TW, Xie C, Govers F, Robatzek S, van der Vossen EA, Jacobsen E, Visser RG, Kamoun

S, Vleeshouwers VG (2015) Elicitin recognition confers enhanced resistance to *Phytophthora infestans* in potato. Nat 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. Plant Journal **18:** 265-276

Fry W (2008) *Phytophthora infestans*: the plant (and *R* gene) destroyer. Mol Plant Pathol **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. Front Plant Sci **7:** 1887

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, McCombie WR (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, Dong GJ, Sang T, Han B (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, Zhou J, Bonierbale M, Hein I, Lindqvist-Kreuze H, Xie C (2018) Potato late blight field resistance from QTL dPI09c is conferred by the NB-LRR gene R8. J Exp Bot **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**

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, Cardle L, Hein I, Jones JDG (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. 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, Thierer T, Ashton B, Meintjes P, Drummond A (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. Nat 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

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, Kamoun S (2005) Patterns of diversifying selection in the phytotoxin-like *scr74* 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. Mol Gen Genet **259:** 233-245

Oh SK, Young C, Lee M, Oliva R, Bozkurt TO, Cano LM, Win J, Bos JIB, Liu HY, van Damme M, Morgan W, Choi D, Van der Vossen EAG, Vleeshouwers VGAA, Kamoun S (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

Pandey MK, Khan AW, Singh VK, Vishwakarma MK, Shasidhar Y, Kumar V, Garg V, Bhat RS, Chitikineni A, Janila P, Guo BZ, Varshney RK (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, Goverse 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, Chitikineni A, Vechalapu S, Kumar CVS, Sharma M, Ghanta A, Yamini KN, Muniswamy S, Varshney RK (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, Periyannan SK, Hernandez-Pinzon I, Witek K, Rouse MN, Yu GT, Hatta A, Ayliffe M, Bariana H, Jones JDG, Lagudah ES, Wulff BBH (2016) Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. Nature Biotechnology **34:** 652-655

Takagi H, Abe A, Yoshida K, Kosugi S, Natsume S, Mitsuoka C, Uemura A, Utsushi H, Tamiru M, Takuno S, Innan H, Cano LM, Kamoun S, Terauchi R (2013) QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. Plant Journal **74:** 174-183

Takken FLW, Luderer R, Gabriels SHEJ, Westerink N, Lu R, de Wit PJGM, Joosten MHAJ (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

Van Weymers PSM, Baker K, Chen XW, Harrower B, Cooke DEL, Gilroy EM, Birch PRJ, Thilliez GJA, Lees AK, Lynott JS, Armstrong MR, McKenzie G, Bryan GJ, Hein I (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, Rietman H, Huigen D, Hoekstra R, Goverse A, Vosman B, Jacobsen E, Visser RG (2011) 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, Pel MA, Kamoun S (2011) Understanding and exploiting late blight resistance in the age of effectors. Annu Rev Phytopathol **49:** 507-531

Wastie RL (1991) Breeding for resistance. In: Ingram, D.S. and Wiliams, D.S., Eds., *Phytophthora infestans*, the cause of late blight of potato. advances in plant pathology, Vol. 7, Academic Press Ltd., London, 193-224. Academic Press, London ; San Diego

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. J Exp Bot **68:** 4125-4135

Zou G, Zhai G, Feng Q, Yan S, Wang A, Zhao Q, Shao J, Zhang Z, Zou J, Han B, Tao Y (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. J Exp Bot **63:** 5451-5462

Supplementary data

For convenience, Supplementary Dataset S1-S5, Supplementary Table S1, S3 and S4 are deposited in 4TU.Centre: https://data.4tu.nl/repository/uuid:9032af40-6cd0-409d-be0d-76fe0f4ed576

Supplementary Dataset S1. Fasta file of the 349 LRR-RLK genes used for generating the RLP/KSeq enrichment library

Supplementary Dataset S2. Fasta file of the 450 LRR-RLP genes used for generating the RLP/KSeq enrichment library

Supplementary Dataset S3. 15 additional known RLP and RLK genes from Solanaceae species

Supplementary Dataset S4. The 2x bait library used in this study

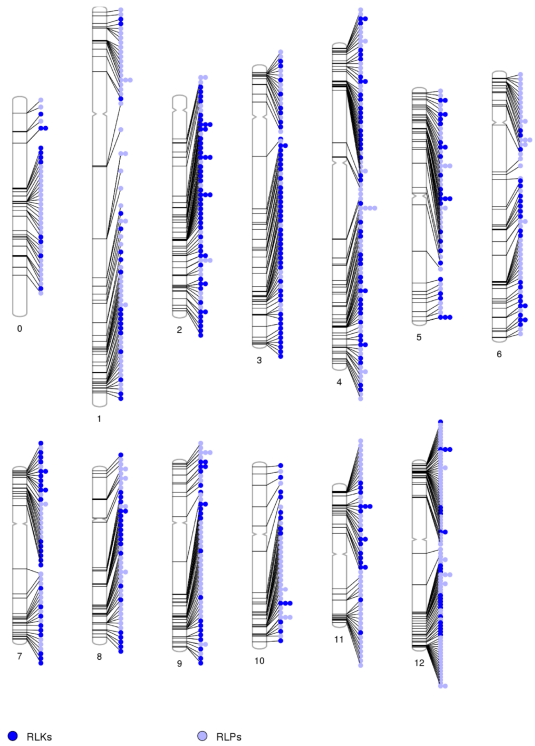
Supplementary Dataset S5. Java script for calling the informative SNPs

Supplementary Table S1. All tested Solanum genotypes for INF1 and SCR74 response by PVX agro-infection

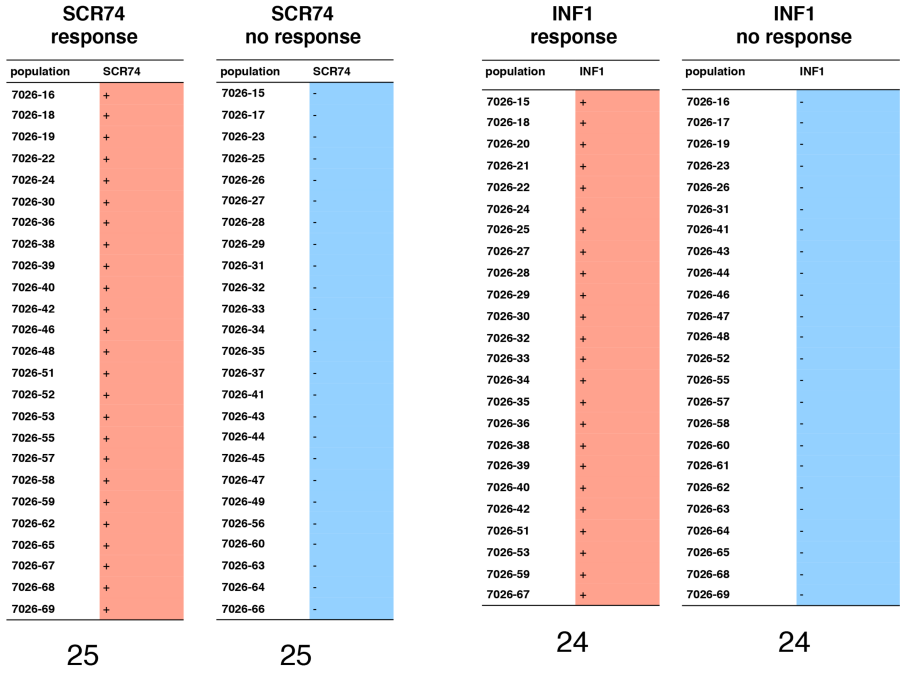
Supplementary Table S3. Summary of the SNP calling outputs under 10% mismatch criterial

Supplementary Table S4. SSR markers used in this study

Supplementary Figure S1. The position of predicted RLP/RLK genes on the reference genome



Supplementary Figure S2. Population 7026 segregates to both INF1 and SCR74 response



SCR74 response

population	SCR74
7026-16	+
7026-18	+
7026-19	+
7026-22	+
7026-24	+
7026-30	+
7026-36	+
7026-38	+
7026-39	+
7026-40	+
7026-42	+
7026-46	+
7026-48	+
7026-51	+
7026-52	+
7026-53	+
7026-55	+
7026-57	+
7026-58	+
7026-59	+
7026-62	+
7026-65	+
7026-67	+
7026-68	+
7026-69	+

25

SCR74 no response

population	SCR74
7026-15	-
7026-17	-
7026-23	-
7026-25	-
7026-26	-
7026-27	-
7026-28	-
7026-29	-
7026-31	-
7026-32	-
7026-33	-
7026-34	-
7026-35	-
7026-37	-
7026-41	-
7026-43	-
7026-44	-
7026-45	-
7026-47	-
7026-49	-
7026-56	-
7026-60	-
7026-63	-
7026-64	-
7026-66	-

25

INF1 response

population	INF1
7026-15	+
7026-18	+
7026-20	+
7026-21	+
7026-22	+
7026-24	+
7026-25	+
7026-27	+
7026-28	+
7026-29	+
7026-30	+
7026-32	+
7026-33	+
7026-34	+
7026-35	+
7026-36	+
7026-38	+
7026-39	+
7026-40	+
7026-42	+
7026-51	+
7026-53	+
7026-59	+
7026-67	+

24

INF1 no response

population	INF1
7026-16	-
7026-17	-
7026-19	-
7026-23	-
7026-26	-
7026-31	-
7026-41	-
7026-43	-
7026-44	-
7026-46	-
7026-47	-
7026-48	-
7026-52	-
7026-55	-
7026-57	-
7026-58	-
7026-60	-
7026-61	-
7026-62	-
7026-63	-
7026-64	-
7026-65	-
7026-68	-
7026-69	-

24

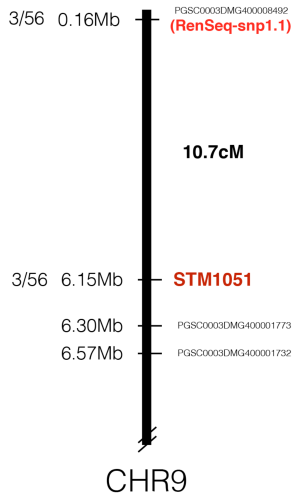
Supplementary Table S2. qPCR for confirming the enrichment output

Primer pairs	Pool	mean ΔCT	mean fold enrichment
ELR-primer1	Pool1	6.133	70.196
ELR-primer2	Pool1	6.382	83.404

Supplementary Table S5. Primers used in this study

Name	Sequence (5'-> 3')	Notes
ELR-qPCR-F1	AGCATGCTTCAAAAGCTAACCG	qPCR after RLK/ RLP enrichment
ELR-qPCR-R1	GCAAAGGCAGGCTAAACACC	qPCR after RLK/ RLP enrichment
ELR-qPCR-F2	TCTGGTTCCATTCCTGATTCCA	qPCR after RLK/ RLP enrichment
ELR-qPCR-R2	TCCAAACGTGTGAGGTGCTT	qPCR after RLK/ RLP enrichment
STM1051-F	TCCCCTTGGCATTTTCTTCTCC	SSR marker
STM1051-R	TTTAGGGTGGGGTGAGGTTGG	SSR marker
SNP1.1-F	ATTCAGTATCCACAATGTACG	RLK/ RLP RenSeq marker
SNP1.1-R	ACGTTAAGCTTTTCAACTTAC	RLK/ RLP RenSeq marker

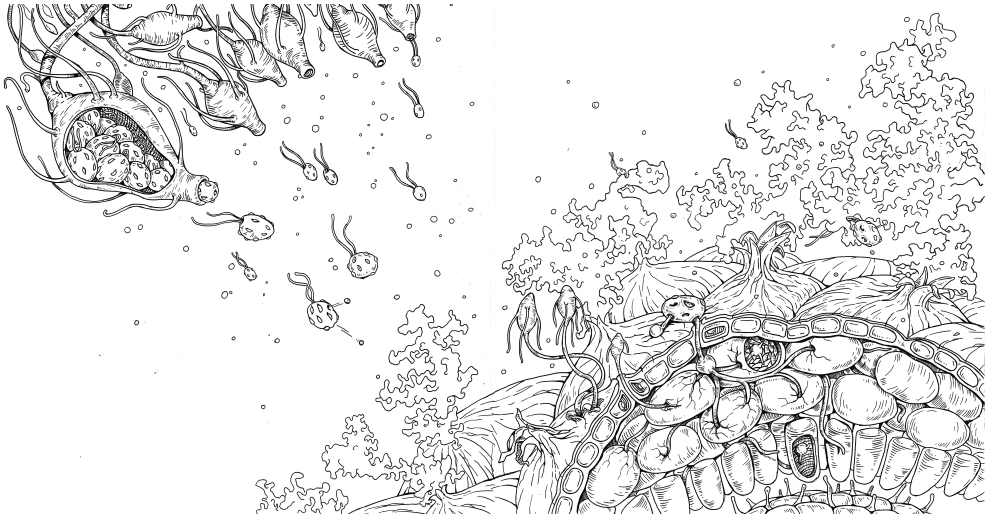
Supplementary Figure S3. The mapping interval of SCR74 receptor



Chapter 4

The candidate receptor of the apoplastic effector SCR74 of *Phytophthora infestans* shows similarity to G-type LecRK proteins that are involved in self-incompatibly in Brassicaceae

Xiao Lin, Doret Wouters, Emmanouil Domazakis, Richard G. F. Visser, Sophien Kamoun, Vivianne G. A. A. Vleeshouwers



Manuscript in preparing

Abstract:

Potato late blight has led to great yield losses of global potato production in the past and at present. The disease is caused by the oomycete pathogen *Phytophthora infestans*. The resistance (*R*) genes cloned so far from potato against late blight represent intracellular nucleotide-binding, leucine-rich repeat (NLR) receptors, however, most NLR genes have been overcome in the field. We argue that exploring surface immunity might help to achieve more durable resistance to *P. infestans*. SCR74 is a family of apoplastic effectors that can be found as one or multiple copies in all strains of *P. infestans*. Here, we fine-mapped a putative SCR74 receptor within a 73kb interval in an F1 segregating population of *Solanum microdontum*. The physical map of the SCR74 responsive genotype GIG362-6 was constructed by isolating and sequencing several BAC clones covering that region. We found that the identified region contains several tandem G-type LecRK (*G-LecRK*) genes with copy number variation and sequence exchanges. RNAseq data indicates that these *G-LecRKs* are up-regulated after *P. infestans* treatment. However, co-expression of these G-LecRKs with SCR74 did not recover the cell death phenotype on SCR74 non-responsive plants. In summary, our findings will lead to cloning a novel potato surface immune receptor recognizing apoplastic effectors from *P. infestans*. Furthermore the striking similarity between the sporophytic self-incompatibility system (SI) in *Brassica* and the plant immunity system in *Solanum* is discussed.

Introduction

Based on their subcellular localization, plant immune receptors can be classified into cytoplasmic receptors, which normally belong to cytoplasmic resistance (*R)* genes with a nucleotide-binding domain and leucine-rich repeats (*NLR*), or to surface immune receptors, which are receptor-like proteins (RLPs) or receptor-like kinases (RLKs) (Win *et al.*, 2013). Surface immune receptors can represent pattern recognition receptors (PRRs) that recognize conserved microbial- or pathogen-associated molecular patterns (MAMPs/PAMPs), for example ELR, FLS2, EFR, CORE, and RLP23 which recognize elicitins, flg22, elf20, csp22 and nlp20, respectively (Kunze, 2004; Chinchilla, 2006; Zipfel *et al.*, 2006; Albert *et al.*, 2015; Wang *et al.*, 2016; Gómez-Gómez *et al.*, Du *et al.*, 2015). In addition, some surface immune receptors can recognize apoplastic effectors, like the CF-9, CF-4, and VE1 from tomato, which recognize AVR9, AVR4, and AVE1, respectively, of the fungal pathogens *Cladosporium fulvum* and *Verticilium daliae* (Thomas *et al.*, 1997; Kawchuk *et al.*, 2001). The surface immune receptors form the first layer of the plant innate immune system are believed to confer broader spectrum and more durable resistance than the NLR receptors (Lacombe *et al.*, 2010).

Most identified surface immune receptors contain an extracellular leucine-rich repeat (eLRR), however, some receptor-like kinases contain different extracellular domains, such as a lectin domain or an EGF-like domain, e.g. TaWAK from wheat (Boutrot and Zipfel, 2017). Lectins are carbohydrate-binding proteins (Lis and Sharon, 1998), the RLKs with an extracellular lectin motif are called LecRKs. These can be classified into four types, based on the lectin domain: 1) The G-type lectin domain (resembling the *Galanthus nivalis* agglutinin, GNA); 2) L-type lectin (legume-like) 3) C-type lectin (calcium-dependent); 4) LysM LecRKs contain lysin motifs, which are thought to bind bacterial peptidoglycan (PGN) and fungal chitins (Lannoo and van Damme, 2014; Wang and Bouwmeester, 2017).

The G-LecRKs were previously called S-domain RLK (SRLK) or B-lectin receptor kinases, derived from the *Brassica oleracea* well-characterized *SRK* gene. SRK is a stigma-specific plasma membrane-localized protein (Stein *et al.*, 1991; Stein *et al.*, 1996), and involved in the sporophytic self-incompatibility system. SRK recognizes the S-haplotype-specific small-cysteine rich protein SCR/SP11 from self-pollen. The recognition will lead to a self-incompatibility response, a cell death response, that will lead to rejection of the self-pollen (Takayama *et al.*, 2001; Chookajorn *et al.*, 2004). Various G-LecRKs have also been found to be involved in plant immunity, for example I3 from tomato confers resistance to the tomato wilt disease caused by *Fusarium oxysporium;* Pi-d2 that confers resistance to rice blast caused by fungal pathogen *Magnaporthe oryzae*; OsLecRK1-3 conferring resistance to the most

destructive pest of rice- the brown hopper, and LORE from *Arabidopsis* that can mediate lipopolysaccharide (LPS) sensing (Chen *et al.*, 2006; Catanzariti *et al.*, 2015; Liu *et al.*, 2015; Ranf *et al.*, 2015). Therefore, the G-LecRKs are versatile cell surface receptors in plants, which are important for sensing the self- or non-self- components from the apoplast.

Potato is an important food crop around the world, but the yield is threatened by pathogens like the pathogenic oomycete *Phytophthora infestans* that causes late blight. For resistance breeding, major *R* genes from wild potatoes have been introduced into modern cultivars, but practically, they have always been quickly defeated by this fast-evolving oomycete. Exploring another layer of the immunity system might help to achieve more durable resistance. So far, only one surface immune receptor, called ELR, has been cloned which can recognize the conserved elicitor INF1 from *P. infestans*. Transformation of the cloned *ELR* gene into the susceptible cultivar Désirée enhanced resistance to late blight under lab conditions (Du *et al.*, 2015). Cloning more surface immune receptors from potato and stacking them together with major *R* genes is believed to help to achieve durable resistance.

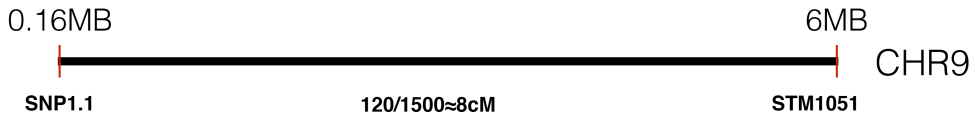
Oomycetes can secrete different types of apoplastic effectors to facilitate their colonization in the host. The PcF/ SCR effectors represent a family of small, cysteine-rich (SCR) proteins. SCR74 variants are under strong positive selection pressure and are considered to co-evolve with the receptors from the host (Liu *et al.*, 2005) (**Chapter 2**). To genetically map the SCR74 receptor, in **Chapter 3**, we developed RLP/RLK enrichment sequencing (RLP/KSeq), and successfully located the SCR74 receptor on the top of linkage group 9.

Here we fine mapped the putative SCR74 receptor in an extended population. We also generated a BAC library of the responsive genotype GIG362-6, isolated and sequenced the BAC clones covering the mapping interval. We found that the putative SCR74 receptor locates in a *G-LecRK* gene cluster, which may reveal the links between plant immune system and self-incompatibility system.

Results

Recombinants screening leads to a mapping interval of 8cM

A



B

Progeny ID	SNP1.1	SCR74 Response	STM1051
43	c	0	0
168	c	0	0
562	c	0	0
282	c	0	0
1287	d	1	1
36	d	1	1
48	d	1	1
291	d	1	1
2562	d	0	1
158	d	0	1
1231	d	0	1
1211	d	0	1
1573	c	1	0
1943	c	1	0
22	c	1	0
665	c	1	0
GIG362-6	a	1	1
MCD360-1	b	0	0

Figure 1: Recombinant screening of population 7026 by flanking markers SNP1.1 and STM1051.

A. The top 6MB of chromosome 9 of the potato DM reference genome is shown by a black bar. Flanking markers SNP1.1 (0.16Mb) and STM1051 (6MB) were used for recombinant screening. In total 120 recombinants were found, the SCR74 receptor was mapped within a 6 MB (8cM) region. **B.** Example of flanking markers screening, 16 progenies of population 7026 are shown. For SSR marker STM1051, 0 or 1 indicates absence/ presence of a specific band on the PAGE gel **(Chapter 3, Figure 5C)**. For HRM marker SNP1.1 a, b, c or d represent the pattern of the melting curve. The SCR74 response of each progeny plant after PVX agro infection are indicated by 0 (non-responsive) and 1 (responsive).

To identify the receptor of SCR74-B3b, we started a fine mapping strategy using population 7026, which was generated by crossing GIG326-2 (responsive) x MCD360-1 (non-responsive). The putative SCR74 receptor was mapped to a 10.7cM region, on the chromosome 9 of DM genome (**Chapter 3**). We extended the mapping population to 1500 progeny plants and genotyped them with two flanking markers: SNP1.1 and STM1051. As a result, 120 recombinants were found between these two markers **(Figure 1)**. Thus, the SCR74 receptor flanking region was reduced from 10.7cM to 8 cM (**Figure 1**).

SNP markers development for fine mapping of the putative SCR74 receptor

For further fine mapping of the putative SCR74 receptor, we designed additional markers between snp1.1 and STM1051 by two rounds of PCR analysis **(Figure 2)**. In brief, we extracted all the exons within the target interval from the potato reference genome, filtered all the exons larger than 1500 bp, and designed primers for 1200 bp amplicons **(Figure 2B)**. The primers were tested on the mapping parents GIG362-6 and MCD360-1, and the successfully amplified PCR products were sent for sequencing from both directions. The sequencing output (.abi) files were analyzed in Geneious R10, and the SNPs only present in GIG362-6, but not in MCD360-1, were selected. A new primer pair, which is flanking the SNPs, were designed for the following high resolution melting (HRM) SNP genotyping by LightScanner (**Figure 2C**; and see **Materials and Methods**). Using this method, many new HRM markers were designed and used for recombinant screening on the F1 progeny plants.

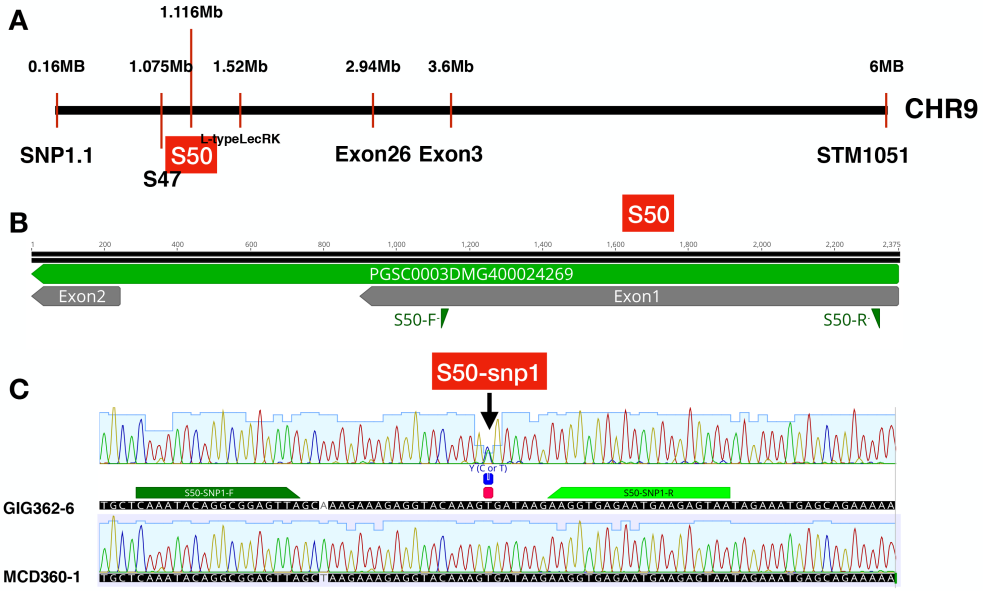


Figure 2. SNP marker development for fine mapping of the SCR74 receptor.

A. The annotated exons between the flanking markers snp1.1 and STM1051 from the potato DM reference genome were used for marker development.

B. An example of marker S50, S50-F and S50-R were designed on exon1 of PGSC0003DMG400024269, which is located on 1.116Mb of the DM reference genome.

C.The PCR products from both parents, GIG362-6 and MCD360-1, were aligned and a new marker, called S50_snp1, was designed based on the SNPs present in the responsive parent GIG362-6.

Fine mapping of the SCR74 receptor to a 355kb (0.13cM) interval

135kb (0.13cM)

1/1500 1/1500

Progeny ID	snp1.1 (0.16Mb)	S47 (1.075Mb)	S50 (1.165Mb)	S55 (1.235Mb)	SCR74 Response	S105-snp2 (1.300Mb)	exon 19 (1.376Mb)	LecRK1snp1 (1.52Mb)	LecRK2snp2 (1.52Mb)	STM1051 (6Mb)
43	c	a	d	b	0	a	a	d	b	0
168	c	a	d	b	0	a	a	d	b	0
562	c	a	d	b	0	a	a	d	b	0
1381	c	a	d	b	0	a	a	d	b	0
1459	c	a	d	b	0	a	a	d	b	0
433	c	b	d	b	0	a	a	d	b	0
204	c	b	c	b	0	a	a	d	b	0
1573	c	b	c	a	1	a	a	d	b	0
1943	c	b	c	a	1	b	a	a	d	0
1435	c	b	c	a	1	b	b	d	b	0
1822	c	b	c	a	1	b	b	c	a	0
1823	c	b	c	a	1	b	b	c	a	0
291	d	b	c	a	1	b	b	c	a	1
1016	d	b	c	a	1	b	b	c	a	1
1030	d	b	c	a	1	b	b	c	a	1
1225	d	a	c	a	1	b	b	c	a	1
639	d	a	c	a	1	b	b	c	a	1
158	d	a	d	b	0		a	c	a	1
1684	d	a	d	b	0		a	c	a	1
1324	d	a	d	b	0		a	d	a	1
1231	d	a	d	b	0		a	d	b	1
1211	d	a	d	b	0		a	d	b	1

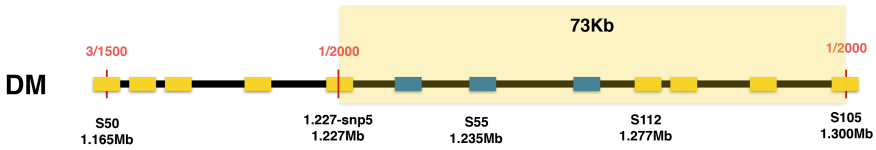
Figure 3. Fine mapping of the putative SCR74 receptor within a region of 135kb (0.13cM)

The interval of 0.16Mb to 6Mb of chromosome 9 of potato DM reference genome is shown. The recombinants between the markers SNP1.1 and STM1051 were selected for fine mapping. The new markers S47, S50, S55, S105-snp2, exon19, LecRK1snp1, LecRK2snp2 were tested on the F1 progenies. The ID of the recombinants are shown in the table. The genotyping output of HRM markers is shown by a, b , c or d. The genotyping output of the SSR markers is indicated by 0, 1. The SCR74 response of each recombinant after PVX agro infection is indicated by 0 (non-responsive) and 1 (responsive), the genotyping output are also visualized by blue (SCR74 non-responsive haplotype) or red (SCR74 responsive haplotype). The grey box indicates the new mapping interval, between markers M50 and S105-snp2, M55 is co-segregating with the SCR74 response.

After using the newly designed HRM markers among the 120 recombinant progenies between the markers snp1.1 and STM1051, a new mapping interval was determined by the markers S50 (1.165Mb) and S105-snp2 (1.520Mb) **(Figure 3)**. Genotype 7026-204 recombines between marker S50 and the SCR74 response, determines the left boundary of the mapping interval. Similarly, a recombination event between marker S105-snp2 and SCR74 response was detected in genotype 7026-1573, which determines the right boundary.

Therefore, based on the DM reference genome, the SCR74 receptor was mapped to a 135kb interval with a genetic distance of around 0.13cM (2/1500) **(Figure 3)**.

To further narrow down the mapping interval of the receptor to be able to clone it, 500 extra seeds were sown for ultra-fine mapping of the receptor. A new recombinant event on the left side was detected by marker M1.227-snp5. Hence, the mapping interval was further narrowed down to a 73kb region based on the DM reference genome **(Figure 4)**, between markers 1.227-snp5 (1.227Mb) and S105 (1.300Mb), the genetic distance is around 0.1 cM (2/2000) **(Figure 4)**.



ID	Annotation	Start	End
PGSC0003DMG400024261	Putative reticulata-related 1-like	1.227.226	1.233.877
PGSC0003DMG400024260	G-type LecRK	1,235,025	1,237,340
PGSC0003DMG400024259	G-type LecRK	1,250,099	1,252,519
PGSC0003DMG400024258	G-type LecRK	1,261,565	1,262,491
PGSC0003DMG400008887	Serine/threonine-protein kinase ATG1c	1,273,335	1,280,882
PGSC0003DMG400008814	Prenylated rab acceptor family protein	1,282,266	1,289,115
PGSC0003DMG400008966	Gene of unknown function	1,287,639	1,288,530
PGSC0003DMG400008888	Uracil phosphoribosyltransferase	1,300,722	1,305,212

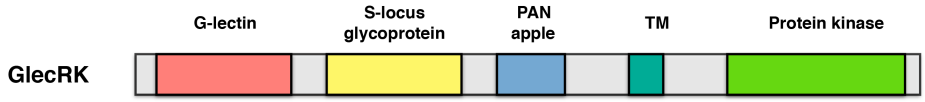
Figure 4. The SCR74 receptor is mapped to a 73kb interval of the DM genome for which a list of candidate genes is presented.

The candidate SCR74 receptor is mapped between marker 1.227-snp5 and S105, which represents a 73kb interval based on the DM reference genome. The genetic distance is 0.1cM (2/2000). 8 genes are annotated in DM, including 3 *G-LecRK* genes (green blocks). Other predicted candidate genes are shown by yellow blocks.

Identification of candidate genes on the potato DM reference genome

To identify the candidate SCR74 receptor, we first checked the annotated genes on the DM reference genome. We found that only 8 genes reside within the mapping interval, and it is

noteworthy that 3 G-type LecRKs (*G-LecRK*) genes are located in this region **(Figure 4)**. These G-LecRKs are receptor-like kinases with a G-type lectin domain. The predicted G-LecRK proteins have a N-terminal G-lectin domain, a PAN-apple domain, followed by a transmembrane (TM) domain and a C-terminal serine/ threonine kinase domain **(Figure 5)**. Many previously cloned *G-LecRK* genes are involved in plant immunity or self-incompatibiity **(Figure 5)**. So we assumed the *G-LecRK* genes to be good candidates for the SCR74 receptor.



G-LecRK	Species	Pathogen/ Pollen	Target	Reference
I3	Tomato	*Fusarium*	Six1 (AVR3)	(Catanzariti, 2015)
Pi-d2	Rice	*Magnaporthe grisea*	unknown	(Chen, 2006)
LORE	Arabidopsis	Gram-negative bacteria	LPS	(Ranf, 2015)
OsLecRK1-3	Rice	Brown planthopper (BPH)	unknown	(Liu, 2015)
SRLK	Brassica	Pollen	SP11/ SCR	(Kemp, 2007)

Figure 5. Structure of the candidate G-LecRK and some examples of functional G-LecRK proteins from literature.

The predicted G-LecRK protein contains 5 domains, including the G-lectin domain (red), S-locus glycoprotein domain (yellow), PAN-apple domain (blue), a transmembrane domain (cyan) and a protein kinase domain. A list of five known G-LecRKs which are involved in plant immunity or sporophytic self-incompatibility are presented.

Besides, 5 other candidate genes are in the mapping interval. PGSC0003DMG400024261 is annotated as "putative reticulate-related 1 like gene" and the gene ontology (GO) shows that this gene may be involved in leaf development, photoperiodism and response to reactive oxygen species. The function of the reticulate-related genes is not entirely clear but they seem to be involved in interconnected photoperiodic growth, amino acid homeostasis, and reactive oxygen species metabolism during *Arabidopsis* leaf growth, and they are located in the chloroplast membrane (Pérez-Pérez *et al.*, 2013). So far, no evidence is provideded that this gene could be involved in plant immunity or effector recognition. PGSC0003DMG400008887 belongs to the receptor-like kinase super family, which has a serine/threonine-protein kinase domain which may be involved in ATP binding and in protein kinase activity. The homologs of this gene were annotated as serine/threonine-protein

kinase ATG1 (Autophagy-related protein). The function of this gene is that it is involved in the cytoplasm to vacuole transport (Cvt) and is found to be essential in autophagy.

Collectively, we consider the *G-LecRK* genes are the most likely candidates for the putative SCR74 receptor, and we started to conduct further experiments with these candidates, however, until we have more evidence, we cannot rule out the other candidate genes.

The SCR74 receptor resides on a highly diverse G-LecRK locus

To obtain the physical map of genotype GIG362-6, which recognizes SCR74-B3b, a bacterial artificial chromosome (BAC) library was generated. New markers for 500 bp amplicons were designed based on the DM reference genome and used for BAC screening **(Materials and Methods)**. Three BACs, covering the mapping interval were isolated, i.e. BAC03-H3, BAC01-3A and BAC01-2B, from both haplotypes of GIG362-6. Based on our marker analysis, the candidate SCR74 receptor is expected to be located on haplotype 1, represented by BAC03-H3 and BAC01-3A **(Figure 6)**.

To see if the *G-LecRK* loci is conserved in different Solanaceae species, we analyzed the *G-LecRK* loci in different genomes of Solanaceae plants. As described previously, there are 3 *G-LecRKs* in the DM reference genome potato. The *G-LecRK* loci were also extracted from the genome of pepper (*C. annuum* v2.0), eggplant (*S. melonggena* Line '67/3') and tomato (*S. lycopersicum* v2.4). We found that there is only one copy of a *G-LecRK* homolog in pepper (*CaGlecRK*) and tomato (*Solyc09g011330.1*), and two copies in eggplant (*324080.1* and *324070.1*) **(Figure 6)**.

Surprisingly, there are 2 and 5 *G-LecRK* genes on the 2 haplotypes of GIG362-6, gigG-LecRK 1 and 2 are on the haplotype 1, 2B-1, 2B-2, 2B-3, 2B-4 and 2B-5 are on haplotype 2 **(Figure 6)**. To see whether this copy number variation (CNV) also exists in other wild potato genomes, we analyzed this locus in 2 sequenced wild potato genomes, *Solanum verrucosum* (personal communication with Dr. Glenn Bryan, James Hutton Institute, Dundee) and a recently published *Solanum chacoense* M6 v4.1 (Leisner *et al.*, 2018). We found 4 full-length or partial *G-LecRK* genes to be present in the *S. verrucosum* genome, we

named them as *ver-24260*, *ver-24259*, *ver-24258* and *ver-partial*. In the *S. chacoense* M6 genome, we detected 4 partial (*chc-1*, *chc-2*, *chc-3* and *chc-6*) and 3 full length *G-LecRK*

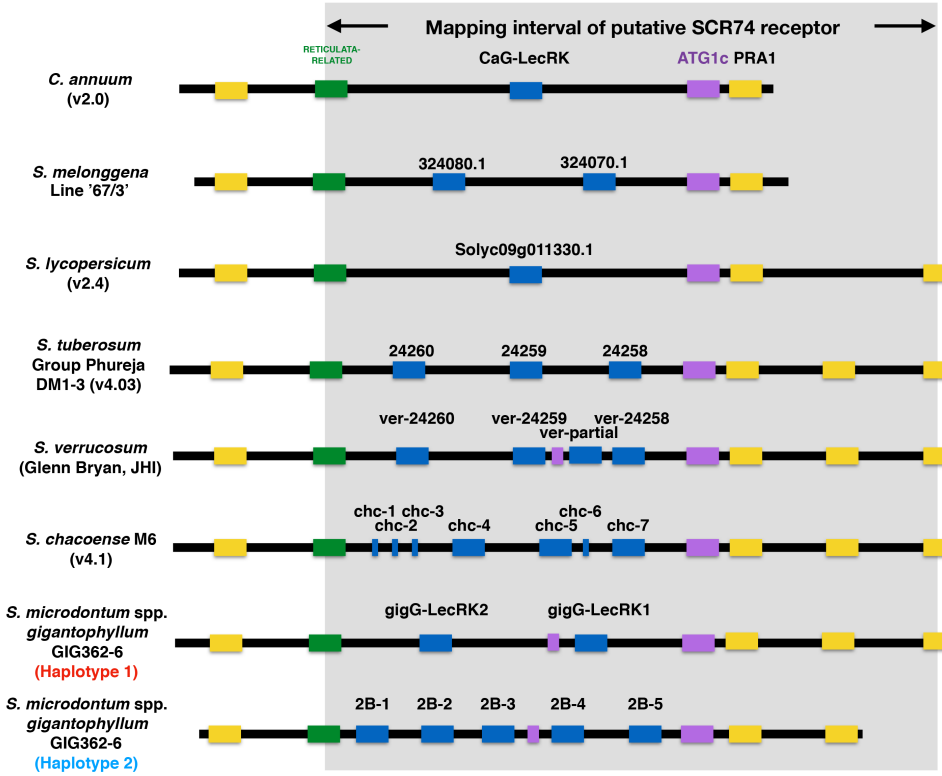


Figure 6. The candidate SCR74 receptor is located on a highly diverse *G-LecRK* locus.

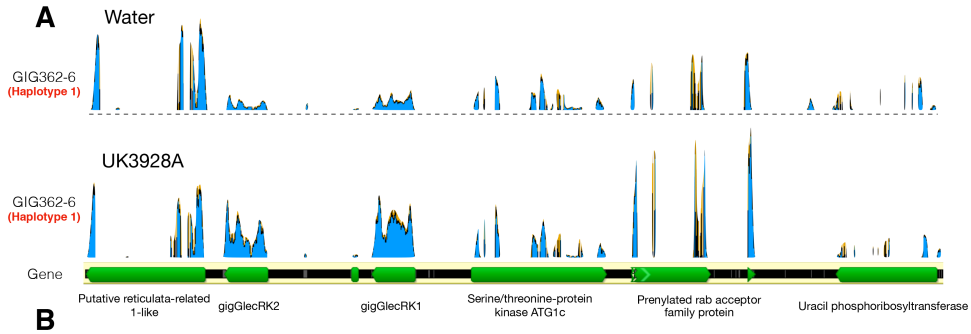
G-LecRK locus of chromosome 9 in pepper (*C. annuum*), eggplant (*S. melongena*), tomato (*S. lycopersicum*), DM potato (*S. tuberosum*), *S. verrucosum*, *S. chacoense* M6, and of 2 haplotypes from *S. microdontum* spp. *gigantophyllum* (genotype GIG362-6) are shown. The genes are shown by colored blocks, the mapping interval is shown by grey shade. The *G-LecRK* genes are shown as blue blocks.

genes (*chc-4*, *chc-5* and *chc-7*). Our data indicates that the *G-LecRK* loci are highly diverse in potatoes.

RNAseq data shows differential expression of the candidate genes

To evaluate the gene expression level of the candidate genes during *P. infestans* infection, we performed an RNAseq experiment on the mapping parents GIG362-6 and MCD360-1. The leaves of GIG362-6 and MCD360-1 plants were detached and spot-inoculated with *P. infestans* isolate UK3928A. A water inoculation was used as control. Leaf discs were sampled

for RNA isolation 48 hours post inoculation (hpi). Four samples were included in this experiment, namely GIG362-6 with water or UK3928A treatment, and MCD360-1 with water and UK3928A treatment. The RNAseq reads from each sample were cleaned and trimmed, then mapped to the BAC sequences to the 2 haplotypes of GIG362-6 (**Materials and Methods, Figure 7**).



B

Treatment	Description	Putative reticulata related 1-like	gigG-LecRK2	gigG-LecRK1	ATG1c-like	Prenylated rab acceptor family protein	Uracil phosphoribosyl transferase
MCD_water	Raw read count	1452.1	12.0	47.5	680.0	1566.1	1402.0
	TPM	43502	949.2	3753.3	18172.8	73194.7	49340.1
MCD_UK3928A	Raw read count	1753.7	146.8	218.0	679.7	2471.5	111.7
	TPM	28606.1	6295.7	9387.3	9890.3	62893.5	21304.2
GIG_water	Raw read count	1268.5	145.1	230.5	574.8	1006.2	1005.3
	TPM	31691.7	9535.3	15203.4	12810.1	39217.0	29503.2
GIG_UK3928A	Raw read count	889.5	516.4	704.7	595.0	1811.8	673.2
	TPM	15306.6	23373.1	32019.0	9134.6	48639.8	13609.6
GIG_water vs GIG_UK3928A	Differential expression p-value	0.39	0.015	0.0056	1	0.12	0.27
	Differential expression ratio	-1.5	3.4	2.9	-1	1.7	-1.6
MCD_UK3928A vs GIG_UK3928A	Differential expression p-value	0.074	0.011	0.0018	0.97	0.15	0.25
	Differential expression ratio	-1.9	3.6	3.3	-1.1	-1.3	-1.6

Figure 7. Differential expression of candidate genes on haplotype 1 of GIG362-6.

A. The mapping interval on GIG362-6 with the RNAseq reads from GIG362-6, mapped to haplotype 1 (BAC03-H3 and BAC01-3A). The blue peaks present the coverage of RNAseq reads. The samples of water treatment and UK3928A inoculation are shown. The 2 G-LecRKs are up-regulated after UK3928A infection. **B.** Differential expression of candidate genes in the mapping interval. The raw read counts and transcript per million (TPM) from each sample as well as the differential expression ratio and the p-value are shown for the four treatments of GIG_water vs GIG_UK3928A and MCD_UK3928A vs GIG_UK3928A. The up-regulated genes are highlighted by red, the down-regulated genes are highlighted by blue. The significant (P<0.05) differential expression p-values are shown in red font, otherwise in blue font.

The RNAseq data shows that the two *G-LecRK* genes are up-regulated after *P. infestans* infection compared to the water controls. The raw read counts of gigG-LecRK2 are 290 (water treatment) and 1029 (blue13 treatment), the normalized transcripts per million (TPM) are 2729.1 vs 10212.6. Transcripts of *gigG-LecRK1* are 460 (water treatment) and 1409 (blue13 treatment), the normalized transcripts per million (TPM) are 4319.6 vs 13958.5. We also mapped the RNAseq reads from MCD360-1 to the GIG362-6 BAC clone. The

expression of *G-LecRKs* from GIG362-6 are much higher than those from MCD360-1, for *gigG-LecRK2*, by comparing the expression after UK3928A infection, the raw read count from MCD360-1 are only 293.5, which is much lower than the expression in GIG362-6, with the raw read counts of 1029 (Normalized TPM is 2334 vs 9284).

Phylogenetic analysis of potato G-LecRK proteins

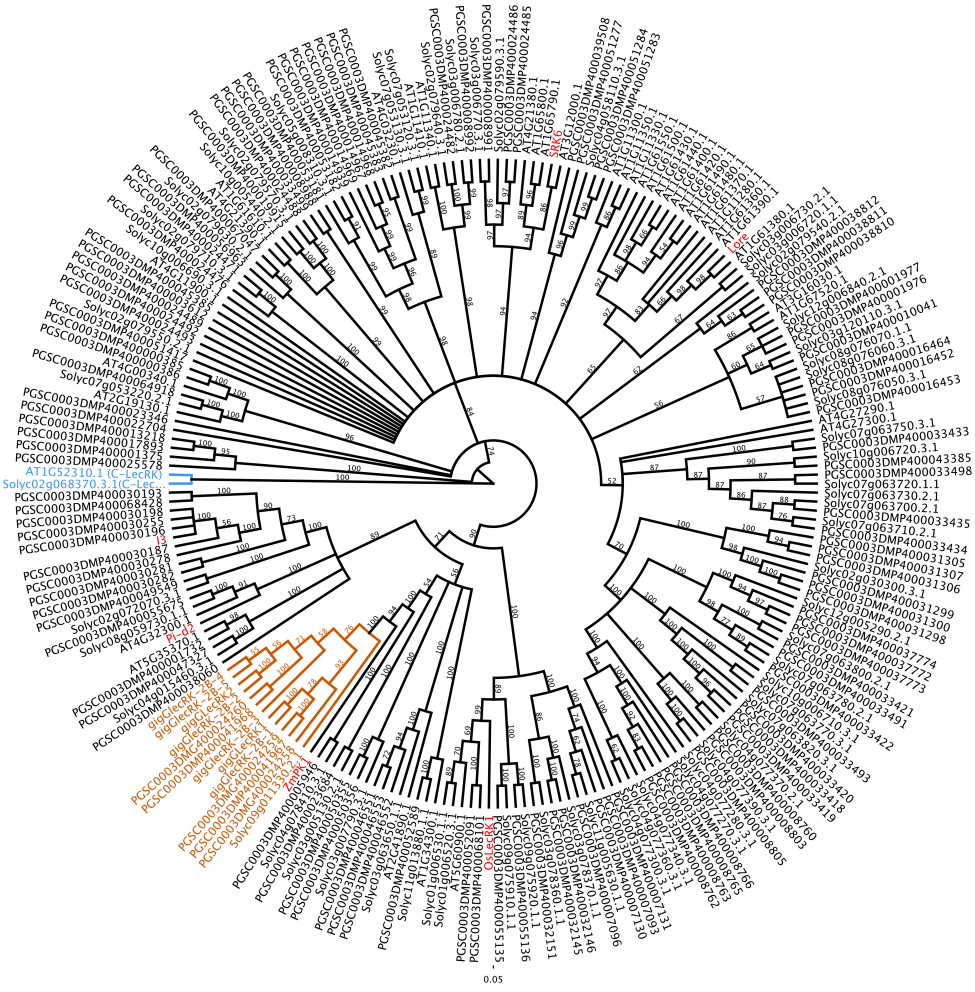


Figure 8. A NJ tree of all predicted G-LecRKs from potato, tomato and *Arabidopsis*

Complemented with the G-LecRKs from genotype GIG362-6 (earthy yellow). 6 well-characterized G-LecRKs (ZmPK1, SRK6, Pi-d2, OsLecRK1, Lore and I3, **Figure 5**) were included in the tree (red). Two C-LecRKs from tomato and *Arabidopsis* were used as outgroup (blue).

In potato, no functional G-LecRK have been characterized so far. To have an overview of the *G-LecRK* family in potato, we predicted all the G-LecRK proteins from the DM reference genome **(Materials and Methods)**. In total, 170 full-length or partial G-LecRK proteins were predicted, which is much higher than the number of G-LecRK proteins in tomato (93) (Teixeira *et al.*, 2018). This suggests that the G-LecRK proteins are expanded in potato.

To understand the phylogeny of the these G-LecRK proteins, We constructed a NJ tree of the 170 predicted G-LecRKs from DM, and added our 7 G-LecRK proteins from genotype GIG362-6, 49 G-LecRKs from *Arabidopsis,* 93 G-LecRKs from tomato, as well as 6 well-characterized G-LecRKs including ZmPK1, SRK6, Pi-d2, OsLecRK1, Lore and I3 from various plants. Two C-LecRKs, AT1G52310.1 and Solyc02g068370.3.1 from *Arabidopsis* and tomato, respectively, were used as outgroup **(Figure 8)**.

Our results indicate that the candidate SCR74 receptors are clustering in a specific clade, together with the homologs in the DM genome and the homolog (Solyc09g11130) in tomato (**Figure 6** and **Figure 8**). Surprisingly, ZmPK1 from maize clusters very close to our candidate G-LecRKs.

Functional test of the candidate G-LecRK genes and SCR74

As described above, many evidences pointed towards the *G-LecRK* as the most promising candidates for the SCR74 receptor.

To test this hypothesis, we firstly cloned the *G-LecRK* genes from gDNA of GIG362-6 **(Table 1)**. We cloned 3 *G-LecRK* genes, and named them *gigG-LecRK5-1*, *gigG-LecRK6-1* and *gigG-LecRK8-1*. The genes were cloned into the Gateway binary expression vector pK7WG2, and agroco-infiltrated with SCR74-B3b-pK7WG2 on *N. benthamiana*, potato cultivar Desiree, Bintje, and a wild potato *Solanum hjertingii* HJT349-3. Co-infiltration of R3b with Avr3b, and the empty pK7WG2 vector were included as positive and negative controls, respectively. However, no visible cell death were detected in any of the co-infiltrated these gigG-LecRKs with SCR74-B3b (data not shown). Later, when the 2 haplotypes of GIG362-6 covering the mapping interval were sequenced, we found that the 3 *G-LecRK* genes are probably chimeric genes, that may be artificial products during the PCR.

It is noteworthy that the *G-LecRK* genes are very difficult to clone in the *E. coli* system. We suspect that they are toxic to *E. coli*. Based on our observation, *E. coli* tend to delete or mutate the plasmid with these *G-LecRK* genes. When the BAC clones were sequenced, attempts were made to clone the *G-LecRK* genes from the BAC DNA. The same issue occurred, some genes

were found to be artificial chimeras and some were very difficult to handle in *E.coli*. By growing the *E. coli* at room temperature to slow the plasmid replication, we got some gigG-LecRK- pK7WG2 constructs, which are with very low concentration but correct genes were verified by sequencing **(Table 1)**. These constructs were transformed to *Agrobacterium* strain AGL1. Again, upon co-agroinfiltration with SCR74-B3b in different plants, no cell death responses were detected (data not shown). However, we cannot rule out that the constructs may not be suitable for the functional tests, due to their potential toxic nature in *E.coli* cells during cloning.

Table 1. Overview of *G-LecRK* from *S. microdontum gigantophyllum* GIG362-2

Gene ID	Cloned from	Notes	Vector	ATTA results	Tested plants
gigG-LecRK5-1	GIG362-6 gDNA	Artificial chimeric	PCR8, PK7WG2	No cell death when co-express with SCR74-B3b	*N. benthamiana*, Bintje, Desiree, HJT349-3
gigG-LecRK6-1	GIG362-6 gDNA	Artificial chimeric	PCR8, PK7WG2	No cell death when co-express with SCR74-B3b	*N. benthamiana*, Bintje, Desiree, HJT349-3
gigG-LecRK8-1	GIG362-6 gDNA	Artificial chimeric	PCR8, PK7WG2	No cell death when co-express with SCR74-B3b	*N. benthamiana*, Bintje, Desiree, HJT349-3
gigG-LecRK1	GIG362-6 BAC03-H3 (Haplotype 1)	Sequence confirmed	PCR8, PK7WG2	No cell death when co-express with SCR74-B3b	*N. benthamiana*, Bintje, Desiree, HJT349-3, PTA767-1, PTA767-8
gigG-LecRK2	GIG362-6 BAC03-H3 (Haplotype 1)	Sequence confirmed	PCR8, PK7WG2	No cell death when co-express with SCR74-B3b	*N. benthamiana*, Bintje, Desiree, HJT349-3, PTA767-1, PTA767-8
gigGlecRK-2B-4	GIG362-6 BAC03-2B (Haplotype 2)	Sequence confirmed	PCR8, PK7WG2		*N. benthamiana*, Bintje, Desiree, HJT349-3, PTA767-1, PTA767-8
gigGlecRK-2B-6-5	GIG362-6 BAC03-2B (Haplotype 2)	Artificial chimeric	PCR8, PK7WG2		*N. benthamiana*, Bintje, Desiree, HJT349-3, PTA767-1, PTA767-8
gigGlecRK-2B-6-7	GIG362-6 BAC03-2B (Haplotype 2)	Artificial chimeric	PCR8, PK7WG2		*N. benthamiana*, Bintje, Desiree, HJT349-3, PTA767-1, PTA767-8
gigGlecRK-2B-6-8	GIG362-6 BAC03-2B (Haplotype 2)	Artificial chimeric	PCR8, PK7WG2		*N. benthamiana*, Bintje, Desiree, HJT349-3, PTA767-1, PTA767-8
gigG-LecRK2	GIG362-6 gDNA	With native promoter and terminator	pBINPASSA		*N. benthamiana*, Bintje, Desiree, HJT349-3, PTA767-1, PTA767-8
gigG-LecRK1	Cloned from pENTR clones	No stop codon, eGFP tag	pBIN-KS-GWY-35s-eGFP	No cell death when co-express with SCR74-B3b	*N. benthamiana*, Bintje, Desiree, HJT349-3, PTA767-1, PTA767-8
gigG-LecRK2	Cloned from pENTR clones	No stop codon, eGFP tag	pBIN-KS-GWY-35s-eGFP	No cell death when co-express with SCR74-B3b	*N. benthamiana*, Bintje, Desiree, HJT349-3, PTA767-1, PTA767-8
Gene ID	**Synthsized**	**Notes**	**Vector**	**ATTA results**	**Tested plants**
gigG-LecRK1	Mutate the Chi motif	Successfully synthesized	PK7WG2	No cell death when co-express with SCR74-B3b	Bintje, Desiree, HJT349-3, PTA767-1, PTA767-8
gigG-LecRK2	Mutate the Chi motif	Success synthesized, but fail to clone to PK7WG2	PUC57		
gigG-LecRK-2B-1	Mutate the Chi motif	Successfully synthesized	PK7WG2	No cell death when co-express with SCR74-B3b	Bintje, Desiree, HJT349-3, PTA767-1, PTA767-8
gigG-LecRK-2B-2	Mutate the Chi motif	Successfully synthesized	PK7WG2	No cell death when co-express with SCR74-B3b	Bintje, Desiree, HJT349-3, PTA767-1, PTA767-8
gigG-LecRK-2B-3	Mutate the Chi motif	Success synthesized, but can not get correct clone in PUC57	Failed		
gigG-LecRK-2B-4	Mutate the Chi motif	Success synthesized, but can not get correct clone in PUC57	Failed		
gigG-LecRK-2B-5	Mutate the Chi motif	Success synthesized, but can not get correct clone in PUC57	Failed		

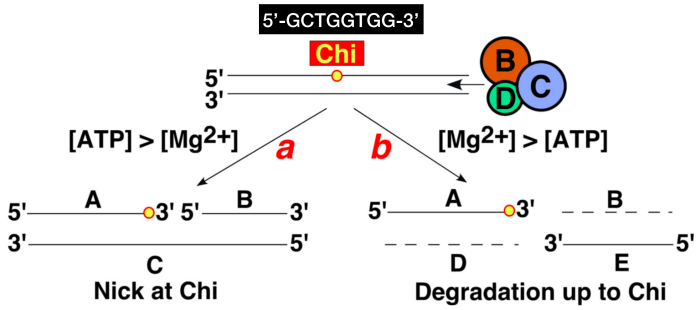


Figure 9. Chi site and RecBCD in *E. coli* (adapted from (Smith, 2012)

Alternative reaction (with excess Mg2+ or excess ATP) of purified RecBCD when the DNA is un-winded at the Chi site.

The "Chi" motif (5'-GCTGGTGG-3') is a recombination hotspot in *E.coli*, which can be recognized by the RecBCD helicase-exonuclease and cut in *E. coli* **(Figure 9)** (Smith, 2012). We found that all our candidate *G-LecRK* genes contain a Chi-motif, and speculated that this could be the nature of potential toxicity to the *E.coli* during the cloning.

Therefore, we decided to synthesize all the 6 *G-LecRK* genes from both haplotypes of GIG362-6, and generated a silent mutation at the Chi motif, from (5'-G**C**TGGTGG-3') to (5'-G**A**TGGTGG-3'), without changing the amino acid (Genewiz). Three *G-LecRK* genes: *syngigG-LecRK1*, *syngigG-LecRK2B-1* and *syngigG-LecRK2B-2* were successfully synthesized and cloned into pK7WG2, however, for *syngigG-LecRK2, syngigG-LecRK2B-3, syngigG-LecRK2B-4* and *syngigG-LecRK2B-5,* the cloning into the intermediate vector PUC57 or destination vector pK7WG2 failed. Based on the report from Genewiz, all of the *G-LecRK* genes are very difficult to clone, therefore, the presence of the Chi motif might not be the only reason to explain the technical cloning difficulties. To test if the three synthesized *G-LecRK* genes recognize SCR74-B3b, *Agrobacterium* transient transformation assay (ATTA) with the candidate genes and SCR74-B3b were performed on different plants which do not respond to SCR74-B3b **(Table 1)**. However, no cell death responses were found for any of these three candidate genes, as well as previously cloned *gig-GlecRK* genes in all tested plants (**Table 2** and **Figure 10**). In addition to infiltration of single G-LecRK, also various combinations of candidate genes were co-infiltrated with the effector, to assess whether heterodimerization of combined G-LecRK would lead to cell death (**Table 2** and **Figure 10**). However, also these experiments did not lead to a complementation phenotype. In summary, many attempts have been made to identify the candidate SCR74-B3b receptor but did not succeed. Other strategies are most probably

Combination	Candidate gene (PK7WG2, AGL1)	Candidate gene ID (PK7WG2, AGL1)	Candidate gene ID (PK7WG2, AGL1)	SCR74-B3b (PK7WG2/ pGR106)	Desiree	HJT349-3	PTA767-1	PTA767-8
1	syngigG-LecRK1			pK7WG2	-	-	-	-
2	syngigG-LecRK2B1			pK7WG2	-	-	-	-
3	syngigG-LecRK2B2			pK7WG2	-	-	-	-
4	gigG-LecRK2			pK7WG2	-	-	-	-
5	syngigG-LecRK1	syngigG-LecRK2B1	syngigG-LecRK2B2	pK7WG2	-	-	-	-
6	syngigG-LecRK1	gigG-LecRK2		pK7WG2	-	-	-	-
7	syngigG-LecRK1	syngigG-LecRK2B1		pK7WG2	-	-	-	-
8	syngigG-LecRK1	syngigG-LecRK2B2		pK7WG2	-	-	-	-
9	gigG-LecRK2	syngigG-LecRK2B1		pK7WG2	-	-	-	-
10	gigG-LecRK2	syngigG-LecRK2B2		pK7WG2	-	-	-	-
11	syngigG-LecRK2B1	syngigG-LecRK2B2		pK7WG2	-	-	-	-
12	syngigG-LecRK1	gigG-LecRK2	syngigG-LecRK2B1	pK7WG2	-	-	-	-
13	syngigG-LecRK1	gigG-LecRK2	syngigG-LecRK2B2	pK7WG2	-	-	-	-
14	gigG-LecRK2	syngigG-LecRK2B1	syngigG-LecRK2B2	pK7WG2	-	-	-	-
15	syngigG-LecRK1			pGR106	-	-	-	-
16	syngigG-LecRK2B1			pGR106	-	-	-	-
17	syngigG-LecRK2B2			pGR106	-	-	-	-
18	gigG-LecRK2			pGR106	-	-	-	-
19	syngigG-LecRK1	syngigG-LecRK2B1	syngigG-LecRK2B2	pGR106	-	-	-	-
20	syngigG-LecRK1	syngigG-LecRK2B1	syngigG-LecRK2B2	pGR106	-	-	-	-
21	syngigG-LecRK1	syngigG-LecRK2B1		pGR106	-	-	-	-
22	syngigG-LecRK1	syngigG-LecRK2B2		pGR106	-	-	-	-
23	gigG-LecRK2	syngigG-LecRK2B1		pGR106	-	-	-	-
24	gigG-LecRK2	syngigG-LecRK2B2		pGR106	-	-	-	-
25	syngigG-LecRK2B1	syngigG-LecRK2B2		pGR106	-	-	-	-
26	syngigG-LecRK1	gigG-LecRK2	syngigG-LecRK2B1	pGR106	-	-	-	-
27	syngigG-LecRK1	gigG-LecRK2	syngigG-LecRK2B2	pGR106	-	-	-	-
28	gigG-LecRK2	syngigG-LecRK2B1	syngigG-LecRK2B2	pGR106	-	-	-	-
+	R3b	Avr3b			+	+	+	+
-	pK7WG2-EV				-	-	-	-

Table 2. Overview of co-expression of the candidate G-LecRKs and SCR74-B3b in plants

The OD used in this study is 0.5 for the agroinfiltration. Data were scored 5 days after infection, "+" indicates strong cell death response, "-" indicates no visible cell death response. The number of combinations are related to **Figure 9**. Desiree: potato cultivar; HJT: *Solanum hjertingii* ; PTA: *Solanum papita*.

required to successfully clone and functional characterize the *G-LecRK* genes, to clearly prove their function as SCR74 receptor. Simultaneously, more efforts should be put in analyzing the other candidate genes in the mapping interval as well because they cannot be ruled out as candidates yet.

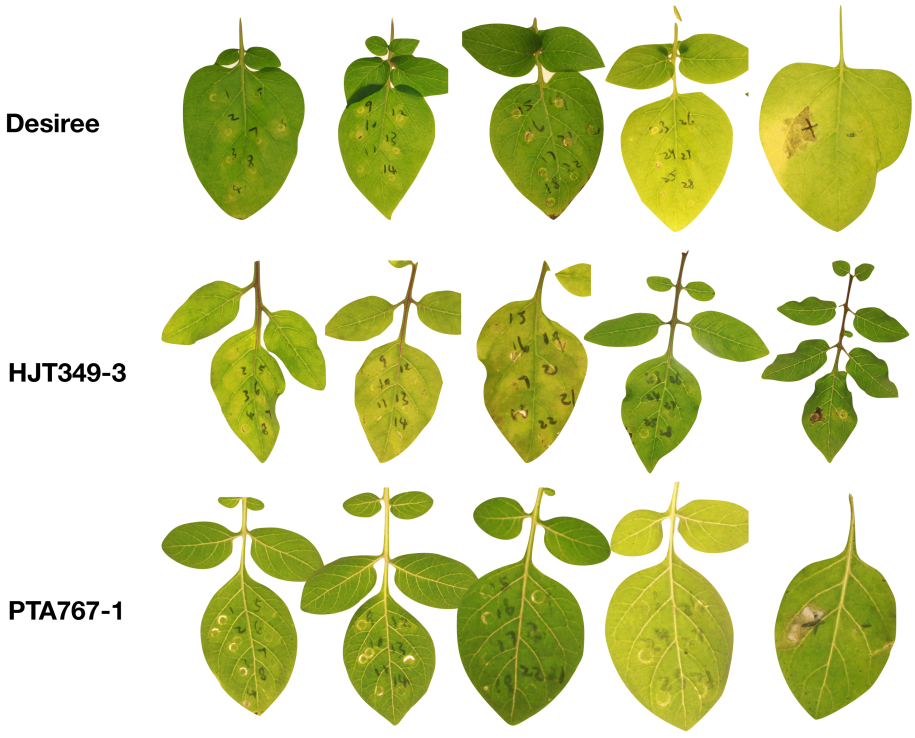


Figure 10. Agro-infiltrated leaves with co-expressed candidate genes and SCR74-B3b.

The agroinfiltration results in leaves of Desiree, HJT349-3 and PTA767-1, photographed at 4 days after infiltration. The combination of candidate genes and effectors are marked on the leaves by numbers and are listed in **(Table 2)**. R3b and Avr3b are used as positive control, empty pK7WG2 vector is used as negative control.

Discussion

In this study, we aimed to identify the receptor of SCR74-B3b of *P. infestans* in a responsive *Solanum* plant, GIG362-6. We fine-mapped the response to SCR74-B3b to a 73kb interval based on the DM reference genome, containing only 8 predicted genes including three *G-LecRK* genes. To construct the physical map of the SCR74 responding genotype GIG362-6, three BAC clones covering the mapping interval from two haplotypes of this diploid plant were isolated and sequenced. There we found 2 and 5 *G-LecRK* genes from the two haplotypes, respectively. RNAseq data suggested that the 2 *G-LecRK* genes in haplotype 1 are up-regulated after *P. infestans* infection, but not any of the other genes in the interval. Genomic analysis indicated that the *G-LecRK* locus is very diverse in different species, which might be caused by copy number variation, duplication or intergenic recombination. This observation fits well with our hypothesis, that SCR74 and the corresponding receptor are coevolving . For complementation, we cloned and synthesized the candidate *G-LecRK* genes into the expression vector pK7WG2. To test which gene is responsible for the SCR74 recognition, we co-infiltrated the candidate genes and SCR74-B3b in different plants, however, we did not identify a combination that leads to cell death.

Hypothesis: the "missing link" between plant immunity and self-incompatibility system

It has been proposed that, the plant self-incompatibility (SI) system shares similarities with the plant-microbial interaction system (Hodgkin *et al.*, 1988; Dickinson, 1994; Elleman and Dickinson, 1999). SI can be found in about 40% of flowering plant species, and has independently evolved in different plant families (Franklin-Tong, 2008; Igic *et al.*, 2008). In many plant species, self-incompatibility is controlled by multiple *S*-haplotypes of a single *S*-locus. For example, in Solanaceae, SI is under gametophytic control (GSI), by S-RNase (Kao and Tsukamoto, 2004). However, in Brassicaceae the SI system is determined by sporophytic SI (SSI), the stigma membrane localized G-LecRK (SRK) receptor perceives a pollen ligand called SCR/ SP11 (Takasaki *et al.*, 2000). Surprisingly, similar to the apoplastic small cysteine-rich protein SCR74 in *P. infestans*, the SCR/ SP11 is also a small-cysteine rich protein. The *SCR/ SP11* alleles are highly polymorphic except the eight cysteine residues. In *Brassica*, the different SCR/ SP11 variants share only 35% protein identity (Schopfer *et al.*, 1999).

In our study, the candidate SCR74 receptor was mapped to a *G-LecRK* locus. It is tempting to speculate that there is similarity between the SRK-SCR/SP11 interaction and the candidate gigG-LecRK-SCR74 interaction. For both systems: 1) they include the invading of the 'host' cell by a tubular cell, i.e. the plant pollen tube or *Phytophthora* hyphae, germinating from a spore-like structure, i.e. pollen or *Phytophthora* cyst; 2) The outcome of the incompatible response excludes undesirable cells or organisms i.e. self-pollen or pathogens; 3) Both interactions are driven by highly diverse receptors and ligands i.e. SRK-SCR/ SP11 and candidate G-LecRK-SCR74) **(Figure 11)**.

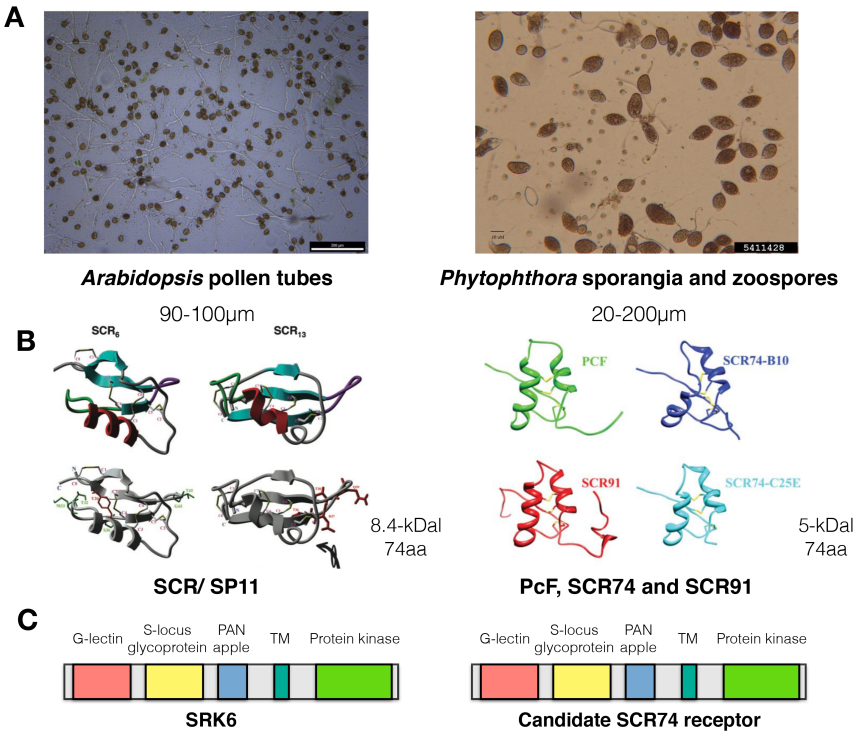


Figure 11. Similarity between pollen and *Phytophthora* zoospores, SCR/ SP11 and SCR74.

A. The morphology of plant pollen is similar to *Phytophthora* sporangia and zoospores. Photos from (Dumont *et al.*, 2015) and Elizabeth Bush.
B. Both SCR/ SP11 and SCR74 are small cysteine-rich proteins, and they share a similar size and molecular weight. Pictures are from (Chookajorn *et al.*, 2004; Orsomando *et al.*, 2011)
C. The pollen SCR/ SP11 receptor is G-LecRK protein (SRK6), same as the candidate SCR74 receptor. They share the same domain architecture.

The *SRK* gene dependent SI system has exclusively evolved in Brassicaceae. From an evolutionary point of view, we speculate that this SRK dependent SI system has evolved from the plant immunity system, because the G-LecRK dependent immunity was present already in monocots, like the *PiD2* and *OsLecRK1-3* genes in rice (Chen *et al.*, 2006; Liu *et al.*, 2015).

If the function of the SCR74 receptor can be confirmed, it would provide evidence that plants deploy similar receptors to recognize self- or non-self- molecules. For more discussion on this topic, we refer to the general discussion of this PhD thesis (**Chapter 6**).

Difficulties to clone the candidate G-LecRK genes

Many technical challenge have been faced during cloning of the candidate *G-LecRK* genes. Firstly, the 7 *gigG-LecRKs* are highly similar. The primers designed for the whole ORF might anneal to several *gigG-LecRK* homologs, therefore, lots of chimeric genes were obtained. We noticed this issue when we had the BAC sequences. When the flanking sequences are known, this issue can be solved by two-steps PCR, firstly amplify the genes with flanking sequences by specific primers, then perform another PCR for the full length ORF.

In addition, the *G-LecRK* genes are extremely difficult to clone in the *E.coli* system. In molecular cloning, it is known that some genes are difficult to clone, and then it is believed that those genes/proteins are toxic to *E. coli* (Saida *et al.*, 2006). Various tricks can deal with this dilemma. For example, using commercially mutated *E. coli* strains, growing *E. coli* at lower temperature, or introducing artificial introns into the gene, etc. In our case, all the *gigG-LecRK* genes contain a Chi motif (5'-GCTGGTGG-3'), which can be recognized by the RecBCD helicase-exonuclease and be cut in *E. coli* **(Figure 9)** (Smith, 2012). Therefore, Chi sites are recombination hot spots in *E. coli*. In our cloning, we also found that the candidate genes tended to be cut in *E. coli*. To deal with this problem, we recovered the competent cells at lower temperature (28°) after transformation, and raised the transformed *E. coli* at room temperature. Using these methods, we could obtain a few clones with the full-length *G-LecRK* genes.

In the end, we decided to synthesize these candidate *G-LecRK* genes and generated a nonsynonymous mutation at the chi site. However, based on the report of the company which synthesized the genes, the mutated *G-LecRK* genes are still difficult to clone into expression vectors, even by a specialized company. The constructs with the *G-LecRK* genes grow very slow and many deletions have been observed in the constructs. This indicates that the Chi

motif is likely not the only reason caused these issues. Surprisingly, other *G-LecRK* genes faced similar issues during cloning, for example, LPS receptor LORE from *Arabidopsis* is also a G-LecRK, it was solved when introduced an artificial intron into the gene (personal communication with Dr. Stefanie Ranf, Technical University of Munich)

Another challenge was functional complementation testing of the candidate genes. The source of the SCR74 receptor *S. microdontum* is not suitable for transient gene expression by agro-infiltration or stable transformation. Therefore we had to test the candidates genes in other plants, such as *N. benthamiana*, or potato genotypes that do not respond to SCR74-B3b. However, if the SCR74-B3b recognition still requires an unknown component from the *S. microdontum* background, the expected cell-death phenotype will not be recovered. For tomato *I-3*, the same problem occurred, the *I-3* transformed susceptible tomato confers resistance to *Fusarium oxysporum f. sp. lycopersici* (Fol) carrying *Avr3*, however, no HR-like cell death could be observed when co-expressed of I-3 and AVR3 in *N. benthamiana* or tomato. Furthermore, Co-immunoprecipitation (Co-IP) experiment did not confirm the direct protein interaction between the I-3 and AVR3 (Catanzariti *et al.*, 2015).

Other candidate genes?

In this study, we aimed to identify the SCR74 receptor. Since SCR74 is a fast-evolving effector under diversifying selection we hypothesized that the candidate receptor would be a fast evolving surface immune receptor as well. The candidate genes of the RLP/ RLK type fit these criteria and therefore we firstly focused on the *G-LecRK* genes. However, the cell death phenotype might also be caused by other types of genes, for example, target of the effectors, or other downstream components. Therefore, other candidate genes in the mapping interval should also be tested by transient or stable transformation assay, and cloning of these genes is ongoing.

In summary, in this study, we genetically fine-mapped the putative SCR74 receptor to a *G-LecRK* locus. The physical map from the wild *Solanum* was generated by sequencing the BAC clone spanning the mapping interval. We found the *G-LecRK* locus are variable in different *Solanum* species, especially in wild potatoes. We found an intriguing similarity between the plant SI system and the immunity system, and we discuss that the *G-LecRK* might be a common receptor to recognize self- and non-self- components. However, The candidate G-*LecRK* genes did not lead to complementation so far.

Materials and Methods

Plant materials and genotyping

Population 7026 was generated by crossing the SCR74-B3b responsive genotype GIG362-6 with the non-responsive genotype MCD360-1 **(Chapter 2)**. The seeds of population 7026 were sown in MS20 medium. A fast genotyping pipeline using *in vitro* plants was developed **(Figure 12)**.

For DNA isolation, 20 µL 0.5 M NaOH was added per well, the caps were closed and put in the TissueLyser II (QIAGEN), grinded for 5 minutes at max speed. 20 µL of 100 mM Tris was added; After spin down, 1 PCR plate was filled with 200 µL 100 mM Tris; 5 µL DNA was added to the PCR plate. Then 1 µl samples was used for genotyping.

The HRM markers were tested on the parents and the F1 progenies using the following protocol for a 10 µL reaction: (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® Plus+ (BioFire), 2 µL 5x Phire Buffer, 0.06 µL Phire taq, 4.44 µL MQ water). 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 was: 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 40s. The LightScanner® System (Biofire) was used for measuring and analyze the melting curve.

For the SSR markers, a 20 µL reaction was used for the PCR reaction (5 µL DNA (=10 ng), 2 µL Gold star buffer, 1.2 µL 25 mM $MgCl_2$, 0.4 µL 25 mM dNTP, 4 µL F+R primers (1 pmol/ µL each), 0.06 µL goldstar polymerase (EUROGENTEC SA, Seraing, Belgium), 7.34 µL MQ. The PCR cycling was : 94 ˚C for 2 min, (94 ˚C for 30 s, 56 ˚C for 30 s, 72 ˚C for 45 s) with 35 cycles, then 72 ˚C for 2 min. The PCR products were analyzed by 4300 DNA analyzer (LI-COR, Lincoln, USA).

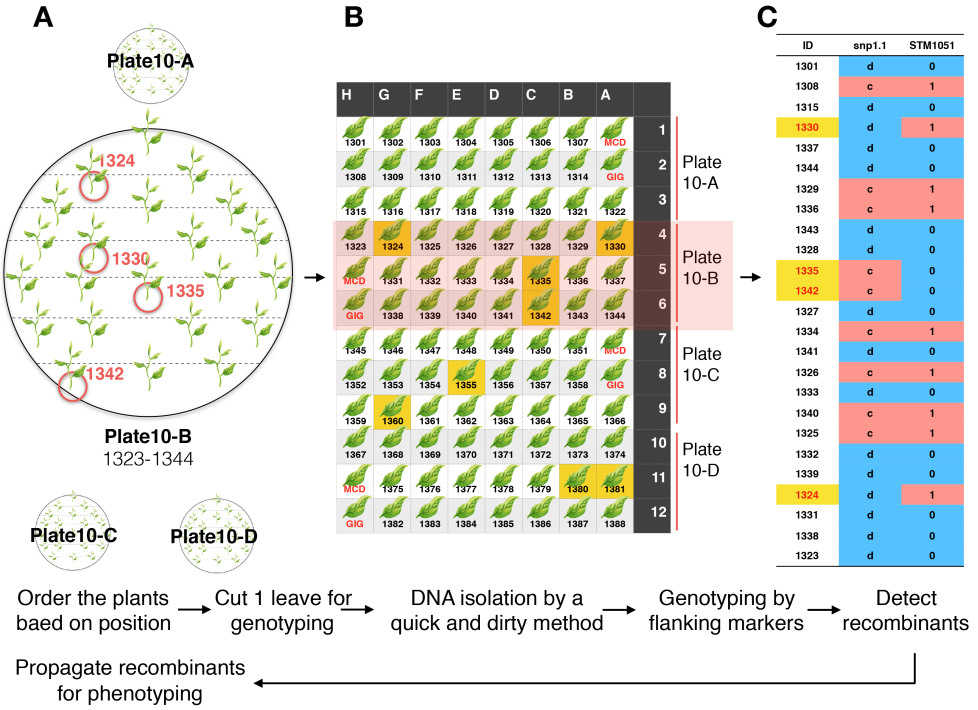


Figure 12: The pipeline of high-throughput genotyping of *in vitro* plants.

A. The seedlings were transferred into round pots with MS20 medium. There were 22 plants in each pot, and the pots were named: PlateX-A, B C or D, corresponding with the code in the 96 plates. The recombinants were transferred in to new pots for *in vitro* maintenance and phenotyping.

B. One leaf from each offspring plantlet was cut and collected into a deep well plate with 96 wells. A "quick and dirty" DNA isolation was performed.

C.The flanking markers were tested after DNA isolation, and the recombinant plants were selected, propagated *in-vitro* and phenotyped.

PVX agro-infection

The recombinant plants from population 7026 were maintained *in vitro* and phenotyped at least 3 times in the greenhouse. We used the same constructs (SCR74-B3b-pGR106) for phenotyping as described in **Chapter 2**; CRN2 and empty pGR106 were used as positive and negative controls. The cell death response was scored 14 days after the agro infection. The details were described in **Chapter 2**.

Markers used for fine mapping of the SCR74-B3b receptor

Markers used for fine mapping the SCR74-B3b responsive receptor are listed in **Table 3**.

Table 3. Markers used in this study

Markers	Type	F	R
STM1051	SSR	TCCCCTTGGCA TTTTCTTCTCC	TTTAGGGTGGGGTGAGGTTGG
RenSeq1.1	SNPs	ATTCAGTATCCACAATGTACG	ACGTTAAGCTTTTCAACTTAC
PRP1.1	SNPs	AGCTTTCTCTTGTTCCTCT	ACTCTAACCAGCTATCGAA
exon19	SNPs	CCAGCATTGCAGCAGAGAAG	GGTGTCCAACAGCCAGATCA
exon26	SNPs	ACCCTACCACCACCACTTCA	AAGTGGCTGCTGAAGGGTTT
exon3	SNPs	TCTCTCCCAACGAGGTCCTT	AGATTCACAGTCCTGCCGTC
S55-snp1	SNPs	GAAGAATTTTGGGCAGAGAT	CTTCAGAACAAAATCCCCACAT
S50-snp1	SNPs	CAAATACAGGCGGAGTTA	TTACTCTTCATTCTCACCTT
LecRK1-snp1	SNPs	GTGTTCGCGATAAGGTCTGAT	CAAAGACCCCTGAAGTCCT
S105-snp2	SNPs	CTAGTGCCACGTGATACTA	GCAGAAACGAAGAAACACT
S103-snp2	SNPs	GTAGTTGTTTGTCTATGGTCCTT	CTAGTAGTACCCCACAGAAAA
S47-snp1	SNPs	TCGGGTAGTCAGATCAAA	CTGGACTAATATGATGGGAA
1.211-snp1	SNPs	GTGAGTACTTTATGACATAGCTTTGT	TCAGAAAAGACTGCAACATGC
1.227-snp4	SNPs	AGTCGACAAATATTCTGGGATTTCC	TCCCCACCCTACACACTCTT
1.227-snp5	SNPs	GTACGCCTTCACTTCATTCCT	CCAGAAGATGGTCTCAGTGGA
1.273-snp5	SNPs	CCTGCTGACCATGTCCAACA	TACACAGTGGAGCGGTTGAC
S112	SNPs	CAGGCTCTATATGTTTGGCAA	CATCTTGATGCTAGTAACATT
S114-snp1	SNPs	GTGTGAGCTGCAAAGAGA	GCCAAATACTAAGCTCCTGTT
S108-snp1	SNP2	GGTGGTAAACGTGTTCTAT	CGCATTTCAATCCACAATT

BAC library of the GIG362-6

The BAC library of GIG362-6 was generated by Bio S&T company (Canada). The BAC clones are stored in three 96 deep well plates, with 490 clones in each well, the total number of clones are 141, 120. The average insert size is 151 kb. The coverage is >11x (GIG362-6 is a highly heterozygous diploid potato, The genome size is around 2 x 900 Mb). The BAC construction vector is pCC1BAC (HindIII) and the *E. coli* strain was DH 10B. The BAC library is stored at -80 °C.

The BAC DNA was isolated by QIAprep miniprep (Qiagen) kit with minor revision. After adding the P3 step in the industrial instruction, the tubes were closed with caps and mixed by inverting, and spin down for 10 minutes, the supernatant was transferred to fresh tubes and 0.6 V isopropanol was added, then mixed and centrifuged for 10 min. The pellets were washed with 500 µL of 70% cold ethanol, centrifuged for 5 minutes. The pellets were air dried and re-suspended in 50 µL MQ.

For BAC clone isolation, the markers in the mapping interval were firstly tested on BAC DNA and positive pools with 490 clones were detected. Then the selected pools were diluted ≈1 x 10^{-6} – 1 x 10^{-8} times with sterilized water. The diluted cultures should contain around 1000 clones/mL. Then 1 mL culture was plated on 1 square LB plate with 25 mg/L chloromycetin and grown overnight at 37 °C. Next day, 1152-2304 single colonies were picked up into 384 plates with liquid LB (25 mg/L chloromycetin) and grown overnight at 37°C. Then the 384 plates were pooled into 96 well plates for PCR. If a positive well was found in the 96 well plates, another PCR was performed to obtain single BAC clones.

The single BAC clone were further plated on solid LB plates, 5 single colonies were tested by PCR. The BAC clones were sequenced by French Plant Genomic Resource Center (INRA-CNRGV) using PacBio technology. The BAC clones were firstly characterized by estimating their insert size and sequencing the BAC ends by Sanger sequencing (BES). Finally, PacBio RS libraries with 8-12 kb insert size (based on a minimum of 60 X with average reads length up to 5 kb) were produced and sequenced.

Genomic and phylogenetic analysis

Potato reference genome DM (v4.03), *C. annuum* genome (v2.0), *S. melonggena* Line '67/3', *S. lycopersicum* (v2.4), *S. verrucosum* (personal communication with Dr. Glenn Bryan, James Hutton Institute, Dundee) and *S. chacoense* M6 (v4.1) genomes were used in this study, the genome was visualized by Geneious R10 (Kearse *et al.*, 2012) (http://www.geneious.com). The BAC sequences were imported into Geneious R10, and the annotation was made by Augustus (v2.5.5) (Stanke and Morgenstern, 2005). Blastn and Blastx were used for further gene annotation. InterPro (https://www.ebi.ac.uk/interpro/interproscan.html) was used for predicting the protein domains. MAFFT (v7.309) was used for sequence alignment (Katoh and Standley, 2013).

The G-LecRK proteins from *Arabidopsis* and tomato were used as query to predict G-LecRKs in potato (Teixeira *et al.*, 2018), all the G-LecRKs were blasted against predicted potato proteins, the hits were annotated by InterProscan and visually checked in Geneious R10. The G-LecRK proteins were aligned by MAFFT (v7.309), and a Neighbor-Joining (NJ) tree was built by Geneious Tree Builder, with Jukes-Cantor model, 1000 replicates of bootstrap were used for resampling.

RNAseq of the mapping parents GIG362-6 and MCD360-1

The mapping parents GIG362-6 and MCD360-1 were treated by water control and a zoospore suspension of *P. infestans* isolate UK3928A. After 48 hours, the leaf disks were sampled into 2 mL RNase free tubes and frozen by liquid nitrogen. The 4 samples were grinded by TissueLyser II (QIAGEN). Then 100mg samples were used for RNA isolation by RNeasy Plus Mini Kit (QIAGEN) following industrial instructions. The gDNA eliminator spin column from the kit can efficiently remove the gDNA. The 4 RNA samples were tested by agarose electrophoresis and quantified by Nanodrop (ThermoFisher). The 4 RNA samples were send to the Genome Analysis Centre (TGAC, UK) with dry ice for RNA sequencing. The samples were sequenced by HiSeq2000 with a 100bp paired end read metric.

The raw reads were first checked with FastQC, the adapters were trimmed with trimmomatic (v0.36) (Bolger *et al.*, 2014). The trimmed reads were then mapped to the BAC sequences using STAR (v2.5) (Dobin *et al.*, 2013). The output BAM files were imported to Geneious R10, the RPKM (Reads Per Kilobase Million), FPKM (Fragments Per Kilobase Million) and TPM (Transcripts Per Kilobase Million) were calculated in Geneious R10. The differential expression levels were also calculated in Geneious R10, the transcripts were compared and normalized by median of gene expression ratios.

Candidate gene cloning

The *G-LecRK* genes were amplified from the isolated BAC clone with Phusion High-fidelity DNA polymerases (ThermoFisher), and A-tailing by DreadTaq DNA polymerases (ThermoFisher), then the PCR products were ligated to gateway entry vector PCR8, using the pCR8/GW/TOPO TA cloning kit (ThermoFisher). Then the genes were shuffled to pK7WG2 vector **(Figure 13)** for transient expression in plants. DH5α competent cells were used for the cloning

The 7 *gigGlecRKs* were synthesized by GeneWiz (UK), the chi site (5'-G**C**TGGTGG-3') was replaced by (5'-G**A**TGGTGG-3'), which does not change any amino-acid on the proteins.

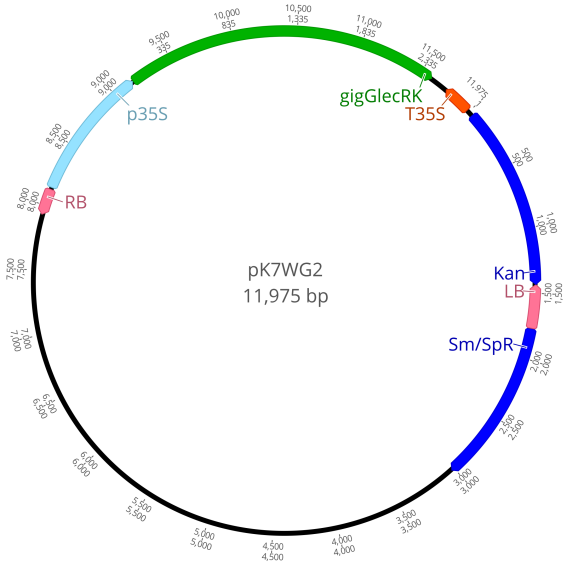


Figure 13: Vector map of pK7WG2 vector

Agro co-infiltration in plants

The protocol of agro co-infiltration were described previously (Du *et al.*, 2014). In brief, we mixed the candidate genes and SCR74-B3b, and co-infiltrated them on plant leaves, we used R3a and Avr3a as positive control and empty pK7WG2 vector as negative control. The phenotype was measured 3-5 days after infiltration.

Acknowledgements

This work was supported by a NWO-VIDI grant 12378 (X.L., D.W., E.D., V.G.A.A.V.), China Scholarship Council (CSC) (X.L). We thank Dr. ir. Klaas Bouwmeester for the inspiring discussions. We thank Dr. Glenn Bryan (James Hutton Institute) for sharing the genome of *S. verrucosum*, Dr. Stefanie Ranf (Technical University of Munich) for discussion. We thank Isolde Bertram-Pereira for culturing *Solanum* plants, Henk Smid and Harm Wiegersma for helping in the greenhouse. We thank Dr. Helene Berges and Caroline Callot from the French Plant Genomic Resource Center (INRA-CNRGV) for their help to sequence the BAC clone.

References

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30**: 2114–2120

Boutrot F, Zipfel C (2017) Function, Discovery, and Exploitation of Plant Pattern Recognition Receptors for Broad-Spectrum Disease Resistance. Annu Rev Phytopathol **55**: 257–286

Catanzariti AM, Lim GTT, Jones DA (2015) The tomato I-3 gene: a novel gene for resistance to Fusarium wilt disease. New Phytol **207**: 106–118

Chen X, Shang J, Chen D, Lei C, Zou Y, Zhai W, Liu G, Xu J, Ling Z, Cao G, et al (2006) A B-lectin receptor kinase gene conferring rice blast resistance. The Plant Journal **46**: 794–804

Chookajorn T, Kachroo A, Ripoll DR, Clark AG, Nasrallah JB (2004) Specificity determinants and diversification of the *Brassica* self-incompatibility pollen ligand. Proceedings of the National Academy of Sciences **101**: 911–917

Dickinson H (1994) Simply a social disease? Nature **367**: 517–518

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics **29**: 15–21

Du J, Rietman H, Vleeshouwers VGAA (2014) Agroinfiltration and PVX agroinfection in potato and *Nicotiana benthamiana*. JoVE e50971–e50971

Du J, Verzaux E, Chaparro-Garcia A, Bijsterbosch G, Keizer LCP, Zhou J, Liebrand TWH, Xie C, Govers F, Robatzek S, et al (2015) Elicitin recognition confers enhanced resistance to *Phytophthora infestans* in potato. NPLANTS 10.1038–nplants.2015.34

Dumont M, Cataye CL, Lehner A, Mar chal E, Lerouge P, Falconet D, Mollet J-C (2015) A Simple Protocol for the Immunolabelling of Arabidopsis Pollen Tube Membranes and Cell Wall Polymers. BIO-PROTOCOL. doi: 10.21769/BioProtoc.1502

Elleman CJ, Dickinson HG Commonalities between pollen/stigma and host/pathogen interactions: calcium accumulation during stigmatic penetration by *Brassica oleracea* pollen tubes. Sex Plant Reprod **12**: 194–202

Franklin-Tong N (2015) Self-incompatibility: Calcium signalling in *Brassica*. NPLANTS **1**: 15129

Franklin-Tong VE (2008) Self-incompatibility in flowering plants: Evolution, Diversity, and Mechanisms. Springer

Hodgkin T, Lyon GD, Dickinson HG (1988) Recognition in flowering plants: A comparison of the Brassica self-incompatibility system and plant pathogen interactions. New Phytol **110**: 557–569

Igic B, Lande R, Kohn JR (2008) Loss of self-incompatibility and its evolutionary consequences. Int J Plant Sci **169**: 93–104

Kao T-H, Tsukamoto T (2004) The Molecular and Genetic Bases of S-RNase-Based Self-Incompatibility. The Plant Cell **16**: S72–S83

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol **30**: 772–780

Kawchuk LM, Hachey J, Lynch DR, Kulcsar F, van Rooijen G, Waterer DR, Robertson A, Kokko E, Byers R, Howard RJ, et al (2001) Tomato *Ve* disease resistance genes encode cell surface-like receptors. Proceedings of the National Academy of Sciences **98**: 6511–6515

Lacombe S, Rougon-Cardoso A, Sherwood E, Peeters N, Dahlbeck D, van Esse HP, Smoker M, Rallapalli G, Thomma BPHJ, Staskawicz B, et al (2010) Interfamily transfer of a plant pattern-recognition receptor confers broad-spectrum bacterial resistance. Nature Biotechnology **28**: 365

Lannoo N, van Damme EJM (2014) Lectin domains at the frontiers of plant defense. Frontiers in Plant Science **5**: 397

Leisner CP, Hamilton JP, Crisovan E, Carpintero NCM, 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

Lis H, Sharon N (1998) Lectins: Carbohydrate-Specific Proteins That Mediate Cellular Recognition †. Chem Rev **98**: 637–674

Liu Y, Wu H, Chen H, Liu Y, He J, Kang H, Sun Z, Pan G, Wang Q, Hu J, et al (2015) A gene cluster encoding lectin receptor kinases confers broad-spectrum and durable insect resistance in rice. Nature Biotechnology **33**: 301–216

Liu Z, 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 *scr74* gene family of *Phytophthora infestans*. Mol Biol Evol **22**: 659–672

Orsomando G, Brunetti L, Pucci K, Ruggeri B, Ruggieri S (2011) Comparative structural and functional characterization of putative protein effectors belonging to the PcF toxin family from *Phytophthora* spp. Protein Sci **20**: 2047–2059

Pérez-Pérez JM, Esteve-Bruna D, González-Bayón R, Kangasjärvi S, Caldana C, Hannah MA, Willmitzer L, Ponce MR, Micol JL (2013) Functional Redundancy and Divergence within the Arabidopsis RETICULATA-RELATED Gene Family. PLANT PHYSIOLOGY **162**: 589–603

Ranf S, Gisch N, Schäffer M, Illig T, Westphal L, Knirel YA, Sánchez-Carballo PM, Zähringer U, Hückelhoven R, Lee J, et al (2015) A lectin S-domain receptor kinase mediates lipopolysaccharide sensing in *Arabidopsis thaliana*. Nat Immunol **16**: 426–433

Saida F, Uzan M, and BOCP (2006) Expression of highly toxic genes in E. coli: special strategies and genetic tools. Current Protein and Peptide Science, 2006, 7, 47-56

Schopfer CR, Nasrallah ME, Nasrallah JB (1999) The male determinant of self-incompatibility in *Brassica*. Science **286**: 1697–1700

Smith GR (2012) How RecBCD Enzyme and Chi Promote DNA Break Repair and Recombination: a Molecular Biologist's View. Microbiol Mol Biol Rev **76**: 217–228

Stanke M, Morgenstern B (2005) AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Research **33**: W465–W467

Stein JC, Dixit R, Nasrallah ME, Nasrallah JB (1996) SRK, the stigma-specific S locus receptor kinase of Brassica, is targeted to the plasma membrane in transgenic tobacco. THE PLANT CELL ONLINE **8**: 429–445

Stein JC, Howlett B, Boyes DC, Nasrallah ME, Nasrallah JB (1991) Molecular cloning of a putative receptor protein kinase gene encoded at the self-incompatibility locus of Brassica oleracea. Proceedings of the National Academy of Sciences **88**: 8816–8820

Takasaki T, Hatakeyama K, Suzuki G, Watanabe M, Isogai A, Hinata K (2000) The S receptor kinase determines self-incompatibility in *Brassica* stigma. Nature **403**: 913–916

Takayama S, Shimosato H, Shiba H, Funato M, Che F-S, Watanabe M, Iwano M, Isogai A (2001) Direct ligand-receptor complex interaction controls *Brassica* self-incompatibility. Nature **413**: 534–538

Teixeira MA, Rajewski A, He J, Castaneda OG, Litt A, Kaloshian I (2018) Classification and phylogenetic analyses of the Arabidopsis and tomato G-type lectin receptor kinases. BMC Genomics **19**: 239

Thomas CM, Jones DA, Parniske M, Harrison K, Balint-Kurti PJ, Hatzixanthis K, Jones JD (1997) Characterization of the tomato Cf-4 gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in Cf-4 and Cf-9. The Plant Cell **9**: 2209–2224

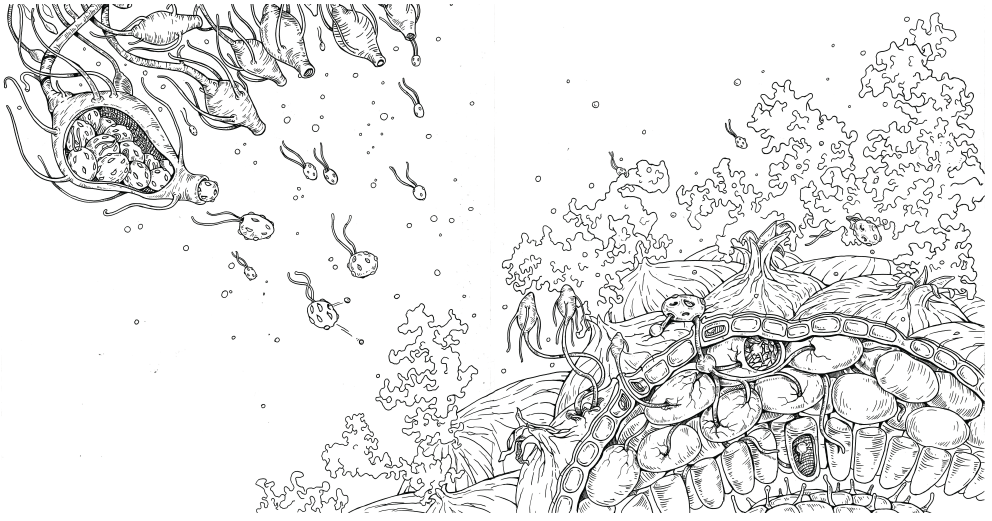
Wang Y, Bouwmeester K (2017) L-type lectin receptor kinases: New forces in plant immunity. PLoS Pathog **13**: e1006433

Win J, Chaparro-Garcia A, Belhaj K, Saunders DGO, Yoshida K, Dong S, Schornack S, Zipfel C, Robatzek S, Hogenhout SA, et al (2013) Effector biology of Plant-associated organisms: concepts and perspectives. Cold Spring Harbor Symposia on Quantitative Biology **77**: 235–247

Chapter 5

Phytophthora MAMPs Pep-13/25 recognition in Solanaceae species and fine mapping of the Pep-25 receptor by BSA-RNAseq

Xiao Lin, Happyka Fillianti, Laura de Rond, Yerisf Carla Torres Ascurra, Emmanouil Domazakis, Carolina Aguilera-Galvez, Afewerki Yohannes Kiros, Evert Jacobsen, Richard G. F. Visser, Thorsten Nürnberger, Vivianne G. A. A. Vleeshouwers



Abstract

MAMPs/PAMPs-triggered immunity (MTI/PTI) is the first layer of the plant immune system and is considered to confer durable disease resistance. Pep-13 and Pep-25 are peptides of 13 and 25 amino acids, respectively, from the oomycete transglutaminase (TGase) GP42. Pep-13/25 are recognized as typical MAMPs/PAMPs of oomycetes and they trigger defense responses in parsley and cell death in potato. However, the receptor of Pep-13/25 remains unknown. Here we screened the Pep-13/25 in various Solanaceae for their ability to mount cell death response. The recognition spectrum was restricted to some members of the *Solanum* section Petota. To understand the genetic basis and map the Pep-25 receptor, a mapping population was generated and a bulked segregant analysis-RNAseq (BSA-RNAseq) approach was applied. The candidate gene was mapped on chromosome 3. Fine mapping resulted in an interval of about 330 kb (0.42 cM) based on the potato reference genome. By isolating and sequencing one BAC clone from the Pep-25 responsive genotype GIG362-6, we found that the Pep-25 receptor is located in a cluster of receptor-like kinases (*RLK*) genes. Together with expression data from BSA-RNAseq, we found 3 *RLK* genes up-regulated after *P. infestans* infection, and these are our top candidate genes. Our results show that BSA-RNAseq is a powerful tool to accelerate map-based cloning. Our finding will lead to the cloning of a novel surface immune receptor in potato which will ultimately contribute to understanding plant durable resistance.

Introduction

Microbial/pathogen-associated molecular patterns (MAMPs/PAMPs) are thought to be widespread, conserved, abundant and essential in pathogens (Pel and Pieterse, 2013). Based on the "zig-zag" model of the plant immune system (Jones and Dangl, 2006), these MAMPs/PAMPs can be recognized by plant pattern recognition receptors (PRRs), which form the first layer of the plant defense system. Many MAMP/PAMP molecules have been identified, like flg22 (flagellin), elf18 (EF-tu), PGN (peptidoglycan), csp22 (Cold shock protein, CSP) and lipopolysaccharides (LPS) from bacteria (Felix *et al.*, 1999; Dow *et al.*, 2003; Felix and Boller, 2003; Kunze, 2004; Gust *et al.*, 2007), and chitin, NLP and EIX (ethylene-inducing xylanase) from fungi (Pearce and Ride, 1982; Bailey *et al.*, 1990; Gijzen and Nürnberger, 2006). Oomycetes are a major class of plant pathogens, including *Phytophthora infestans* that causes the notorious potato late blight disease. In oomycetes, elicitins (e.g. INF1) are the best studied MAMPs/PAMPs (Derevnina *et al.*, 2016). Other oomycete MAMPs/PAMPs include glycoside hydrolase (XEG1), cellulose-binding elicitor lectin (CBEL), GP42 (Pep-13), and the ethylene-inducing peptide 1-like protein (NLP, nlp20) (Séjalon *et al.*, 1995; Brunner *et al.*, 2002; Böhm *et al.*, 2014; Ma *et al.*, 2015).

To protect plants against oomycetes, nucleotide-binding domain and leucine-rich repeat containing (*NLR)* genes based resistance breeding has been practiced for many years (Haverkort *et al.*, 2016). However, due to the fast evolving nature of oomycetes, most *NLR* genes have been quickly defeated in the field. Finding more plant surface immune receptors and stacking them with *NLR* genes is thought to be a preferred way to achieve more durable resistance (Dangl *et al.*, 2013). Some surface immune receptors confer interfamily broad-spectrum resistance, for example, NLP receptor RLP23 from *Arabidopsis* enhances resistance to *P. infestans* in *S. tuberosum* (Albert *et al.*, 2015). In potato, the INF1 receptor ELR is the only MAMP/PAMP receptor which has been cloned so far (Du *et al.*, 2015). ELR was identified from a wild potato *Solanum microdontum* (Verzaux *et al.*, 2010), and when expressed in cultivated potato, ELR reduced the growth rate of *P. infestans* on detached leaves under laboratory conditions (Du *et al.*, 2015).

GP42 is a 42kDa protein that was originally isolated from culture filtrates of *Phytophthora megasperma* f. sp. *glycinea* and found to trigger plant defense responses in cell cultures of parsley (Parker *et al.*, 1991). To find the minimal elicitor peptide that has GP42 elicitor activity, several endo- and exopeptidases were used for digesting the protein. The obtained peptides were purified by reverse-phase high pressure liquid chromatography (HPLC), and a peptide of 13 amino acid residues (Pep-13) was found to be the smallest peptide with elicitor-

activity in parsley. Pep-25 is a longer version containing Pep-13, which showed the same kind of elicitor activity (Nürnberger *et al.*, 1994). The motif of Pep-13 is highly conserved and exclusively present in all tested *Phytophthora* species, including *P. infestans* (Figure 1A). In parsley, defense responses include ion fluxes (H^+/Ca^{2+} influxes, K^+/Cl^-effluxes), oxidative burst, defense-related gene activation and phytoalexin formation (Nürnberger *et al.*, 1994). Infiltration of Pep-13 and Pep-25 into potato (*S. tuberosum* cv. Désirée) leaves (Brunner *et al.*, 2002) leads to visible hypersensitive responses like cell death, as well as up-regulation of defense-related genes (Brunner *et al.*, 2002). Another study, after using RNA interference constructs targeting jasmonic acid (JA) biosynthesis genes in potato, demonstrated that Pep-13 could also induce accumulation of salicylic acid (SA) (Halim *et al.*, 2009). So, the Pep-13 response in potato requires both JA and SA pathways, induces hydrogen peroxide and expression of defense genes, including HR-like cell death response.

In addition to plant defense-eliciting activity, Pep-13 was found to have calcium-dependent cell wall transglutaminase (TGase) activity. TGases catalyze post-translational modification of proteins by crosslinking between glutamine and lysine residues, therefore, it is also called 'biological glue' (Martin *et al.*, 2002). TGases are ubiquitous enzymes present in micro-organisms, plants as well as animals. In oomycetes, TGases are involved in the cell wall biogenesis and might protect cell walls against plant proteases (Meijer *et al.*, 2006; Raaymakers and Van den Ackerveken, 2016). The TGase activity was found in all tested *Phytophthora* species, including *P. infestans* (Brunner *et al.*, 2002). GP42 does not have homologs of other known TGases from plants or animals, but a database search indicated that some TGase-like sequences from marine *Vibrio* bacteria did show similarity with GP42 from oomycetes. Probably, evolutionary lateral gene transfer between bacteria and oomycetes may have taken place (Reiss *et al.*, 2011).

Despite the fact that Pep-13/25 was found in 1994 (Nürnberger *et al.*, 1994), the receptor of Pep-13 and Pep-25 has remained unidentified. Previous studies in the parsley cell induction system showed that the high affinity binding site to Pep-13 is present on the plant plasma membrane, indicating the presence of a surface receptor (Nürnberger *et al.*, 1994). By chemical crosslinking and photo-affinity labeling assays, it was found that the putative receptor could be a monomeric 100 kDa protein on the plasma membrane of parsley (Nennstiel *et al.*, 1998). In addition, a downstream component of Pep-13 response was found to be a mitogen-activated protein (MAP) kinase, which is translocated into the nucleus after Pep-13 treatment (Ligterink *et al.*, 1997).

Map-based cloning is a powerful genetic approach to clone genes, and most potato resistance genes against *Phytophthora infestans* (*Rpi* genes) were cloned by map-based cloning (Haverkort *et al.*, 2016). Traditionally, map-based cloning in potato is a time-consuming work. Approximately 80 available SSR markers that are wide spread on potato genome, could potentially be used to map genes onto a particular chromosome (Domazakis *et al.*, 2017), and fine mapping could be performed using an extended segregating population. However, with the fast development of sequencing technology, genotyping by sequencing is becoming a more efficient and economical way to map genes in segregating populations (Liu *et al.*, 2012). Several genes have been reported to be mapped by bulked segregant RNA-Seq (BSA-RNAseq) analysis (Zou *et al.*, 2016). In potato, enrichment sequencing technology has been used for mapping resistance genes and the surface immune receptors (Jupe *et al.*, 2013; Witek *et al.*, 2016) (**Chapter 3**), however, whether BSA-RNAseq could be used for mapping potato genes is still unknown.

In this study, we investigate two research questions, 1) Whether Pep-13/25 can be recognized by a wild-range of Solanaceae plants, and 2) What is the receptor of Pep-13 and Pep-25 in potato? To address these questions, we used the effectoromics approach (Vleeshouwers *et al.*, 2008) to study the Pep-13 and Pep-25 recognition pattern in Solanaceae family. Pep-13 and Pep-25 peptides were infiltrated on wild and cultivated Solanaceae species, including 8 *Nicotiana tabacum* cultivars, *Nicotiana benthamiana*, 10 pepper accessions, 7 eggplant accessions, 24 tomato accessions, 24 potato cultivars and 146 wild potato genotypes. Furthermore, a potato population that is segregating for Pep-25 response was generated. In combination with a BSA-RNAseq analysis approach, we fine mapped the candidate Pep25 receptor to a 330kb interval based on the DM potato genome. One bacterial artificial chromosome (BAC) clone within the mapping interval from GIG362-6 was isolated and sequenced, and RNAseq data indicates that 3 candidate *RLK* genes from the sequenced region are up-regulated after *P. infestans* infection.

Results

Recognition of Pep-13 and Pep-25 in plants of the Solanaceae family

Pep-13 and Pep-25 are conserved motifs of the cell wall associated transglutaminase (GP42) in many *Phytophthora* species (Brunner *et al.*, 2002). Pep-13 is completely conserved in 6 tested *Phytophthora* species and in 5 of these species Pep-25 was also completely conserved.

Amino acid 23 in Pep-25 of *P. cinnamomi* and *P. sojae* were found to be different. Pep-13 was entirely covered by the amino acid order of Pep-25 **(Figure 1A)**.

Pep-13 can elicit defense response or cell death in parsley and potato (cultivar Désirée; (Hahlbrock *et al.*, 1995; Halim *et al.*, 2004). It is considered to be a MAMP/PAMP molecule. However, whether both peptides could be universally recognized by more species belonging to the Solanaceae family remains unknown. To answer this question, we used plants belonging to the Solanaceae family to determine whether both peptides can be widely recognized in this plant family. Both Pep-13 and Pep-25 peptides were synthesized by Genscript (USA), dissolved by MQ_ water and diluted to a proper concentration by tap water for functional screening **(Materials and Methods) (Figure 1B)**.

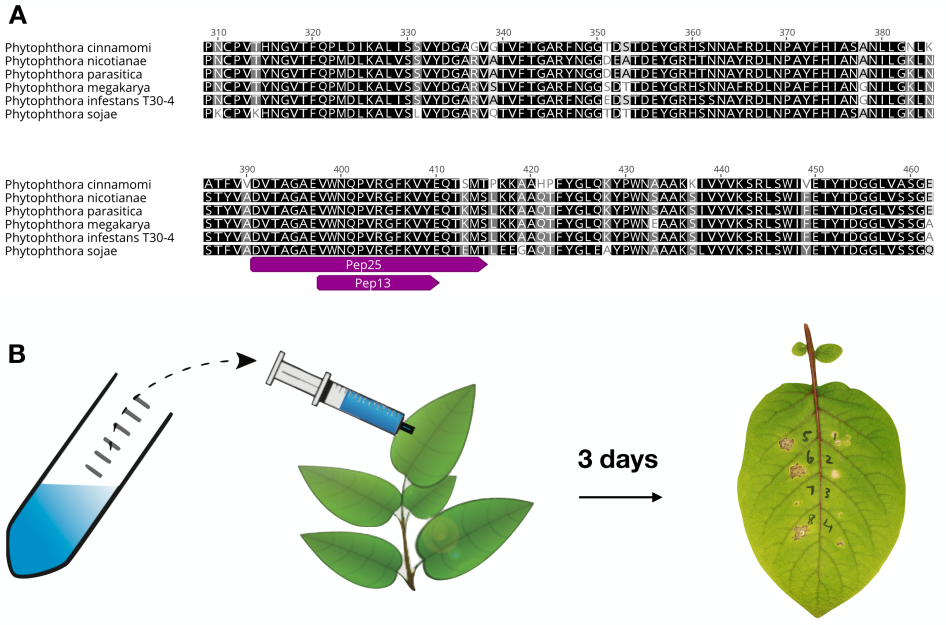


Figure 1. Sequence of Pep-13/ 25 and the workflow of Pep-13/25 peptide infiltration in Solanaceae species.

A. Part of the GP42 proteins (309 aa to 462 aa) from 6 *Phytophthora* species are shown, the consensus amino acids are highlighted in black or grey based on the similarity, the polymorphic amino acids are not highlighted. Pep-13 and Pep-25 from *P. sojae*, which are used in this study, are shown by purple arrows.
B. The peptides are dissolved in tap water, adjusted to a proper concentration, and infiltrated into leaves by needleless syringe. The leaves are visually scored for occurrence of cell death responses 3 days after infiltration.

Peptide infiltration was performed in 146 wild potato genotypes belonging to 56 species **(Supplementary Table S1)**. Among the 146 genotypes, 14 genotypes that belong to at least 8

different tuber-bearing *Solanum* species/ sub-species showed cell death after infiltration of Pep-13 and/or Pep-25 **(Table 1, Figure 2)**. These 8 wild potato species/ sub-species include *Solanum doddsii* (DDS), *Solanum demissum* (DMS), *Solanum edinense* (EDN), *Solanum hondelmannii* (HDM), *Solanum leptophyes* (LPH), *Solanum microdontum* (MCD), *Solanum microdontum* subsp. *gigantophyllum* (GIG), *Solanum chacoense* (CHC), *Solanum ehrenbergii* (EHR) and two unknown species (SPEC).

Table 1. 14 tuber-bearing *Solanum* genotypes recognize Pep-13 , Pep-25 or both.

Series	Genotypes	Species	Contries	Ploidy and EBN	Pep-13	Pep-25	water
Tuberosa II	LPH 680-5	*S. leptophyes*	Bolivia	2x (2EBN)	-	+	-
Tuberosa III	DDS 144-3	*S. doddsii*	Bolivia	2x (2EBN)	+	+	-
Tuberosa III	HDM 646-4	*S. hondelmannii*	Bolivia	2x (2EBN)	+	+	-
Tuberosa III	MCD 958-3	*S. microdontum*	Bolivia	2x (2EBN)	-	+	-
Tuberosa III	GIG 362-6	*S. microdontum subsp. gigantophyllum*	Bolivia	2x (2EBN)	-	+	-
Tuberosa III	GIG 956-1	*S. microdontum subsp. gigantophyllum*	Bolivia	2x (2EBN)	+	+	-
Demissa	DMS 364-1	*S. demissum*	Mexico	6x (4EBN)	-	+	-
Demissa	EDN 150-4	*S. edinense*	Mexico	5x	+	+	-
Demissa	EDN 151-1	*S. edinense*	Mexico	5x	+	+	-
Demissa	EDN 151-3	*S. edinense*	Mexico	5x	+	+	-
Yungasensa	CHC 544-1	*S. chacoense*	Bolivia	2x (2EBN)	+	+	-
Pinnatisecta	EHR 153-3	*S. ehrenbergii*	Mexico	2x (1EBN)	+	+	-
	SPEC 165-2	Unknown species	N/A	N/A	+	+	-
	SPEC 381-2	Unknown species	N/A	N/A	-	+	-

The series (Spooner *et al.*, 2014), genotypes and species/sub-species are shown in this table, "+" : cell death response; "-": no cell death response; "N/A": not applicable.

The results in **Table 1** and **Supplementary Table S1** show that Pep-13 and/or Pep-25 recognition is only found in 14 (out of 146) tested wild potato genotypes. All of the 14 genotypes responded to Pep-25, 9 genotypes responded to both Pep-13 and Pep-25. This result indicates that both peptides do not have a broad recognition spectrum, which supports the probable existence of a narrow spectrum MAMP/PAMP concept (Thomma *et al.*, 2011).

The Pep-25 responding genotype *Solanum microdontum* subsp. *gigantophyllum* (GIG362-6) **(Figure 2)** was earlier used for mapping the receptor of SCR74-B3b (**Chapter 3 and Chapter 4).** Therefore several populations were already made previously, they were used here to map the Pep-25 receptor in potato.

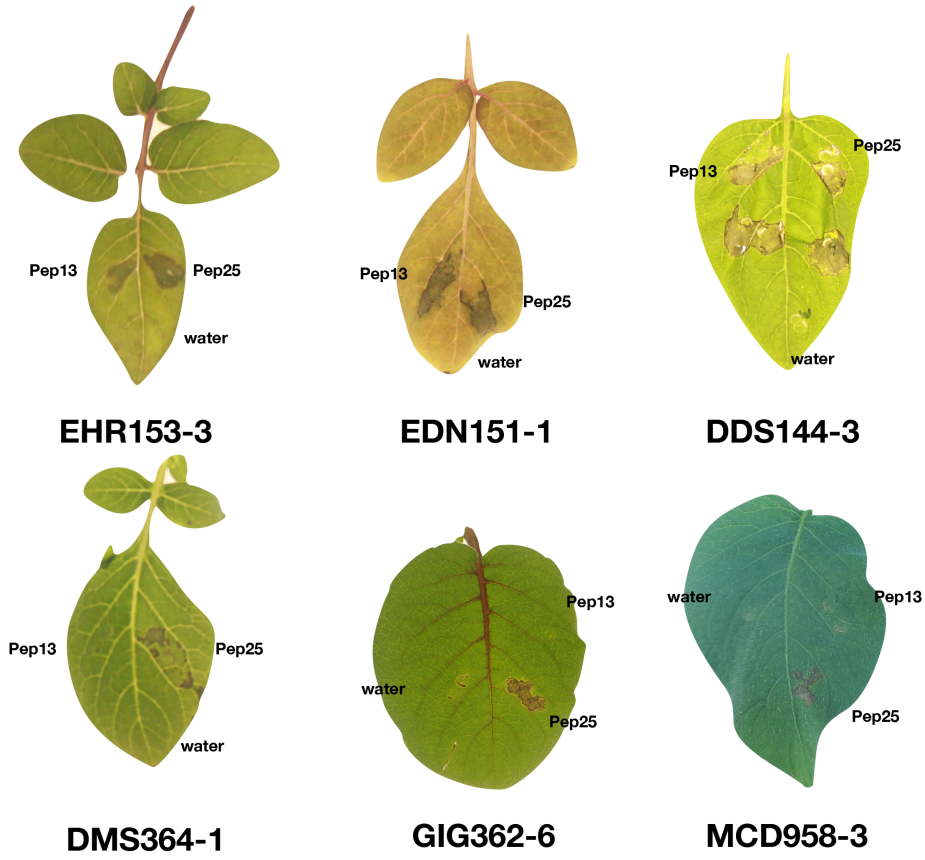


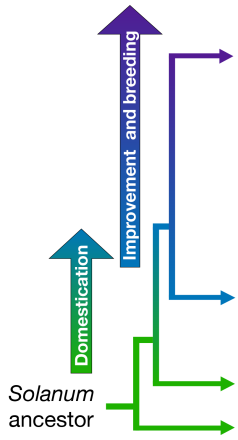
Figure 2. Examples of the Pep-13 and/or Pep-25 responsive genotypes.

Leaves from 6 wild *Solanum* accessions are showing cell death responses after infiltrate Pep-13 and /or Pep-25 peptides. The genotypes EHR153-3, EDH151-1, DDS144-3 responded to both peptides, but other genotypes like DMS364-1, GIG362-6 and MCD956-3 responded to Pep-25 only.

5

Pep-13/ Pep-25 recognition in cultivated potatoes

The results from the large screening show that recognition of Pep-13 and/or Pep-25 is found in a rather limited number of wild potato genotypes, with most responding species being related to *S. tuberosum* or *S. demissum* (i.e.. Series Tuberosa and Demissa). A previous report showed that potato cultivar Désirée recognizes both Pep-13 and Pep-25 (Halim *et al.*, 2004). We speculated whether Pep-13/25 recognition would be relatively more common in *S. tuberosum* related potato cultivars, cultivated landraces or proposed progenitors compared to the wild *Petota* germplasm.



Group	Species	Genotype	Pep13	Pep25	Water
Cultivated potato	*S. tuberosum* Group Tuberosum	cv. Aveka	-	+-	-
	S. tuberosum Group Tuberosum	cv. Aventra AL2865	+	+	-
	S. tuberosum Group Tuberosum	cv. Craigs Snowwhite	+	+	-
	S. tuberosum Group Tuberosum	cv. Désirée	+	+	-
	S. tuberosum Group Tuberosum	cv. Bintje	+	+	-
	S. tuberosum Group Tuberosum	cv. Impala	+	+	-
	S. tuberosum Group Tuberosum	cv. Lady Rosetta	+	+	-
	S. tuberosum Group Tuberosum	cv. Monalisa	+	+	-
	S. tuberosum Group Tuberosum	cv. Rosset Burbank	+	+	-
	S. tuberosum Group Tuberosum	cv. Shepody	+	+	-
	S. tuberosum Group Tuberosum	cv. Zhongshu3	+	+-	-
	S. tuberosum Group Tuberosum	cv. Atlantic	-	-	-
	S. tuberosum Group Tuberosum	cv. Carolus	+	+	-
	S. tuberosum Group Tuberosum	cv. Kathadin	+	+-	-
	S. tuberosum Group Tuberosum	cv. Nicola	+	+	-
	S. tuberosum Group Tuberosum	cv. Premiere	-	-	-
	S. tuberosum Group Tuberosum	cv. Sarpo Mira	-	-	-
	S. tuberosum Group Tuberosum	cv. Toluca	+	+	-
	S. tuberosum Group Tuberosum	RH89-039-16	-	-	-
Cultivated landraces	*S. tuberosum* Group andigena	tub ADG 240-2	+	+	-
	S. phureja	DM (1-3 516 R44)	+	+	-
	S. phureja	PHU 372-8	+	+	-
	S. phureja	PHU 200-4	+	+	-
	S. stenotomum	STN 829-3	+	+	-
Proposed progenitor of cultivar potato	*S. candolleanum*	CN D531-3	-	-	-

Wild species

Figure 3. Pep-13/25 cause cell death in most potato cultivars, cultivated landraces and a proposed progenitor of potato cultivars.

The flow chart on the left reflects the history of potato domestication (Hardigan *et al.*, 2017). On the right side of the Figure, the *Solanum* species are listed, as well as the response to Pep13/25 infiltration of each genotype. The clear cell death responses are represented by "+" and highlighted in red, relatively weak cell death phenotype or variations between repeats are shown by "+-" and highlighted in orange, absence of cell death is marked as "-" . All data points are supported by at least 6 repeats.

To address this question, we deployed the same protein infiltration workflow **(Figure 1)** to screen 19 potato cultivars, including the previous reported potato cultivar Désirée as a positive control, 5 progenitor species of cultivated potato, i.e., *S. tuberosum andigena* (ADG240-2), *S. phureja* (DM 1-3 516R44), PHU372-8 and PHU200-4, *S. stenotomum* (STN829-3), and the most likely wild species thought to be a progenitor of cultivars *S. candolleanum* (CND531-3).

Our results show that 14 out of 19 potato cultivars did recognize both peptides and five did not at all, or showed weaker cell death phenotype to Pep-25 **(Figure 3)**. All 5 tested landraces recognized both peptides, however, the proposed progenitor of cultivated potato *S. candolleanum* (CND531-3) did not.

Some examples of responses to both peptides in cultivated potatoes are shown in **Figure 4**. The sequenced double monoploid potato DM 1-3 516R44 (Xu *et al.*, 2011) can recognize both peptides. This indicates that the candidate receptor should be functionally present in the

plant with a sequenced genome. This observation will greatly increase the speed of mapping and cloning of the putative Pep-13/25 receptor gene.

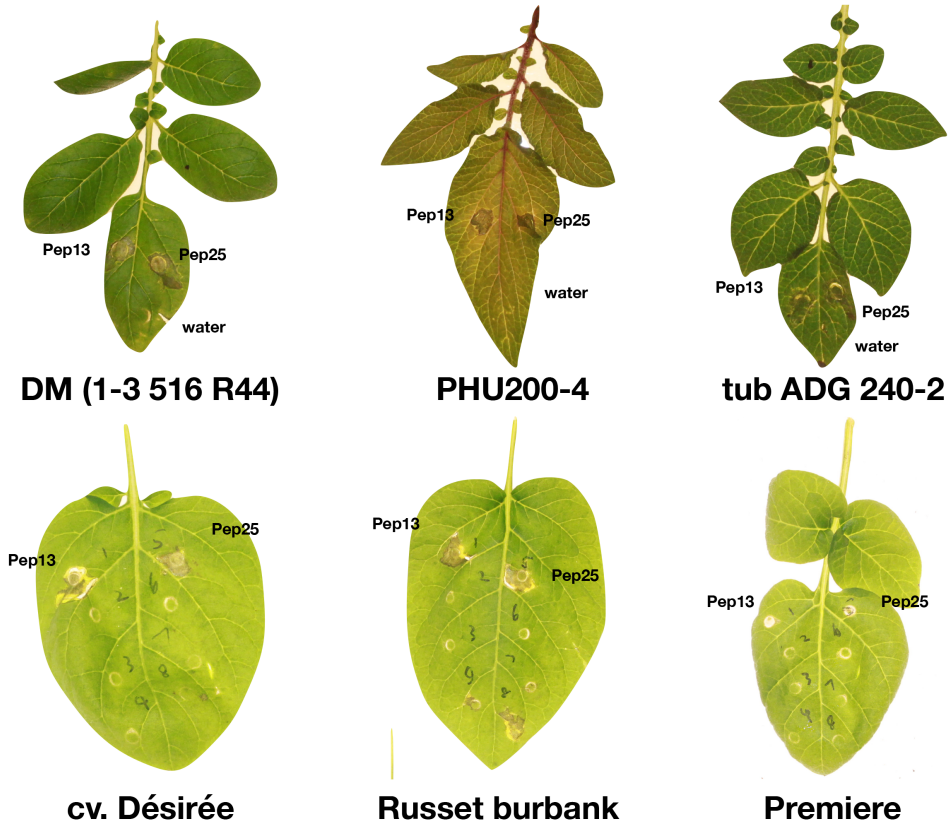


Figure 4. Examples of Pep-13/ 25 recognition in potato cultivars and cultivated landraces. The photos were taken 3 days after the Pep-13/25 infiltration.

Pep-13/25 recognition is restricted to Solanum section Petota

To further determine whether Pep-13 and Pep-25 are broad or narrow spectrum MAMP, additional Solanaceae species, including 24 tomato accessions, 7 eggplant accessions, 10 pepper accessions and 8 Nicotiana accessions were tested for Pep-13 and Pep-25 recognition. We infiltrated the Pep-13/25 peptides as described previously (Figure 1) , Pep-13/25 responsive potato cultivar Désirée was used as positive control.

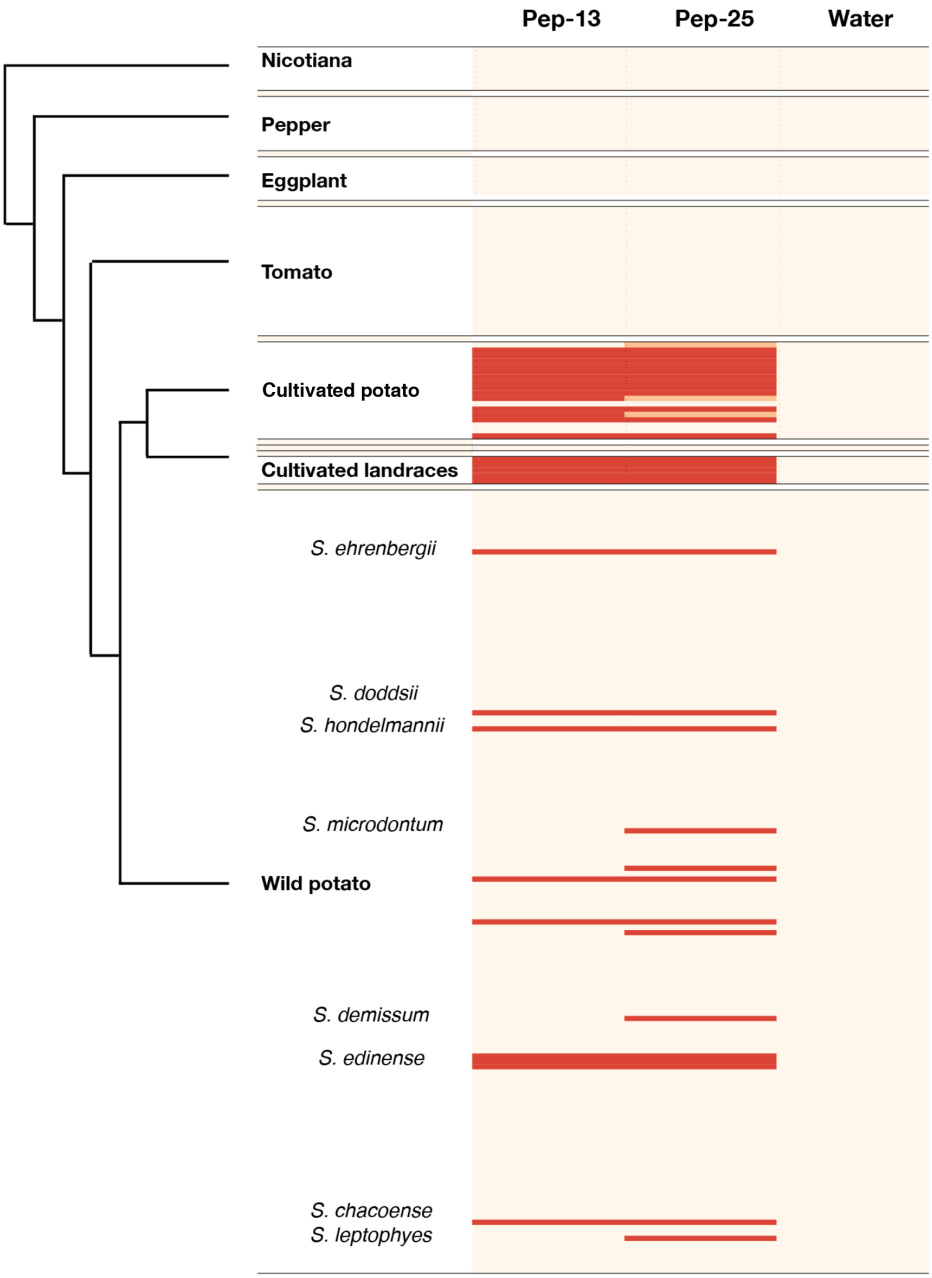


Figure 5. Heatmap of Pep-13/ 25 response in Solanaceae species. The taxonomy of these Solanaceae species is reflected by the tree-like illustration on the left. The Pep- 13/ 25 responding genotypes are shown in red in the heat map, otherwise are beige, some intermediate phenotype is shown by orange. The wild potatoes are sorted based on a Bayesian tree generated by AFLP data **(Chapter 3)**.

However, it was found that none of the tested plants of tomato, eggplant, pepper or *Nicotiana* can recognize either Pep-13 or Pep-25. This result indicates that both peptides are representing a narrow spectrum MAMP/PAMP, which only triggers cell death responses in a subset of *Solanum* species **(Figure 5)**.

Figure 5 provides a comprehensive picture of our Pep-13/ 25 screening on Solanaceae species. To sum up, Pep-13/ 25 recognition is mainly present in cultivated potatoes and a restricted number of wild species from the *Solanum* section *Petota*.

Genetic basis of Pep-25 recognition

Although the responses of Pep-13 and Pep-25 in plants are well characterized, the corresponding plant receptor is still unknown. As described above and illustrated in **Figure 2**, *Solanum microdontum* subsp. *gigantophyllum* (GIG362-6) was found to recognize Pep-25, but not Pep-13. To study the genetic basis of Pep-13/25 recognition, two mapping populations (7026 and 3521) were tested for their segregation patterns after Pep-25 infiltration **(Figure 6)**.

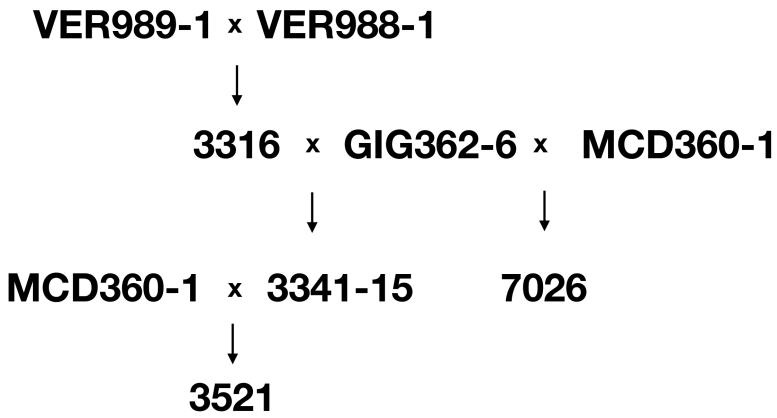


Figure 6. Crossing scheme of Pep-25 segregating populations.

Population 7026 was crossed by GIG362-6 (Pep-25 responding) and MCD360-1 (Pep-25 non-responding). 3341-15 was crossed by 3316-12 x GIG362-6, and crossed to MCD360-1 to generate population 3521 (3341-15 x MCD360-1).

By crossing with *Solanum microdontum* (MCD360-1), which does not respond to either Pep-13 or Pep-25, population 7026 (GIG362-6 x MCD360-1) was generated. Upon testing

this population for Pep-25 recognition, we found that all the 7026 progeny plants responded to Pep25. This suggest that the Pep-25 receptor locus is homozygous in GIG362-6.

Then, another population 3521 (3341-15 x MCD360-1) was tested for segregating to Pep-25 response (**Figure 6** and **Figure 7**). 110 progeny plants were tested and 49 of them responded to Pep-25, but 61 did not. This segregating ratio is close to 1:1 (χ^2 p=0.252), suggesting that the response to Pep25 is a dominant trait and the receptor locus is heterozygous in the diploid parent 3341-15.

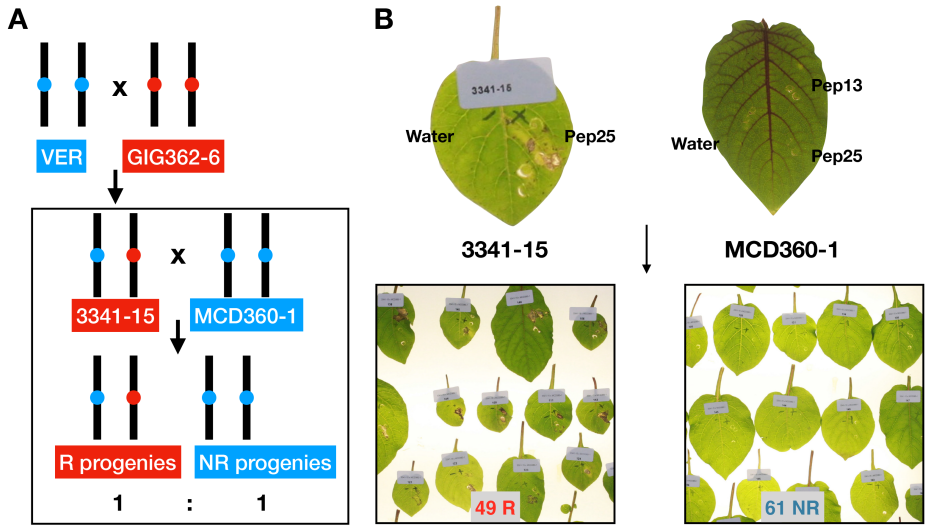


Figure 7. Population 3521 (3341-15 x MCD360-1) segregates for the Pep-25 response.

A. Genetic model of Pep-25 receptor, the Pep-25 responsive (R) genotypes are highlighted in red, the Pep-25 non-responsive (NR) genotypes are highlighted in blue. The Pep-25 receptor locus is represented by the red dots, otherwise is blue. There are two functional Pep-25 receptor loci present in GIG362-6 (homozygous), all the F1 progenies (population 3341) are heterozygous for the Pep-25 receptor loci. 3341-15 was selected to be crossed with the Pep-25 non-responsive genotype MCD360-1, population 3521 (3341-15 x MCD360-1) segregates in a 1: 1 ratio for Pep-25 response;

B. Pep-25 recognition of 3341-15 (heterozygous), MCD360-1 and the F1 progenies of this population, R: Pep-25 responsive; NR: Pep-25 non-responsive.

Fast mapping of the Pep-25 receptor by BSA-RNAseq

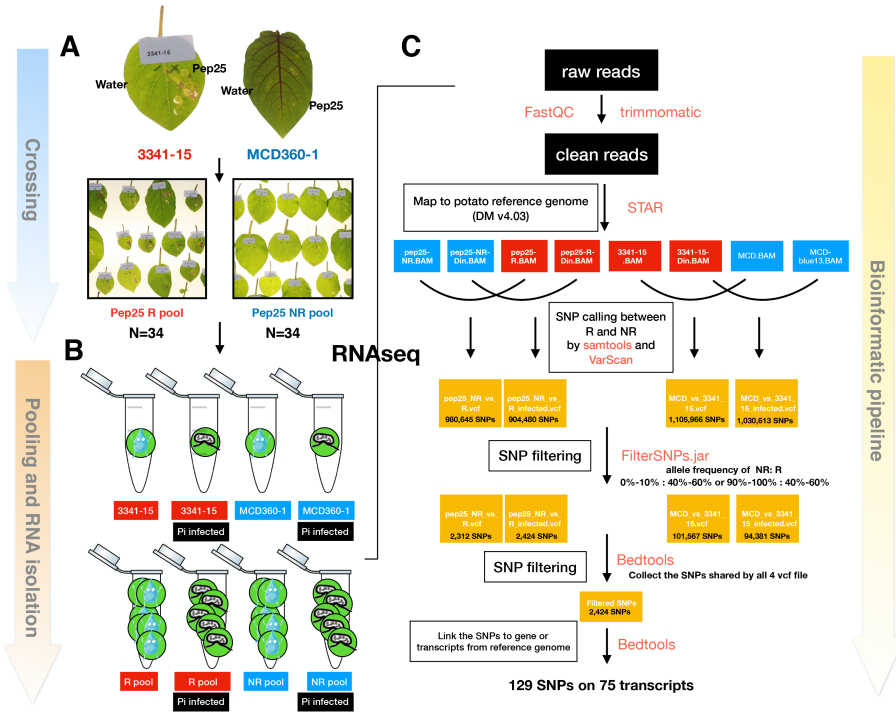


Figure 8. Pipeline of the BSA-RNAseq

A. Segregating population was generated for mapping the Pep-25 receptor, 34 Pep-25 responsive progenies and 34 Pep-25 non-responsive progenies were selected and pooled for RNA isolation.
B. Untreated samples and *P. infestans* infected samples are included for potential differential expression data. In total, 8 samples were included: 1. 3341-15 without infection; 2. 3341-15 with *P. infestans* isolate Dinteloord infection; 3. MCD360-1 without infection; 4. MCD360-1 with *P. infestans* isolate UK3928A infection; 5. Pep-25 responding pool without infection; 6. Pep-25 with *P. infestans* isolate Dinteloord infection; 7. Pep-25 non-responding pool without infection; 8. Pep-25 with *P. infestans* isolate Dinteloord infection.
C. Bioinformatics pipeline of the BSA-RNAseq, the raw reads are cleaned and mapped to potato reference (v4.03) genome by STAR, then SNPs are called and filtered **(Materials and Methods)**.

The segregating population 3521 (3341-15 x MCD360-1) is suitable for identifying the Pep-25 receptor by map-based cloning. To accelerate the map-based cloning, a BSA-RNAseq approach was used for fast mapping the Pep-25 receptor in population 3521 (3341-15 x MCD360-1) **(Figure 8)**.

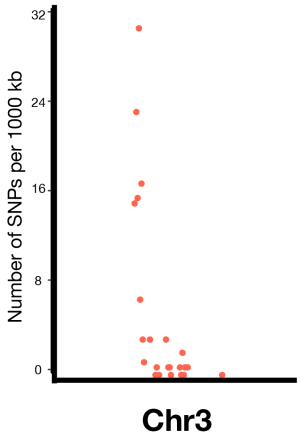


Figure 9. Pep-25 receptor is mapped to chromosome 3.

Chromosome 3 is shown on the X axis. The red dots present the number of SNPs per 1000 kb bin, most of SNPs are located on chromosome 3.

Two pools were made of both 34 Pep-25 responding plants and 34 non-responding plants, together with the responsive parent 3341-15 and the non-responsive parent MCD360-1. All these samples were treated with water control as well as with zoospores of *P. infestans* isolate Dinteloord or UK3928A, and sampling was done two days after inoculation **(Figure 8A and 8B)**. We pooled, as earlier indicated, the samples before RNA isolation, hole punches were used for sampling, the same amount of leaf discs were pooled based on Pep-25 response or non-response. The RNA was sent to Novogene for quality control and sequencing (12GB raw data, HiSeq, pair-end). The raw RNAseq reads were cleaned for mapping to reference genome, the adapters and low quality reads were trimmed by trimmomatic (v0.36), the quality of reads was test by FastQC (v0.10.0) and visualized by MultiQC (v1.4) **(Figure 8C)**. Then the reads from different samples were mapped to the potato reference genome DM v4.03 by STAR **(Materials and Methods)**. Eventually 129 SNPs were filtered, which are located on 75 transcripts, and 126 SNPs were located on chromosome 3 of the potato reference genome **(Figure 9)** and 3 SNPs ending up on chromosome 1 (data not shown).

Marker development based on the BSA-RNAseq

The BSA-RNAseq method provided a promising fast mapping pipeline. The putative Pep-25 receptor was mapped on the chromosome 3, and all these SNPs can be converted to HRM markers. To test the reliability of the BSA-RNAseq methods, and develop markers for fine mapping and positional cloning of the Pep-25 receptor **(Figure 10)**.

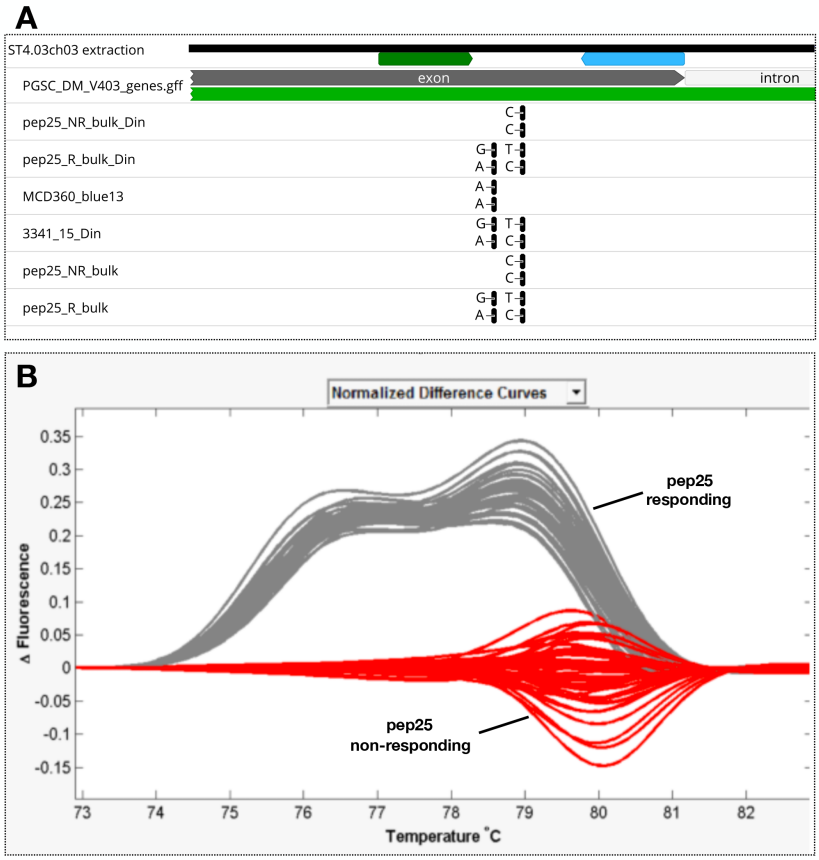


Figure 10. Conversion of BSA-RNAseq markers to HRM markers for application in the lightscanner.

A. The genome and SNP information are visualized by Geneious R10. The SNPs and their positions are shown in black (blocks), the SNPs are filtered based on our criterion, indicating that the SNPs are homozygous on the non-responding progenies or parent, and the allele frequency is close to 1:1 in the responding progenies . Primers (forward primer: green arrow, reverse primer: blue arrow) are designed and flanking the filtered SNPs.
B. This marker (M2) was tested in population 3521 (3341-15 x MCD360-1), and it can distinguish the Pep-25 responding progenies (grey) and Pep-25 non-responding progenies (red). Recombinants are not shown in this figure.

Firstly, 22 markers based on the BSA-RNAseq were designed and tested on population 3521 (3341-15 x MCD360-1) and mapping parents. Among those, 6 markers could distinguish the Pep-25 responding /non-responding progenies **(Materials and Methods)**. To obtain flanking markers for fast genotyping of a large population, a small population (n=176) was phenotyped after Pep-25 infiltration and genotyped by the HRM markers. Two flanking markers M2 and M13 were obtained. This result indicates that the Pep-25 receptor is located between these two flanking markers. The size of the mapping interval is 1.73 Mb, based on the physical map of the potato DM reference genome (data not shown).

Fine mapping of the Pep-25 receptor within a 330kb (0.42cM) interval

To fine-map the Pep-25 receptor in population 3521 (3341-15 x MCD360-1), additional ≈1700 seeds were sown in the greenhouse, the DNA from 1550 progeny plants was isolated, and the flanking HRM markers M2 and M13 were used for genotyping the progeny plants **(Materials and Methods)**. In total, 170 recombinants were selected, and additional markers were tested on these plants to narrow down the mapping interval. By adding more markers within the mapping interval, 19 recombinants were selected **(Figure 11B)** and tested by 12 selected HRM markers **(Materials and Methods)**. Two new flanking markers M71and M80 are representing the new boundary of the Pep-25 receptor, and three co-segregating markers with Pep-25 response were detected: M66, M4 and M63. The candidate gene should be located very close to these co-segregating markers.

According to our genetic model **(Figure 7)**, the Pep-25 receptor is inherited from GIG362-6, It is noteworthy that the marker order in our mapping population is not exactly the same as in the potato reference genome. This finding suggests an inversion event occurred in this region **(Figure 11A)**. Additionally, there are a few gaps in the reference genome which is present within the mapping interval, therefore, the precise physical mapping interval is still unknown.

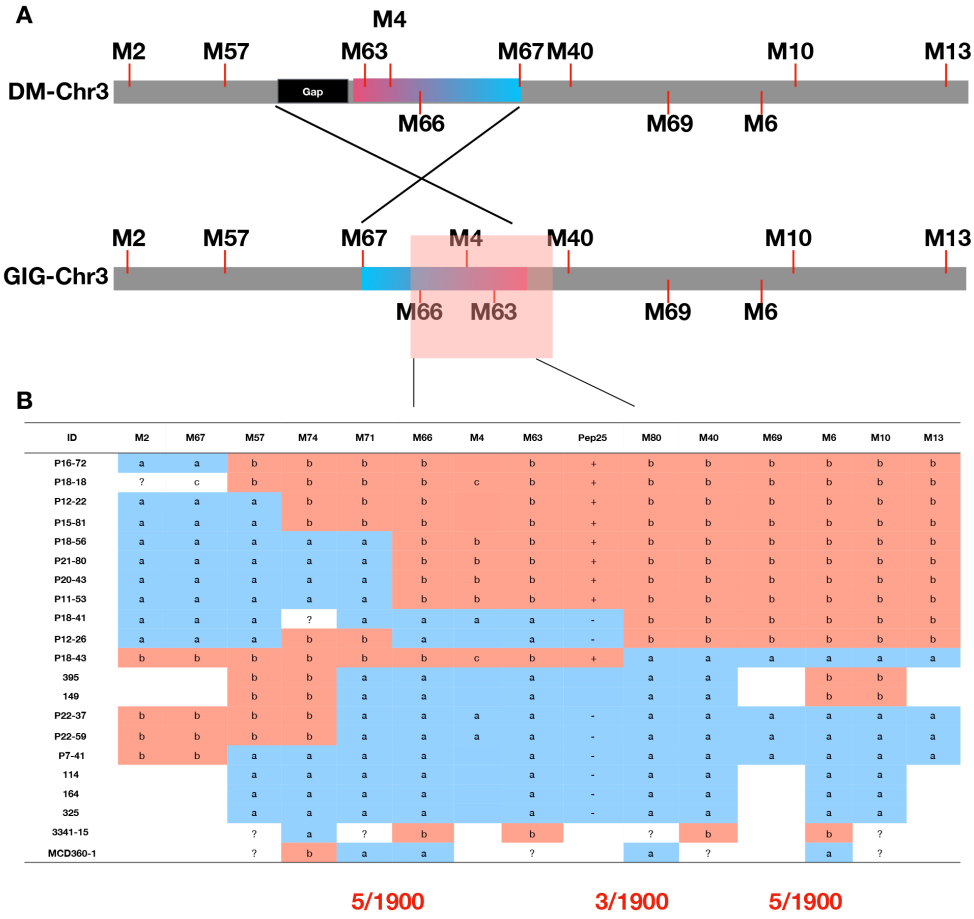


Figure 11. Fine mapping of the Pep-25 receptor into a ≈330 kb (0.42 cM) interval

A. Diagram of the physical map of the DM genome, and the predicted physical map of one of the GIG362-6 haplotypes. Based on the marker position, an inversion event (indicated by lines and gradient pink to blue block) might be present on the GIG362-6 haplotype. A gap is present in the mapping interval of DM (black block). The mapping interval is highlighted by red shading.

B. Diagram with recombinants, the lines are selected recombinant progeny genotypes from population 3521 (3341-15 x MCD360-1), the columns are the markers. The genotyping results are marked by "a" or "b", genotype "b" linked with the Pep-25 receptor. The Pep-25 recognition is recorded by "+" (responsive) and "-" (non-responsive). Number of recombinant events are shown in the bottom.

Candidate genes in the mapping interval

Based on our genetic analysis, the Pep-25 receptor was mapped to a ≈330 kb interval. To have more insight into the genes present in the mapping interval, the physical map of the potato reference genome is visualized **(Figure 12)**. The region contains less than 25 candidate genes in the reference genome, including several receptor-like kinases with an LRR domain (LRR-RLK) and proteases, which might represent candidate genes that mediate the response to Pep-13/25.

To predict the candidate Pep-25 receptor, we also include our BSA-RNAseq data into this analysis. It was hypothesized that: 1) The Pep-25 receptor is a *RLP/RLK* gene; 2) the Pep-25 receptor is up-regulated after *P. infestans* infection; and 3) the informative SNPs from BSA-RNAseq land on the Pep-25 receptor. By combining these criteria, only 3 *RLK* genes did fit (pink arrows in **Figure 12**), based on the DM reference genome.

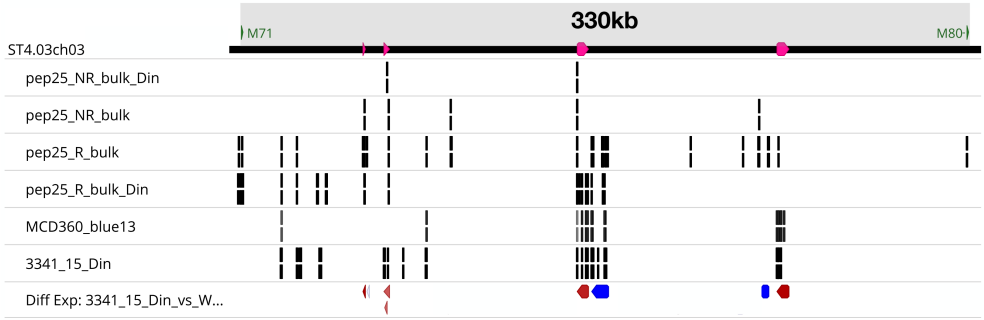


Figure12. The physical map of the target mapping interval on the potato reference genome.

The flanking markers M71 and M80 are shown by green arrows. The black strips are the informative SNPs obtained by BSA-RNAseq from different samples/ pools, under *P. infestans* challenge or water treatment. The differential expression between water inoculation and *P. infestans* challenge was calculated and visualized by Geneious R10, the gene differential expression of 3341-15, between water control and Dinteloord (Din) challenge were shown. Three genes are up-regulated (red arrows) or down-regulated (blue arrows) after Dinteloord (Din) infection.

BAC screening to generate a physical map of GIG362-6

Based on our genetic analysis, the Pep-25 receptor in 3341-15 is inherited from GIG362-6 **(Figure 7)**. The BAC library of GIG362-6 was already generated for cloning the SCR74-B3b receptor (Chapter 4). According to our genetic model, GIG362-6 is homozygous for the functional Pep-25 receptor **(Figure 9)**.

To isolate the BAC clones covering the target mapping interval, several markers were designed for the BAC screening **(Materials and Methods)**. The markers were firstly tested on the responsive parental lines GIG362-6 and 3341-15 and screened in the BAC library. One clone from BAC01-10G was isolated from the GIG362-6 BAC library **(Materials and Methods)** and sequenced. However, this BAC did not cover the whole mapping interval **(Figure 13)**.

The BAC clone was sequenced by PacBio (CNRGV, INRA, France), it contains 131kb, 19 genes are predicted by the program Augustus **(Materials and Methods)**, and the RNAseq reads from 10 samples/ pools (Pep-25 responsive pool with water/ Pi treatment; Pep-25 non-responsive pool with water/ Pi treatment, 3341-15 with water/ Pi treatment, and two previous described samples of GIG-362-6 and MCD360-1 with water/ Pi treatment) were mapped to gigBAC01-10G **(Figure 13)**. The automatic annotation was corrected manually based on the RNAseq data. The expression level and differential expression between samples/ pools were calculated and visualized by Geneious R10 **(Figure 13)**.

On gigBAC01-10G, there is a large region (≈60 kb) with transposable elements related genes like reverse transcriptase or gag-pol polyprotein, this region might be also represented in the reference genome, which caused the "gaps" in the reference genome. In addition to this, this region contains at least 9 full-length or partial receptor-like kinase (*RLK*) genes (g1, g2, g3, g11, g12, g14.1, g14.2, g16, g17), which are normally involved in the MAMPs/PAMPs recognition in the plant immunity system. By comparing the gene expression level between the water treatment and *P. infestans* treatment samples, or the samples with or without Pep-25 recognition, we found that candidate genes g14.1, g14.2 and g16 are highly expressed after *P. infestans* treatment **(Figure 12)**. Especially for g14.2, which is up-regulated 5.2 and 2.4 fold after *P. infestans* infection, in 3341-15 and GIG362-6, respectively, and it is also up-regulated in the Pep-25 responsive pool after *P. infestans* infection.

Collectively, 3 *RLK* genes g14.1, g14.2 and g16 are the top candidate genes of Pep-25 receptor, cloning these candidate genes and functional testing on plants will be the next step.

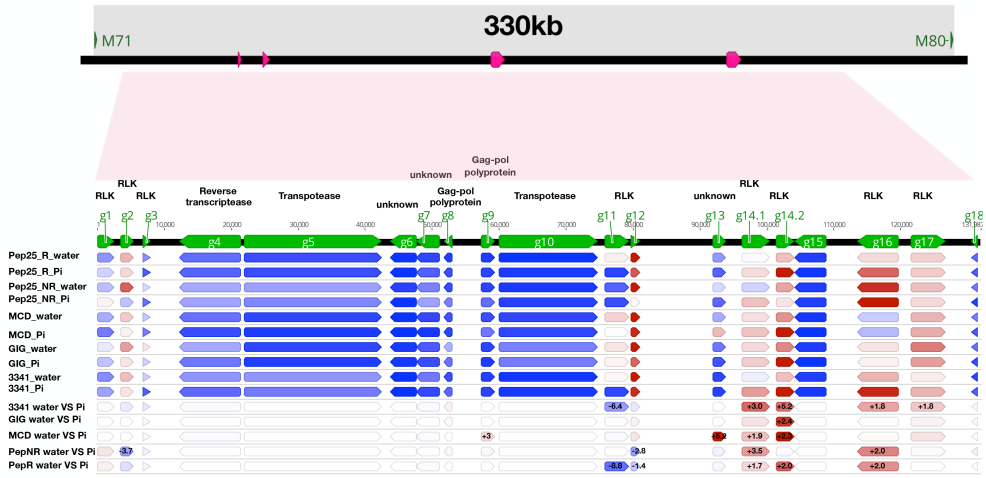


Figure 13. Diagram of BAC clone gigBAC01-10G

The 330 kb mapping interval from DM is shown, gigBAC01-10G was sequenced and is visualized. The 19 genes are shown by green arrows, the types of genes are annotated by Blastx and marked above the genes. The RNAseq reads from each sample/pool are mapped to gigBAC01-10G, and the expression levels are calculated in Geneious R10. The blue arrows are genes with low expression, the red arrows are genes with high expression. The differential expressions between samples are calculated, the fold of differential expression are shown by number if the $P < 0.05$. GIG: GIG362-6, MCD: MCD360-1, 3341: 3341-15, PepNR: Pep-25 non-responsive pool; PepR: Pep-25 responsive pool. Pi: *P. infestans* treatment; water: water treatment.

Discussion

MAMPs/PAMPs are classical terms of plant-microbe interaction, and it refers to the conserved molecular patterns from pathogens and microbes. They can be widely recognized by plants from different families and are essential for the pathogens/microbes involved (Jones and Dangl, 2006), like the flg22 (Felix *et al.*, 1999). However, with the better knowledge on plant-microbe interaction, it's becoming difficult to make an easy definition of MAMPs/PAMPs. The boundary between PAMPs/MAMPs and typical effectors is getting blurred (**Chapter 2**, and Thomma *et al.*, 2011).

In pathogens, Pep-13 and Pep-25 are showing a relatively narrow distribution spectrum (Thomma *et al.*, 2011), because they are conserved in *Phytophthora* species but not in other oomycetes. The TGase activity suggests that the proteins are important to the fitness of *Phytophthora*. Attempts to knock-down/ out TGase in *P. infestans* did not succeed (Parker,

2003). On the other hand, the defense response was so far only found in parsley cell cultures (Nürnberger *et al.*, 1994) and visible cell death in potato leaves (Brunner *et al.*, 2002; Halim *et al.*, 2004). In our study, we tested several species from the Solanaceae family. Our findings indicated that the cell death phenotype to Pep-13/25 recognition is restricted to potato and not found in tomato, eggplant, pepper and *Nicotiana* species. However, the connection between defense response and cell death is still under discussion (Coll *et al.*, 2011), it is also possible that Pep-13/25 can be recognized by multiple plant families but the cell death phenotype is limited to potato. To address this question, more plant families need to be tested for both occurrence of cell death and defense responses like expression of *PR* genes and ROS after Pep-13/25 treatment. It is foreseen that the cloning of the Pep-13/25 receptor from potato will help to answer this question.

From our screening, we have another intriguing finding. Generally, the responding potato genotypes, including most potato cultivars, can recognize both Pep-13 and Pep-15. However, 36% of the responding wild potatoes, only recognize Pep-25 but not Pep-13 **(Table1)**. This suggests that Pep-13 doesn't always contain the minimal units for eliciting cell death response in potato. This observation indicates that the receptor from these plants can only recognize Pep-25, not Pep-13. On the other hand, the data from this large scale screen would benefit from further confirmations, and observed differences might disappear by e.g. altering the infiltrated protein concentrations. These questions can be answered after cloning the receptor from GIG362-6, which seems to be Pep-25 specific, as well as from DM that responds to both Pep-13/Pep-25.

Our Pep-13/25 screening raised another question, whether the Pep-13/Pep-25 receptor from potato is the same gene as parsley? Previously, it was reported that the potential Pep-13 receptor in parsley might be a monomeric 100 kDa protein on the plasma membrane (Nennstiel *et al.*, 1998). Based on our screening data, the recognition of Pep-13 and Pep-25 might have evolved independently in potatoes, but not on other Solanaceae. In line with this, the Pep-13/25 receptor in parsley may also have evolved independently, or, the Pep-13/25 receptor might be the same, but the mechanism of the defense response eliciting and the HR-like cell death may be different in potato and parsley.

Another observation from our Pep-13/25 screening is that most cultivars (*S. tuberosum* Group Tuberosum) or cultivated landraces (*S. tuberosum* Group Andigena, *S. phureja* and *S. stenotomum*) responded to both peptides **(Figure 3)**. Therefore, we specilate that the Pep-13/25 receptor is expected to originate from wild species during domestication, however, most potato cultivars are not resistant to *P. infestans*, so the question has to be answered why

active Pep-13/25 receptors are present in cultivars? It could either be that the Pep-13/25 receptor showed a strong resistance to *P. infestans* but that it was broken by the fast evolving *P. infestans*, potentially by evolving effectors which could inhibit or block the Pep13/25 receptor. For example, a well-studied *P. infestans* RXLR effector Avr3a^{KI} can inhibit cell death response induced by INF1 from *P. infestans* (Bos *et al.*, 2006). Alternatively, the Pep-13/25 receptor might be closely linked to some domestication traits which were preferred in potato cultivars.

To test if the Pep-25 receptor contributes to late blight resistance, we performed a detached leaf analysis on the segregating population to Pep-25 response (data not shown). The Pep-25 responsive group was not more resistant than the Pep-25 non-responsive group, and our data do not support that Pep-25 receptor would enhance late blight resistance. However, the MAMPs/PAMPs triggered immunity acts in the early stage of pathogen invasion, so other measurement methods might be needed to detect the role of Pep-25 receptor in immunity. It is still worthwhile to test the disease resistance to late blight when the Pep-25 receptor is cloned after stable transformation into non-responsive genotype. If Pep-25 receptor is showing a minor effect on late blight, it can be stacked with other surface immune receptors, and with *NLR* genes to achieve a more durable resistance.

In this study, we also developed a pipeline for fast mapping potato genes by BSA-RNAseq **(Figure 8)**. It's the first time to apply BSA-RNAseq on potato, and our result demonstrated that BSA-RNAseq can dramatically accelerate the speed of map-based cloning. Once the segregating population is ready, the BSA-RNAseq can rapidly map the gene to a given chromosome within two months. And it is noteworthy that all the markers needed for fine mapping are ready to use **(Figure 10)**. Meantime, with the infected and non-infected treatment, the differential expression data would also facilitate to obtain the candidate gene.

In this study, the candidate gene has been mapped to a ≈330 kb of the potato reference genome. However, there are some gaps on the reference genome, and they might be caused by the presence of repetitive sequences. We sequenced a BAC clone from GIG362-6 **(Figure 7A)**, and more BAC clones are needed to cover the whole mapping interval. Based on the DM genome and the 1 sequenced BAC clone, our finding indicates that the Pep-25 receptor is located in a *RLK* cluster that contains at least 9 *RLK* genes. By combining with the SNPs information and the differential expression data from BSA-RNAseq, 3 *RLK* genes are our top candidate Pep-25 receptors **(Figure 13)**. The next step is to clone the candidate genes into expression vector and transient express them in plants which do not respond to Pep-25, then infiltrate Pep-25 peptides to see if the cell-death phenotype will be recovered.

Materials and Methods

Synthesis of Pep-13 and Pep-25

Pep-13 (VWN QPV RGF KVY E) and Pep-25 (DVT AGA EVW NQP VRG FKY EQT EMT E) were synthesized by GenScript, the peptides were dissolved by MQ water, the stock concentrations were 2.58 mmolar for Pep13 and 1.68mmolar for Pep25. The working concentration was 1000x dilution of the stock by tap water. For the protein infiltration, the peptides Pep-13 and Pep-25 were infiltrated into abaxial side of plant leaves by needleless syringe and then 3 days after infiltration, scored for symptom development (cell death).

Plant materials used in this study

The seeds of 24 tomato, 7 eggplant and 10 pepper accessions were obtained from the Centre for Genetic Resources, the Netherlands (CGN; see **Materials and Methods** in **Chapter 4**). They were sown in the greenhouse, and the protein infiltration was performed 6 weeks after germination. Three leaves per plant, and 3 plants per peptide were used.

In addition, *Nicotiana benthamiana Nicotiana glutinosa* and 6 cultivars of *Nicotiana tabacum,* including cv. Rustica, cv. White burley, cv. Cleveland, cv. Samsun, cv. Xanthii and cv. SR1 were also tested with the peptides Pep-13 and Pep-25 for response. The seeds were obtained from Unifarm of Wageningen University and Research. Six weeks plants were used for protein infiltration.

All the wild species and potato cultivars were from the *in vitro* collection of Plant Breeding, Wageningen University and Research. These plants are maintained *in vitro* on MS20 medium at 25 ˚C. Top shoots of plants were cut and clonally propagated 2 weeks before transfer to the soil in a climate controlled greenhouse compartment at 22 ˚C/18 ˚C day/night temperature regime under long day conditions.

Generation of the mapping population

The crossings between selected clones were made in different years and the seedlings obtained were propagated and grown *in vitro* on MS20 medium for 2 weeks, before moving to

seedlings into the greenhouse. When the plants were flowering, the mother plant was emasculated before the flower opened, 3 days after, the pollen from the father plant were collected on a small glass slide, then dipped on the pistil of the mother plant. Ripe berries were collected, the seeds were collected and cleaned by water, then they air-dried on filter paper before packaging.

The seeds needed to be sterilized before germinating, 70% ethanol was used for rinsing the seeds, followed by treatment with a solution of 1.5% hypochlorite with one droplet of tween for sterilization. Autoclaved water was used for washing to removing the hypochlorite. For germination the seeds were placed on MS20 medium (with 1000 ppm GA3 if the seeds were new for breaking the dormancy) and put in the dark.

Alternatively, the seeds could be sown in the greenhouse directly for large recombinant screening. After genotyping, the selected recombinants were sterilized and moved into the *in vitro* collection for later analyses.

Sample preparation and RNA isolation for BSA-RNAseq

100 seeds from population 3521 (3341-15 x MCD360-1) were sown in the greenhouse. When they were 6 weeks old, Pep-25 peptides and water were infiltrated for phenotyping. The infiltration was performed on 3 leaves and was repeated at least for 3 times on the same plant. When the phenotyping was clear, the leaves from 34 Pep-25 responsive and 34 Pep-25 non-responsive progenies and the mapping parents 3341-15 and MCD360-1 were collected for inoculation with *P. infestans*. The RNAseq data from Pep-25 non-responsive parent MCD360-1 were ready from previous studies in Chapter 4. 48 hours after inoculation, hole punch with 1cm diameter was used per leaf disc, 2 leaf discs from the same plant are collected into 2ml tubes with 2 small metal beads and immediately frozen by liquid nitrogen. In the end, 6 samples were collected, i.e., 34 Pep-25 responsive progenies with water inoculation or inoculation with isolate Dinteloord , 34 Pep-25 non-responsive progenies with water inoculation or inoculation with isolate Dinteloord and 3341-15 with water inoculation or inoculation with isolate Dinteloord.

The 6 samples were grinded by TissueLyser II (QIAGEN). Then 100mg samples were used for RNA isolation by RNeasy Plus Mini Kit from QIAGEN following industrial instructions. The gDNA eliminator spin column from the kit could efficiently remove the gDNA. The 6 RNA samples were tested by agarose electrophoresis and quantified by Nanodrop

(ThermoFisher). The 6 RNA samples were send to Novogene (Beijing, China) with dry ice for RNA sequencing.

Bioinformatic analysis

Paired-end Illumina HiSeq reads were first checked with FastQC (v0.10.0; http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and the adapters were trimmed with trimmomatic v0.36 (Bolger *et al.*, 2014). The trimmed reads were then mapped on the potato DM genome (v4.03) using STAR v2.5 (Dobin *et al.*, 2013). Pile up files were generated for the bulk and parents using SAMtools mpileup with default settings and piped into VarScan mpileup2snp (v2.3.7) (Koboldt *et al.*, 2012)

SNPs were filtered using the custom Java code (Chapter 3) to retain informative SNPs present in both bulks and both parents. SNPs were filtered based on expected allele ratios in responsive/non-responsive (responsive: Rr; non-responsive: rr) samples. 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 non-responsive and 40-60% alternate allele for responsive. BEDTools intersect (v2.20.1) (Quinlan and Hall, 2010) was used to extract SNPs present in both bulks and parents (informative SNPs) and to relate the informative SNP locations to transcripts of the reference genome. The number of parental, bulk and informative SNPs and variant genes were plotted in 1 Mb bins across each chromosome and visualized using R (Chapter 3).

High resolution melting (HRM) marker development and analysis

The BAM file and VCF file for the filtered informative SNPs were visualized by Geneious R10 (Kearse *et al.*, 2012) (http://www.geneious.com). The primers were designed in Geneious R10, ideally the PCR product should only contain 1 informative SNP and the size is between 80-150 bp. Primers flanking the informative SNPs were manually selected on the conserved sequences of both parents, R and NR bulks. The protocol for DNA isolation and HRM markers can be found in Chapter 4.

The BAC library

The BAC library of GIG362-6 was generated as described in **Chapter 4**. PCR primers were designed in the mapping interval based on the DM genome for detecting the positive BAC clones.

Detached leaf assay

P. infestans isolates were propagated on rye medium for 14 days in a 15°C climate chamber. The zoospores were collected in cold water, and 10 µl suspension (5 x 10^4 zoospores/ml) was used to inoculate plants. The leaves were sampled from 10 weeks-old plants. Samples from Pep-25 responsive progenies and Pep-25 non-responsive progenies from population 3521 (3341-15 x MCD360-1) were picked separately after phenotyping. The experimental procedure was as mentioned previously (**Chapter 2**).

Acknowledgements

This work was supported by a NWO-VIDI grant 12378 (X.L., E.D., V.G.A.A.V.), China Scholarship Council (CSC) (X.L). We thank Isolde Bertram-Pereira for culturing *Solanum* plants, Henk Smid and Harm Wiegersma for help in the greenhouse. We thank the centre for genetic resources, the Netherlands (CGN) for providing the seeds of Solanaceae plants used in this study. We thank Dr. Helene Berges and Caroline Callot from the French Plant Genomic Resource Center (INRA-CNRGV) for their help to sequence the BAC clone.

References

Albert I, Böhm H, Albert M, Feiler CE, Imkampe J, Wallmeroth N, Brancato C, Raaymakers TM, Oome S, Zhang H, et al (2015) An RLP23–SOBIR1–BAK1 complex mediates NLP-triggered immunity. NPLANTS **1**: 15140

Bailey BA, Dean JFD, Anderson JD (1990) An Ethylene Biosynthesis-Inducing Endoxylanase Elicits Electrolyte Leakage and Necrosis in *Nicotiana tabacum* cv Xanthi Leaves. PLANT PHYSIOLOGY **94**: 1849–1854

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30**: 2114–2120

Bos JIB, Kanneganti T-D, Young C, Cakir C, Huitema E, Win J, Armstrong MR, Birch PRJ, Kamoun S (2006) 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*. The Plant Journal **48**: 165–176

Böhm H, Albert I, Oome S, Raaymakers TM, Van den Ackerveken G, Nürnberger T (2014) A conserved peptide pattern from a widespread microbial virulence factor triggers pattern-induced immunity in *Arabidopsis*. PLoS Pathog **10**: e1004491

Brunner F, Rosahl S, Lee J, Rudd JJ, Geiler C, Kauppinen S, Rasmussen G, Scheel D, Nürnberger T (2002) Pep-13, a plant defense-inducing pathogen-associated pattern from *Phytophthora* transglutaminases. The EMBO Journal **21**: 6681–6688

Coll NS, Epple P, Dangl JL (2011) Programmed cell death in the plant immune system. Cell Death and Differentiation **18**: 1247

Dangl JL, Horvath DM, Staskawicz BJ (2013) Pivoting the plant immune system from dissection to deployment. Science **341**: 746–751

Derevnina L, Dagdas YF, la Concepcion De JC, Bialas A, Kellner R, Petre B, Domazakis E, Du J, Wu C-H, Lin X, et al (2016) Nine things to know about elicitins. New Phytol 1–8

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics **29**: 15–21

Domazakis E, Lin X, Aguilera-Galvez C, Wouters D, Bijsterbosch G, Wolters PJ, Vleeshouwers VGAA (2017) Effectoromics-Based Identification of Cell Surface Receptors in Potato. *In* Plant Pattern Recognition Receptors. Humana Press, New York, NY, pp 337–353

Du J, Verzaux E, Chaparro-Garcia A, Bijsterbosch G, Keizer LCP, Zhou J, Liebrand TWH, Xie C, Govers F, Robatzek S, et al (2015) Elicitin recognition confers enhanced resistance to *Phytophthora infestans* in potato. NPLANTS 10.1038–nplants.2015.34

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

Gijzen M, Nürnberger T (2006) Nep1-like proteins from plant pathogens: recruitment and diversification of the NPP1 domain across taxa. Phytochemistry **67**: 1800–1807

Halim VA, Altmann S, Ellinger D, Eschen Lippold L, Miersch O, Scheel D, Rosahl S (2009) PAMP-induced defense responses in potato require both salicylic acid and jasmonic acid. The Plant Journal **57**: 230–242

Halim VA, Hunger A, Macioszek V, Landgraf P, Nürnberger T, Scheel D, Rosahl S (2004) The oligopeptide elicitor Pep-13 induces salicylic acid-dependent and -independent defense reactions in potato. Physiological and Molecular Plant Pathology **64**: 311–318

Hardigan MA, Laimbeer FPE, Newton L, Crisovan E, Hamilton JP, Vaillancourt B, Wiegert-Rininger K, Wood JC, Douches DS, Farré EM, et al (2017) Genome diversity of tuber-bearing *Solanum* uncovers complex evolutionary history and targets of domestication in the cultivated potato. Proceedings of the National Academy of Sciences **114**: E9999–E10008

Haverkort AJ, Boonekamp PM, Hutten R, Jacobsen E, Lotz LAP, Kessel GJT, Vossen JH, Visser RGF (2016) Durable Late Blight Resistance in Potato Through Dynamic Varieties Obtained by Cisgenesis: Scientific and Societal Advances in the DuRPh Project. Potato Res **59**: 35–66

Jones JDG, Dangl JL (2006) The plant immune system. Nature **444**: 323–329

Jupe F, Witek K, Verweij W, Śliwka 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

Ligterink W, Kroj T, Nieden zur U, Hirt H, Scheel D (1997) Receptor-mediated activation of a MAP kinase in pathogen defense of plants. Science **276**: 2054–2057

Liu S, Yeh C-T, Tang HM, Nettleton D, Schnable PS (2012) Gene Mapping via Bulked Segregant RNA-Seq (BSR-Seq). PLoS ONE **7**: e36406

Ma Z, Song T, Zhu L, Ye W, Wang Y, Shao Y, Dong S, Zhang Z, Dou D, Zheng X, et al (2015) A *Phytophthora sojae* glycoside hydrolase 12 protein is a major virulence factor during soybean infection and is recognized as a PAMP. The Plant Cell **27**: 2057–2072

Martin G, Rita C, J MCB (2002) Transglutaminases: Nature's biological glue. Biochem. J

Meijer HJG, van de Vondervoort PJI, Yin QY, De Koster CG, Klis FM, Govers F, de Groot PWJ (2006) Identification of cell wall-associated proteins from *Phytophthora ramorum*. MPMI **19**: 1348–1358

Nennstiel D, Scheel D, Nürnberger T (1998) Characterization and partial purification of an oligopeptide elicitor receptor from parsley (*Petroselinum crispum*). FEBS Letters **431**: 405–410

Nürnberger T, Nennstiel D, Jabs T, Sacks WR, Hahlbrock K, Scheel D (1994) High affinity binding of a fungal oligopeptide elicitor to parsley plasma membranes triggers multiple defense responses. Cell **78**: 449–460

Parker JE (2003) Plant recognition of microbial patterns. Trends in Plant Science. doi: 10.1016/S1360-1385(03)00105-5

Parker JE, Schulte W, Hahlbrock K (1991) An extracellular glycoprotein from *Phytophthora megasperma* f. sp. *glycinea* elicits phytoalexin synthesis in cultured parsley cells and protoplasts. MPMI

Pearce RB, Ride JP (1982) Chitin and related compounds as elicitors of the lignification response in wounded wheat leaves. Physiological Plant Pathology **20**: 119–123

Pel MJC, Pieterse CMJ (2013) Microbial recognition and evasion of host immunity. Journal of Experimental Botany **64**: 1237–1248

Raaymakers TM, Van den Ackerveken G (2016) Extracellular Recognition of Oomycetes during Biotrophic Infection of Plants. Frontiers in Plant Science **7**: 906

Reiss K, Kirchner E, Gijzen M, Zocher G, Löffelhardt B, Nürnberger T, Stehle T, Brunner F (2011) Structural and phylogenetic analyses of the GP42 transglutaminase from *Phytophthora sojae* reveal an evolutionary relationship between oomycetes and marine *Vibrio* bacteria. J Biol Chem **286**: 42585–42593

Séjalon N, Dargent R, Villalba F, Bottin A, Rickauer M, Esquerré-Tugayé MT (1995) Characterization of a cell-surface antigen isolated from the plant pathogen *Phytophthora parasitica* var. *nicotianae*. Can J Bot **73**: 1104–1108

Spooner DM, Ghislain M, Simon R, Jansky SH, Gavrilenko T (2014) Systematics, diversity, genetics, and evolution of wild and cultivated potatoes. Bot Rev **80**: 283–383

Thomma BPHJ, Nürnberger T, Joosten MHAJ (2011) Of PAMPs and Effectors: the blurred PTI-ETI dichotomy. The Plant Cell **23**: 4–15

Verzaux E, Van Arkel G, Vleeshouwers VGAA (2010) High resolution mapping of genes conferring resistance to late blight from two wild species of the Solanum circaeifolium group. edepot.wur.nl

Vleeshouwers VGAA, Rietman H, Krenek P, Champouret N, Young C, Oh S-K, Wang M, Bouwmeester K, Ben Vosman, Visser RGF, et al (2008) Effector genomics accelerates discovery and functional profiling of potato disease resistance and *Phytophthora infestans* avirulence Genes. PLoS ONE **3**: e2875

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

Xu X, Pan S, Cheng S, Zhang B, Mu D, Ni P, Zhang G, Yang S, Li R, Wang J, et al (2011) Genome sequence and analysis of the tuber crop potato. Nature **475**: 189–195

Zou C, Wang P, Xu Y (2016) Bulked sample analysis in genetics, genomics and crop improvement. Plant Biotechnology J **14**: 1941–1955

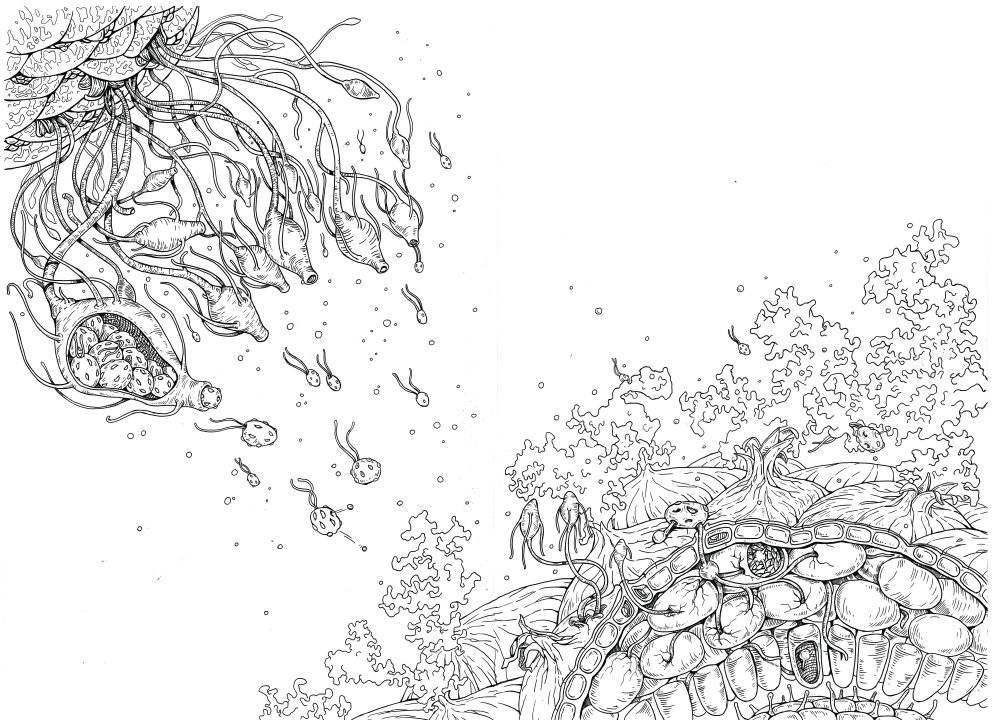
Supplementary data

Supplementary Table S1. Pep-13/25 screening on Solanaceae species

For convenience, this supplementary table is deposited in 4TU.Centre:
https://data.4tu.nl/repository/uuid:9032af40-6cd0-409d-be0d-76fe0f4ed576

Chapter 6

General Discussion



In this thesis, we used the *Phytophthora infestans*- potato pathosystem as a model to study the surface immunity in crop plants. The apoplastic effectors from the Irish famine pathogen *P. infestans* were used as tools to hunt for the surface immune receptors from *Solanum* germplasm. Wild *Solanum* species have been found to respond to a wide spectrum of effectors of *P. infestans*. Here, we are targeting the receptors of a wide diversity of apoplastic effectors, i.e. 1) elicitins, small cysteine-rich (SCR) proteins that are highly conserved and interact with a known (conserved) receptor ELR, 2) SCR74, belongs to a fast-evolving small cysteine-rich (SCR) effector family of *P. infestans* and under positive selection, 3) PcF, which is related to SCR74 but seems to represent a typical MAMP/PAMP in various *Phytophthora* species, 4) Pep-13/25, MAMPs/PAMPs from *Phytophthora* transglutaminase enzyme GP42.

In **Chapter 2**, we used an effectoromics strategy to functionally screen the cell death responses to members of the PcF/SCR family from *Phytopthora*, in a panel of wild Solanaceae species. Our data indicates that PcF and SCR74 act as MAMP/PAMP and effector, respectively. In **Chapter 3**, a novel receptor-like proteins (*RLP)/*receptor-like kinases (*RLK*) gene enrichment sequencing (RLP/KSeq) method was developed to accelerate mapping of plant surface immune receptors. We used the INF1 receptor ELR as a proof of concept and we localized a novel SCR74 receptor on chromosome 9. In **Chapter 4**, we expanded the mapping population and fine-mapped the SCR74 receptor to a *G-LecRK* locus. The candidate *G-LecRK* genes were functionally tested in plants by co-expressing them with the SCR74 effector. In **Chapter 5**, we screened Pep-13/25 peptides for inducing cell death in Solanaceae species. A segregating population was generated for Pep-25 responsiveness and a bulked-segregant analysis and RNAseq (BSA-RNAseq) method was applied, which allowed us to quickly map the Pep-25 receptor to a *RLK* gene cluster on chromosome 3.

I. MAMPs/PAMPs, effectors, surface immune receptors and durable resistance

The plant innate immune system was described by a classic "Zigzag" model that consists of two layers, namely MAMP/PAMP-triggered immunity (MTI/PTI) and effector-triggered immunity (ETI) (Jones and Dangl, 2006). In most cases, the MTI/PTI is mediated by plant pattern recognition receptors (PRRs), which are cell surface localized RLKs and RLPs that typically recognize MAMPs/PAMPs that are stereotypically conserved. However, the term

PRR is used incorrectly sometimes (Zipfel, 2009), because pathogens can also secrete apoplastic effectors that are fast evolving. In the meantime, with the growing knowledge of this effector biology, the boundary between the MAMPs/PAMPs and effectors, and consequently between PTI/MTI and ETI, is less strict in many cases. Therefore, an invasion model was proposed, which describes those ligand/ receptor molecules in a continuous manner (Thomma *et al.*, 2011; Cook *et al.*, 2015).

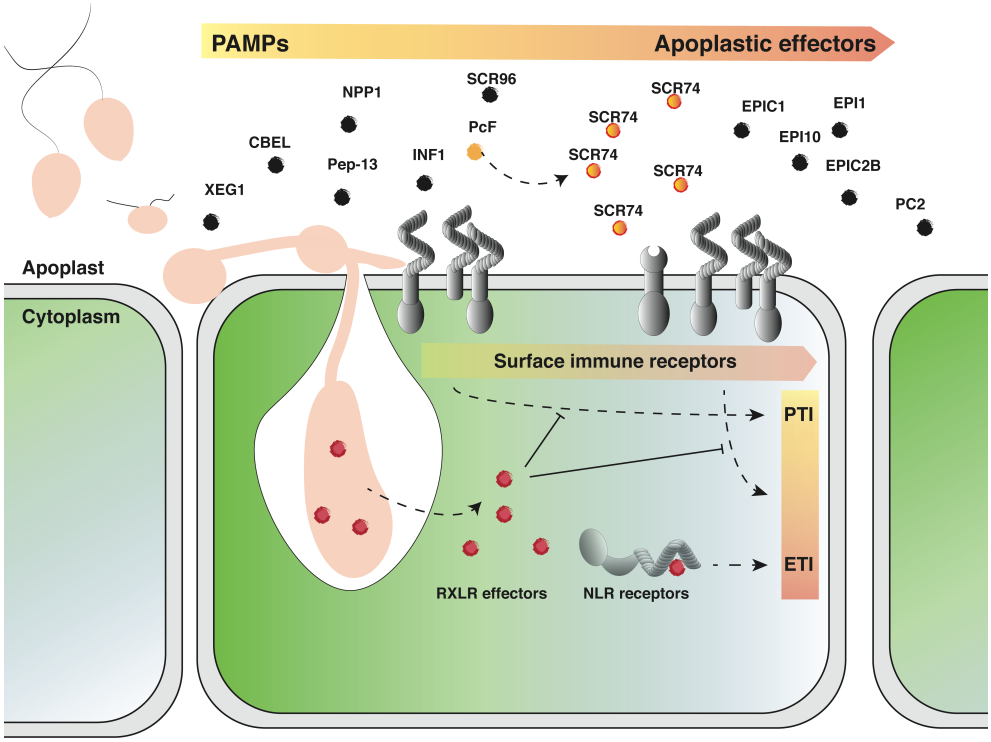


Figure 1. Schematic representation of the *Phytophthora* - host plant interaction, combining the zigzag model and invasion model

Some MAMPs/PAMPs from *Phytophthora* including XEG1, CBEL, Pep-13, NPP1, INF1 and PcF, and some apoplastic effectors like SCR74, SCR96, EPIC1, EPI10, EPI1, EPIC2B. The gradient arrays reflect the invasion model, the MAMPs/PAMPs and effectors are in a continuous state. The PcF/SCR family provides an example that an effector family can possess both typical MAMPs/PAMP and effector features, and we hypothesize that SCR74 family has evolved from PcF, and has been shaped by co-evolution with receptor genes from host plants (dotted line connect PcF and SCR74). Similarly, the surface immune receptors can recognize proteins that fit the criteria of MAMPs/PAMPs or apoplastic effectors, and consequently, they remain conserved, or co-evolve with the apoplastic effectors, respectively. To suppress the surface immunity, *Phytophthora* secretes RXLR effectors into plant cell, which might be recognized by plant NLR receptors and trigger ETI.

In this thesis, we use the term 'surface immune receptors', for receptors that recognize the extracellular molecules from pathogens, including conserved MAMPs/PAMPs and fast evolving apoplastic effectors. These surface immune receptors form the first layer of active defense **(Figure 1)**.

PcF and SCR74, from MAMP to effector and co-evolution with the host receptor

In **Chapter 2**, we studied the PcF/SCR family, proteins that share a common PcF domain, which are specific present in oomycetes. We found that *PcF* is conserved in *Phytophthora*, and can be widely recognized by different plant species, like strawberry, tomato and potato. So, we concluded that PcF fits the concept of MAMPs/PAMPs. On the contrary, we found that the SCR74 family is lineage specific. SCR74 variants are exclusively present in *P. infestans*, their sequences are highly diverse and the genes are under strong positive selection pressure. For these reasons, we consider the SCR74 family do not act as typical MAMPs/PAMPs, but rather representing apoplastic effectors. Therefore, we propose that some MAMP-like molecules, such as PcF, can develop into an apoplastic effector, such as SCR74, that co-evolves with the corresponding receptors of their hosts. In turn, the respective host receptor is expected to be fast-evolving as well.

In this thesis, the putative SCR74-B3b receptor is mapped to a *G-LecRK* locus of potato, as described in **Chapter 4**. The candidate *G-LecRK* genes show copy number variation in the two haplotypes of GIG362-6, with 2 or 5 copies, respectively. By comparing this locus with other sequenced genomes of Solanaceae, we found only 1 *G-lecRK* gene in tomato and pepper genome, and 2 *G-lecRK* in eggplant genome. Intriguingly, the copy number of these *G-lecRKs* in different potato genomes varies dramatically. Namely, 3, 4 and 7 full-length or partial *G-LecRK* genes were found in the DM potato, *Solanum verrucosum* and *Solanum chacoense* genomes, respectively (**Chapter 4 Figure 5**). This observation fits well with our hypothesis that the highly diverse apoplastic SCR74 co-evolves with their host receptor in potato.

Pep13/25, how to not be a good MAMPs/PAMPs

Previously, MAMPs/PAMPs were thought to be conserved epitopes from pathogen proteins, like flg22 or elf18 (Jones and Dangl, 2006), that interact with the corresponding surface immune receptors, like FLS2 or EFR (Gómez-Gómez and Boller, 2000; Zipfel *et al.*, 2006), respectively. However, MAMPs/PAMPs are not invariable, even for the well-studied flg22, a genomic study of *Pseudomonas syringae* pv. *Tomato* (*Pto*) revealed that an ancestral allele of flg22 triggered stronger immune response, but it is rarely present in current *Pto* populations (Cai *et al.*, 2011). And in plants, new surface immune receptors have evolved against different epitopes of the same protein, e.g. FLS3 that recognizes flgII-28 of flagellin (Hind *et al.*, 2016).

In **Chapter 5**, we studied a typical *Phytophthora* MAMPs/PAMPs Pep-13/25, they are conserved in *Phytophthora* species. Pep-13/25 peptides trigger defense and cell death responses in parsley and potato, respectively, but if Pep-13/25 recognition is common in different plant families remained unclear. We screened many Solanaceae species, including potato, tomato, pepper, eggplant and *Nicotiana spp.*, for cell-death responses after Pep-13/25 infiltration, but the cell death was limited to cultivated and wild potatoes. Our data indicates that the ability to respond to Pep-13/25 is very common in the cultivated potatoes (*S. tuberosum*) and some landraces, but relatively rare in the wild potatoes. These data indicate that the surface immune receptor might not be very conserved among plant species. This also raises the question whether the Pep-13/25 receptors from parsley and potato represent the same genes, or, whether the Pep-13/25 recognition evolved independently in parsley and potato? To address this question, more plant species need to be tested for Pep-13/25 recognition, and the cloning of Pep-13/25 receptors will reveal the answer.

Do surface immune receptors confer enhanced disease resistance?

Apoplastic immunity is the first line of active plant defense, a basic immune response which is triggered upon recognition of MAMPs/PAMPs or apoplastic effectors from pathogens. Vice versa, losing this layer of immunity is expected to increase the susceptibility of plants. Indeed,

in a recent study of virus-induced gene silencing (VIGS) of SOBIR1, a co-receptor of RLPs, led to increased susceptibility to *P. infestans* in *N. benthamiana* (Domazakis *et al.*, 2018). This suggests that RLPs play an important role in enhancing disease resistance.

However, with the emerging knowledge on surface immunity, can we expect that introducing new surface immune receptors will increase disease resistance? In principle, PTI and ETI share many common pathways and trigger similar immune responses, like ROS production, MAP kinases activation, callose deposition, and consequently up-regulation of pathogenesis-related (*PR*) genes and plant hormones related genes (Tsuda and Katagiri, 2010). A large-scale transcriptome study shows that PTI and ETI shares many common features at the transcriptional level (Navarro, 2004). However, the PTI responses triggered by PRR occur typically in the early stage of infection, but disappear quickly, whereas the ETI response triggered by NLR receptors are generally prolonged (Tsuda and Katagiri, 2010).

So far, various surface immune receptors were reported to enhance resistance in the laboratory or under greenhouse conditions, e.g. FLS2. EFR, RLP23, CORE, LORE, ELR and RXEG1 (Wang *et al.*, 2016; Zipfel *et al.*, 2006; Lacombe *et al.*, 2010; Albert *et al.*, 2015; Du *et al.*, 2015; Ranf *et al.*, 2015; Zipfel *et al.*, 2004; Hind *et al.*, 2016; Wang *et al.*, 2018) **(Table 1)**. Some other surface immune receptors were shown to enhance disease or pest resistance in the field, like Xa21, and OsLecRK1-OsLecRK3 (Tu *et al.*, 2000; Liu *et al.*, 2015).

Remarkably, many studies show that interfamily transformation of these surface immune receptors may enhance the disease resistance, like EFR and RLP23 (Lacombe *et al.*, 2010; Albert *et al.*, 2015) **(Table 1)**. The practices of interfamily transformation promises to further expand the gene pool for resistance breeding. Theoretically, non-adapted pathogens may have failed to evolve the effectors that inhibit the surface immunity of a non-host plants, and this could lead to a huge practical potential.

Currently, breeding for resistance to late blight is still mainly depending on the *R* genes of the NLR class. Attempt has been made to stack multiple *NLR* genes (Jo *et al.*, 2014), which is more efficient and expected to last longer than single *NLR* gene, however, the pathogens still have a chance to defeat the stacked *R* genes by many strategies (Vleeshouwers and Oliver,

2014). On the other hand, some *NLR* genes have a "fitness cost" for the plants (Tian *et al.*, 2003), so how to fine-tune the plant immune system is crucial to achieve the goal of durable resistance.

Table 1. Surface immune receptors enhance plant disease resistance.

Surface immune receptor	From	Type	MAMPs	Endogenous/ heterologous expression	Pathogens	Reference
CORE	tomato	RLK	csp22	*Arabidopsis*	*Pseudomonas syringae* pv. tomato DC 3000	(Wang, 2016)
EFR	*Arabidopsis*	RLK	EF-Tu	*N. benthamiana*	*Pseudomonas syringae* pv. *syringae* (*Pss*) B728a	(Zipfel, 2006)
				N. benthamiana	*Pseudomonas syringae pv. tabaci (Pta) 11528*	(Lacombe, 2010)
				N. benthamiana	*Agrobacterium tumefaciens*	
				Tomato	*Xanthomonas perforans*	
				Tomato	*Ralstonia solanacearum* GMI1000	
RLP23	tomato	RLP	NLPs (nlp20)	potato	*Phytophthora infestans*	(Albert, 2015)
RLP23				potato	*Sclerotinia sclerotiorum*	
ELR	potato	RLP	INF1	potato	*Phytophthora infestans*	(Du, 2015)
LORE	*Arabidopsis*	G-LecRK	LPS	*Arabidopsis*	*Pseudomonas syringae*	(Ranf, 2015)
FLS2	*Arabidopsis*	RLK	flg22	*Arabidopsis*	*Pseudomonas syringae* pv. tomato DC 3000	(Zipfel, 2004)
FLS3	Tomato	RLK	flgII-28	Tomato	*Pseudomonas syringae* pv. tomato DC 3000	(Hind, 2016)
Xa21	Wild rice	RLK	RaxX	Rice cultivar	*Xanthomonas oryzae* pv. *oryzae* (*Xoo*)	(Wang, 1996)
				Banana	*Xanthomonas campestris* pv. *musacearum* (*Xcm*)	(Tripathi, 2014)
				Sweet orange	*Xanthomonas axonopodis* pv. *citri*	(Mendes, 2010)
				Tomato	Ralstonia solanacearum,	(Afroz, 2011)
Ve1	tomato	RLP	tomato	Cotton	*Verticillium* spp.	(Song, 2018)
				Tobacco	*Verticillium* spp.	(Song, 2018)
				Arabidopsis	*Verticillium* spp.	(Fradin, 2011)
LecRK-I.9	*Arabidopsis*	L-LecRK	eATP	Potato and *N. benthamiana*	*Phytophthora infestans*	(Bouwmeester, 2014)
LecRK-VI.2	*Arabidopsis*	L-LecRK		*N. benthamiana*	*Psrudomonas*	(Huang 2014)
RXEG1	*N. benthamiana*	RLK	XEG1	*N. benthamiana*	*Phytophthora parasitica*	(Wang, 2018)

Except the interfamily transformation, pyramiding *NLR* genes and surface immune receptors can potentially confer strong and more durable resistance. For example, *Bs2* is a *NLR* gene from pepper, which can recognize AvrBs2 from *Xanthomonas campestris* pv. *Vesicatoria*. Transferring *Bs2* to tomato resulted in strong resistance to bacterial spot (Tai *et al.*, 1999) when pyramided with the surface immune receptor EFR from *Arabidopsis.* The transformed tomato shows enhanced and broad-spectrum resistance to many bacterial pathogens (Lacombe *et al.*, 2010). In a more recent study, the combination of *Bs2* and *EFR* in tomato was proven also efficient in the field (Kunwar *et al.*, 2018). Therefore, recruiting both layers of immune systems provide an alternative strategy to achieve more durable resistance.

The most ideal situation is to turn the crops from a host to a nonhost of the deadly pathogens. Nonhost resistance (NHR) can be divided into two classes. Type I NHR does not show any visible HR, whereas Type II NHR results in a rapid HR. *NLR* genes are sometimes involved

in the Type II NHR, by recognition of the *Avr* genes from the pathogen, but the Type II NHR can still be broken by the pathogen (Gassmann *et al.*, 2000). On the other hand, the surface immune receptors might be involved in the Type I NHR. For example, parsley is not a host of *P. infestans*, so the Pep-13/25 induced cell-death was thought to be the source of NHR in parsley (Kamoun, 2001). Therefore, studying the plant surface immunity might reveal the mechanism of type I NHR, stop the pathogen at first sight.

Do the putative SCR74 and Pep-13/25 receptor confer resistance?

To test if the surface immune receptors studied in this thesis can enhance the resistance to *P. infestans*, we performed preliminary detached leaf assay (DLA) experiments to test whether response to SCR74 and Pep-25, respectively, co-segregates with resistance to *P. infestans*, however, such correlation was not detected (**Chapter 2** and **Chapter 5)**. This observation can be explained by the fact that the resistance is too low to be detected among other immune receptors in the wild *Solanum* background, or that the surface immune receptors are inhibited by the *P. infestans* RXLR effectors, or a combination of both.

However, more variables need to be considered in such experiments. For SCR74, most *P. infestans* isolates carry multiple SCR74 homologs, and indeed for the SCR74-responsiveness genotype GIG362-6, which is highly resistant to *P. infestans*, it is expected to carry multiple immune receptors that recognize other effectors or MAMPs/PAMPs. To dissect the contribution of SCR74 among the wide-ranging effector repertoire that is potentially detected in these wild potatoes, the experiment would benefit from testing other *P. infestans* isolates that cause more extreme differences in resistance level of the parental plants (**Chapter 2**). And obviously, a better way to address the question would be to test transgenic potato plants that express the receptor, once isolated, as shown by (Du *et al.*, 2015).

So far, the real biological function of SCR74 remains unknown. Theoretically, the interaction between the apoplastic effectors and surface immune receptors is in the early stage of the infection (Zipfel, 2009). Therefore, whether the DLA that monitors lesion size at later time points can detect the effect of SCR74 is questionable. More detailed studies of early infection stages at the microscopic level might shed more light on the effect of SCR74 and the

matching receptor. On the other hand, a clean genetic background is important for both the *P. infestans* and plants which can be achieved for instance, by generating *P. infestans* strains with only 1 *scr74* gene, and test this strain on the plants with or without the corresponding receptor.

For the Pep-13/25, we found that the Pep-13/25 receptor is likely present in most modern potato cultivars, regardless whether they are resistant or susceptible to *P. infestans.* However, the Pep-13/25 receptor might still contribute to the basic immunity of the plants. Once the receptor has been identified, gene editing technologies like CRISPR/ Cas9 can be used to knock out the receptor in potato cultivars to study the effect. Another possibility is that *P. infestans* has evolved specific RXLR effectors that inhibit the Pep-13/25-triggered immunity (effector-triggered susceptibility, ETS). In such scenario, interfamily transfer of the Pep-13/ Pep-25 receptor to other crops might enhance resistance. For instance, transfer of Pep-13/25 receptor from potato to soybean might enhance the resistance to *P. sojae*, which also carries the Pep-13/25 but might not have evolved the effectors that cause ETS. Further studied need to be performed to see if the RXLR effectors from different *Phytophthora* species can inhibit the Pep-13/25 response.

II. The missing link between plant immunity and plant fertilization system

Plants have an ancient immunity system

The origin of the plant immune system can be traced back to unicellular organisms. Currently, genomics-based phylogeny data indicate that RLKs are the most ancient immune receptors, that even were found back in algae (544 MYA) (Yue *et al.*, 2012). In a recent study, the earliest complete *TIR-NBS-LRR* genes were identified in Bryophytes (360 MYA), indicating that these *NLR* genes are much younger than the *RLK* genes. These phylogenetic analyses agree well with the "zig-zag" model of Jones and Dangl (2006), indicating ETI was deployed when the PTI had been overcome by effectors from pathogens.

Reproduction and self-incompatibility systems in flowering plants

In flowering plants, successful fertilization, or self-incompatibility (SI) requires many steps in the cell-to-cell communication. This is mainly mediated by SCR proteins as signals, and their corresponding receptors for signaling. From an evolutionary point of view, self-fertilization has been seen to be an "evolutionary dead end" for plants (Takebayashi and Morrell, 2001; Wright *et al.*, 2013). In contrast, self-incompatibility and outbreeding of plants increases genetic variation as well as the adaptation and fitness of plants. Plant SI systems have evolved in flowering plants over a long period of time (130 MYA) **(Figure 2)**, leading to greater genetic diversity. It might be the answer of Darwin's "abominable mystery", which was used to describe the booming of flowering plants in the middle of the Cretaceous period (Whitehouse, 1950). Up to date, SI systems have been found in around 40% of flowering plant species and in at least 100 plant families (Fujii *et al.*, 2016). It has been speculated that SI systems evolved independently many times in flowering plants (Igic *et al.*, 2008).

In many plant species, the SI system is controlled by a single *S*-locus from different haplotypes (Takayama and Isogai, 2005) with tightly linked male and female S-determinants, which are specifically expressed in pollen and pistils, respectively. So far, there are three well-studied plant SI systems. Two of them are gametophytic SI (GSI), in which the pollen itself determines the S-specificity, including the *Papaver rhoeas* pollen or stigma *S* gene PrpS-PrsS in Papaveraceae and *S*-locus F-box gene (SLFs)-S-RNase in Solanaceae, respectively (Murfett *et al.,* 1994; Wheeler *et al.*, 2009; Kubo *et al.*, 2010). In the Brassicaceae, a sporophytic SI (SSI) system is active, and in this case, the S-specificity is determined by SRK-SCR proteins (Schopfer *et al.*, 1999; Takasaki *et al.*, 2000) .

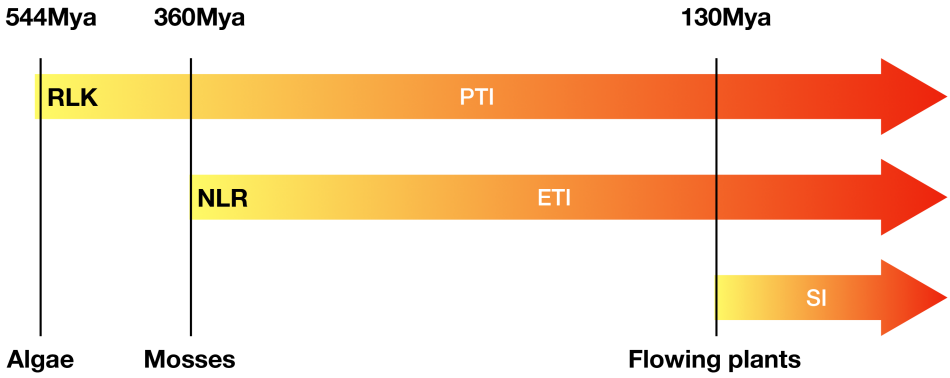


Figure 2. Time scale of the emergence in plants of the PTI and ETI disease immune systems and the self-incompatibility (SI) for fertilization system.

RLK receptors that are involved in PTI plant immunity can be traced back 544 Mya, as receptor-like kinases were found in algae. The first *NLR* genes were found in Mosses (360 Mya), The SI system has evolved in flowering plants 130 Mya ago.

Missing link between plant fertilization and immune system

How SI systems have evolved in plants remain unclear, but it is difficult to expect a de-novo development of such a complicated system. Thus it is tempting to speculate that the plant SI system evolved from another already existing non-self-recognition system, like the plant immune system (Hodgkin *et al.*, 1988; Dickinson, 1994). Except for the plant SI system as a whole, many other processes during plant fertilization, like the pollen tube germination and tip growth, pollen tube guidance and reception, also share similarities with the plant immune system. With the growing knowledge about the plant fertilization system and the plant immune system, we here provide new insight into both systems and try to establish a link. The understanding of the plant self- and non-self-recognition systems enables us to utilize the SI system at one hand and to achieve durable disease resistance on the other hand, or vice versa.

Small cysteine-rich proteins in plants and pathogens

SCR proteins are widely distributed and versatile in different organisms, here we define SCR proteins based on the following criteria: 1. They are small; 2. Have a conserved N-terminal signal peptide; and 3. Contain 4-16 cysteines to form disulfide bridges (Marshall *et al.*, 2011). In plants, small cysteine-rich (SCR) proteins are very versatile, they can act as antimicrobial peptides, such as defensins (DEFs), like RsAFP2, MsDef1 and ZmESR-6 (Terras *et al.*, 1992; Balandín *et al.*, 2005; Sagaram *et al.*, 2011; Van Der Weerden *et al.*, 2012).

Many SCR proteins also participate in multiple steps during double fertilization in flowering plants, for example, LAT52, LeSTIG, SCA and LTP5 are proteins involved in pollen tube germination (Muschietti *et al.*, 1994; Park *et al.*, 2000; Tang *et al.*, 2004; Chae *et al.*, 2010). So far, only the receptor of LAT52 and LeSTIG have been found and they are encoded by the *LRR-RLK* gene LePRKs (Tang *et al.*, 2002). LURE is a pollen attractant which is secreted from synergid cells (Okuda *et al.*, 2009), and has found to be recognized by multiple RLKs, including MIK1-MDIS1 and PRK3/PRK6-PRK6 (Takeuchi and Higashiyama, 2016; Wang *et al.*, 2016). Specific SCR proteins also do participate in pollen tube reception and gamete activation, like RALFs, ZmES1-4, PMEI1 and EC1 (Pearce *et al.*, 2001; Amien *et al.*, 2010; Sprunck *et al.*, 2012; Woriedh *et al.*, 2013). **(Table 2 and Figure 4).**

Additionally, several SCR proteins are involved in plant self-incompatibility, for example, SCR/SP11 and PrsS, which determine SSI and GSI in Brassicaceae and Papaveraceae, respectively (Foote *et al.*, 1994; Shiba *et al.*, 2001). The receptor of SCR/SP11 is SRK, encoded by a *G-LecRK* gene (Takayama *et al.*, 2001) and the receptor of PrsS is PrpS (Wheeler *et al.*, 2009) **(Table 2 and Figure 3)**.

In plant pathogens, many pathogen-encoded apoplastic effectors are also SCR proteins, and the disulfide bridges formed by cysteines can keep them stable to resist the harsh environment in the apoplast. Examples are AVR2, AVR4, AVR4E, AVR5, AVR9 and ECP6 from *Cladosporium fulvum* (Joosten *et al.*, 1994; Van den Ackerveken *et al.*, 1994;

Table 2. SCR proteins in plants and pathogens

Function	Plant SCR	reference	Receptor	Type	reference
Self-incompatibility	SCR/ SP11	(Shiba, 2001)	SRK	G-LecRK	(Takayama, 2001)
	PrsS	(Foote, 1994)	PrpS		(Wheeler, 2009)
Pollen tube germination and tip growth	LAT52 (Ole e I)	(Muschietti, 1994)	LePRKs	LRR-RLK	(Tang, 2002)
	LeSTIG	(Tang, 2004)	LePRKs	LRR-RLK	(Tang, 2002)
	SCA	(Park, 2000)			
	LTP5	(Chae, 2010)			
Micropylar pollen tube guidance	LURE	(Okuda, 2009)	MIK1/ MDIS1, PRK3 or PRK6	RLK	(Wang, 2016) (Takeuchi, 2016)
Pollen tube reception	RALFs	(Pearce, 2001)	FER	Malectin-like receptor kinase	(Haruta, 2014)
	ZmES1-4	(Amien, 2010)			
	PMEI1	(Woriedh, 2013)			
Gamete activation	EC1	(Sprunck, 2012)			
Function	Fungi/ oomycete SCR	Reference	Receptor	Type	Reference
Apoplastic effector	***Cladosporium fulvum***				
	Avr2	(Luderer, 2002)	Cf-2	RLP	(Dixon, 1996)
	Avr4	(Joosten, 1994)	Cf-4	RLP	(Thomas, 1997)
	Avr4E	(Westerink, 2004)	Cf-4E	RLP	(Takken, 1999)
	Avr5	(Mesarich, 2014)	Cf-5	RLP	(Dixon, 1998)
	Avr9	(van der Ackerveken, 1994)	Cf-9	RLP	(Jones, 1994)
	Ecp6	(de Jonge, 2010)			
	Phytophthora				
	PcF	(Orsomando, 2003)	Unknown		
	SCR74	(Liu, 2005)	Fine mapped to G-LecRK locus		(unpublished)
	SCR91	(Win, 2006)	Unknown		
	SCR96	(Chen, 2015)	Unknown		
	Fusarium oxysporum f. sp. lycopersici				
	Avr1 (Six4)	(Houterman, 2008)	Unknown		
	Avr3 (Six1)	(Rep, 2004)	I-3	G-LecRK	(Catanzariti, 2015)
	Avr7 (unkown)		I-7	RLP	(Cendales, 2016)
	Fungi RALFs				
	Ustilago maydis				
	Pep1	(Doehlemann, 2009)			

Luderer *et al.*, 2002; Westerink *et al.*, 2004; de Jonge *et al.*, 2010; Mesarich *et al.*, 2014), and their corresponding receptors are the RLP, Cf-2, Cf-4, Cf-4E, Cf-5 and Cf-9, respectively (Jones *et al.*, 1994; Dixon *et al.*, 1996; Thomas *et al.*, 1997; Dixon *et al.*, 1998; Takken *et al.*, 1999); In *Fusarium oxysporum* f. sp. *lycopersici*, SCR proteins including AVR1, AVR3 and AVR7 have been found (Rep *et al.*, 2004; Houterman *et al.*, 2008). The AVR3 receptor I-3 is encoded by a *G-LecRK* gene **(Table 2 and Figure 3)**.

SCR/SP11 from Brassicaceae and SCR74 from P. infestans

In *Brassica* species, the *S*-locus Cys-rich (SCR/SP11) gene encodes highly polymorphic SCR proteins that are specifically expressed in the pollen and localized at the surface of the pollen coat (Schopfer *et al.*, 1999; Takayama *et al.*, 2000; Shiba *et al.*, 2001). The receptor of SCR/SP11 is encoded by a *G-LecRK* gene, namely SRK **(Figure 3)**. It physically interacts

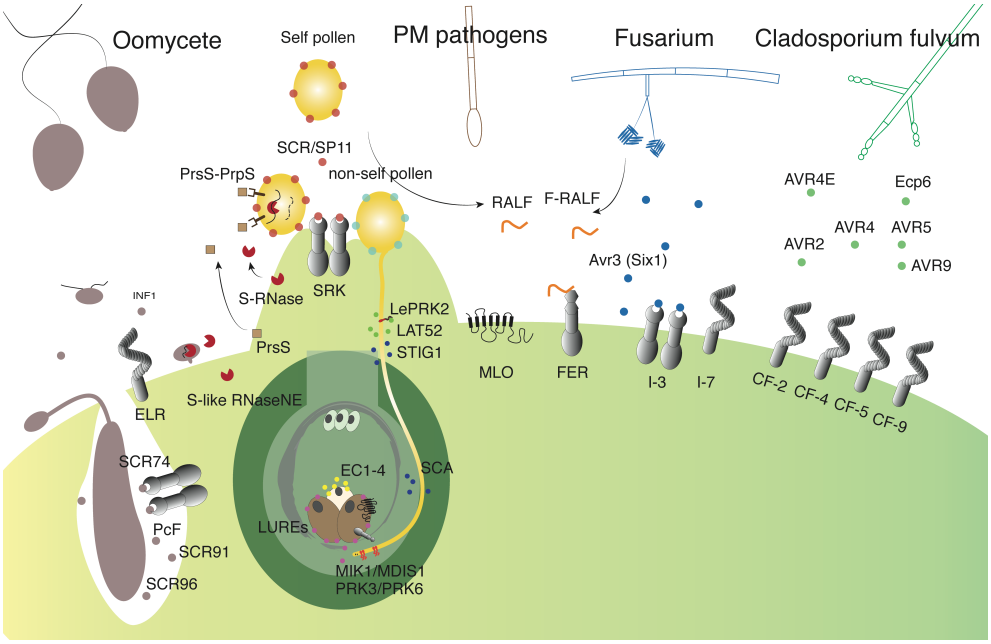


Figure 3. A combined illustration of the first layer of the plant immune system and the plant fertilization system.

Oomycetes secrete apoplastic effectors like SCR74, PcF, SCR91 and SCF96. The SCR74 receptor has been mapped to a *G-LecRK* locus. The growth and germination of *Phytophthora* zoospores can be inhibited by the S-like RNases (red) from plants. SCR/SP11 (red) is a small cysteine-rich protein on the pollen coat, it controls the self-incompatibility reaction in Brassicaceae plants. SRK (G-LecRK) on the stigma and the non-self SCR/SP11 (cyan) protein will not trigger the SI response. Many SCR proteins are involved in plant fertilization, like LAT52, STIG1, SCA, LURE, EC1-4, and some of the receptors have been found, like LePRK2, MIK1/MDIS1 and PRK3/PRK6. RALF and F-RALF are from plants and *Fusarium*, both of them can be recognized by FER. *Mlo* is a susceptibility gene facilitating plant pathogens. The *mlo* mutants showed the same phenotype as *fer* mutants, suggesting that it is also involved in the plant fertilization process. *Fusarium* secretes many SCR proteins into the apoplast, like AVR3, and a G-LecRK I-3 confer resistance when AVR3 is detected. I-7 is an RLP, but the corresponding *Avr* gene is yet unknown. In *Cladosporium fulvum*, many SCR proteins have been identified, like AVR2, AVR4, AVR4E, AVR5, AVR9 and Ecp6, their receptors are mainly RLP like CF-2, CF-4, CF-5 and CF-9.

with SCR/SP11 at a S-haplotype-specific manner. The recognition leads to the self-incompatibility reaction (Kachroo *et al.*, 2001; Takayama *et al.*, 2001; Chookajorn *et al.*, 2004). Genetically, *SRK* and *SCR/ SP11* are tightly linked (Boyes *et al.*, 1997). SCR/SP11 sequences from different haplotypes are highly diverse and under strong positive selection, as well as the hypervariable region of SRK (Sato *et al.*, 2002).

In *Phytophthora* species, many small SCR proteins act as apoplastic effectors (Kamoun, 2006) Some SCR proteins contain a PcF domain, and these are classified into the PcF/SCR family (**Chapter 2**). PcF was first found in *Phytophthora cactorum*, it triggers a cell death response in

the host plant strawberry, tomato and potato (Orsomando *et al.*, 2003) (Chapter 5). The structure of *PcF* is similar to the pollen allergen Ole e 6 of olive tree (*Olea europaca* L), reminiscent of a molecular mimic hypothesis (Nicastro *et al.*, 2009). Similarly, the SCR protein LAT52 of tomato is a homolog of Ole e 1, which is another allergen in pollen of the olive tree (Alché *et al.*, 2004). Another PcF domain containing SCR protein is the SCR74, which has been found in *P. infestans* as a multi-copy and highly polymorphic protein family that is under positive selection, and are presented in different strains (Liu *et al.*, 2005).

Table 3. Features of SCR/SP11 from Brassicaceae and PcF/SCR74 family in *Phytopthora*

	SCR/ SP11	**PcF/ SCR**
Origin	*Brassicaceae*	*Phytophthora*
Expression pattern	Pollen specific	Germinating cysts
Type	Small cysteine-rich	Small cysteine-rich
Length	< 160 aa	< 160 aa
Signal peptide	Yes	Yes
Molecular weight	≈8.4-kDal	≈5-kDal
Positive selection	Yes	Yes
Polymorphism	High	High
Number of cysteine	8	6-8
Localization	Apoplast	Apoplast
Receptor	SRK (G-LecRK)	Fine mapping to G-LecRK locus

SCR/SP11 from Brassicaceae and the PcF/SCR family from *Phytophthora* species share many common features **(Table 3)**. For example, they are both SCR proteins smaller than 160 amino acids (aa) and with 6-8 cysteines, both of them have a signal peptide and are localized in the apoplastic space **(Figure 4)** and display tissue-specific expression, in pollen and germinating cysts, respectively. Both SCR proteins are highly diverse and under positive selection. The receptor of SCR/ SP11 is a G-LecRK protein SRK, and the SCR74 receptor has been fine mapped to a *G-LecRK* locus **(Chapter 4)**.

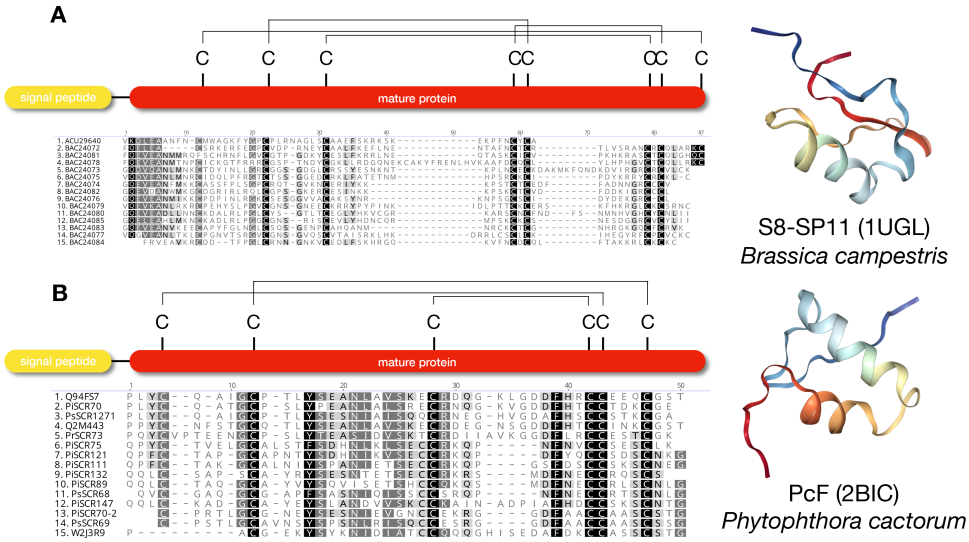


Figure 4. Molecular comparison of SCR/SP11 from Brassicaceae and PcF/SCR from *Phytopthora* species.

For both families, the proteins are highly diverse, except for the presence of several cysteines, the structure is shown in the right panels. **A.** SCR/ SP11 proteins from *Brassicaeae* and the structure of S8-SP11 from *Brassica campestris*. **B.** PcF/ SCR effector family from *Phytophthora* species, the structure of PcF from *P. cactorum* is shown.

RALF, FERONIA (FER) and Nortia (Mlo)

FERONIA (FER) is a RLK kinase with a malectin-like domain, which is specifically expressed in synergid cells from *Arabidopsis thaliana*. FER is found to regulate pollen reception and may be involved in breaking reproduction barriers. In *fer* mutants, pollen tubes continue to grow after entering the synergid cells, and do not release the sperm cell (Escobar-Restrepo *et al.*, 2007). Later, a secreted peptide hormone, rapid alkalinization factor (RALF), was found to interact directly with FER, leading to suppression of cell elongation (Haruta *et al.*, 2014). Surprisingly, the plant alkalinizing peptides were also found in the fungal pathogen *Fusarium oxyporum*, namely F-RALF, that can regulate the pH of the host apoplast and facilitate infection. The plant receptor FER might act as a negative factor of F-RALF. *Arabidopsis fer* mutants showed enhanced resistance against *Fusarium* (Masachis *et al.*, 2016). Furthermore, FER was found to be a scaffold of the surface immune receptor complex of FLS2 and BAK1, to initiate plant immune signaling (Stegmann *et al.*, 2017) (**Figure 3)**.

Mildew resistance locus o (*Mlo*) is another example, it is a well-known "susceptibility gene", the loss of function mutants causes powdery mildew resistance. It was first cloned from barley, and the loss of function mutations of *Mlo* led to durable resistance to powdery mildew *Blumeria graminis* f.sp. *hordei* (Bgh). After inoculation of a *mlo* mutant, the interaction with powdery mildew is stopped at the very beginning, already during cell wall penetration (Jørgensen, 1992; Büschges *et al.*, 1997). *Mlo* genes have several homologs in many plant genomes, and phylogenetic analysis indicated that this family originated from green algae (Kusch *et al.*, 2016) **(Figure 3)**. In literature it is shown that *mlo* causes resistance to other mildews in crop plants like tomato, pea and grape (Bai *et al.*, 2007; Humphry *et al.*, 2011; Pessina *et al.*, 2016). Further studies revealed that the function of *Mlo* is not restricted to plant immunity. In *Arabidopsis*, the *Atmlo7* mutation showed pollen tube overgrowth in the synergid cells, and the fertility was reduced. This phenotype is similar with the *Feronia* (*Fer*) mutation mentioned above. Surprisingly, the *fer/fer* mutant in *Arabidopsis* also showed powdery mildew resistance (Kessler *et al.*, 2010). Taken together, FER and MLO provide additional examples of a direct link between the plant fertilization and immunity systems.

S-RNase in Solanaceae

In some genera of Solanaceae, Rosaceae and Plantaginaceae, the SI is gametophytic and controlled by female S-RNase. The RNA of self-pollen will be degraded, but the non-self-pollen carrying the S locus F-box (*SLF*) gene can detoxify the specific S-RNase (Lai *et al.*, 2002). Similar like other genes involved in fertilization of flowering plants, the S-Rnases have been believed to protect the nutrient-rich and vulnerable pistil from pathogens (Dickinson, 1994). In tobacco, a S-like RNase NE was found to control the pathogen growth, when inoculated with *Phytophthora parasitica* (Galiana *et al.*, 1997). A further study showed that purified RNase NE has antimicrobial activity against two plant pathogens, *P. parasitica* and *F. oxysporum* (Hugot *et al.*, 2007) **(Figure 3)**. Additionally, a recent study reported that the host can deliver small RNAs (sRNAs) into pathogen cell, to silence pathogen genes that are important for pathogenicity (Cai *et al.*, 2018).

Conclusions and perspectives

Better understanding of the plant fertilization and SI system are important for seed production and breeding for yield. For example, although the SI seems to be a better strategy in nature to keep the genetic diversity as broad as possible, sometimes humans need to break or circumvent the SI to develop homozygous parental lines, for breeding hybrid varieties with higher yields. Lectures from the ancient immune system would help to develop a deeper insight. And due to the fast evolving genetic variation of plant pathogens, it is important to know how to achieve durable resistance, which still is a puzzle to be solved. By studying the similarity between the two systems at the molecular level, we found various molecules playing important roles in both systems, like the SCR proteins, RLK with different extracellular domains like LRR, G-lectin and malectin-like, as well as S-Rnases and MLO. Upon future cloning and characterization of additional plant surface receptors, more evidence is expected to be found to reveal the mechanisms of both systems leading to plant fertilization and immunity.

III. Cloning plant *R* genes in the "Omic" era

With the fast evolution of sequencing technology, mapping-by-sequencing strategies are now widely used in many model plants and crops. In **Chapter 3**, we developed the RLP/KSeq for fast mapping plant surface immune receptors genes, and successfully mapped the INF1 and SCR74 receptors on chromosome 12 and 9, respectively. In **Chapter 5**, we first applied BSA-RNAseq on potato, and quickly mapped the Pep-25 receptor on chromosome 3 in a segregating population. The BSA-RNAseq markers further allow fine mapping of the Pep-25 receptor in an expanded population.

In **Table 4**, we summarize several "mapping-by-sequencing" methods that have been used for mapping and cloning of plant disease resistance genes, like genomic resequencing, Indel-Seq, BSA-RNAseq, MutChromSeq, targeted chromosome-based cloning via long-range assembly (TACCA) (Takagi *et al.*, 2013; Ramirez Gonzalez *et al.*, 2015; Sánchez-Martín *et al.*, 2016; Singh *et al.*, 2017; Thind *et al.*, 2017), and the RenSeq series including RenSeq, dRenSeq, AgRenSeq, RLP/KSeq, SMRT-RenSeq and MutRenSeq (Jupe *et al.*, 2013; Steuernagel *et al.*,

2016; Van Weymers *et al.*, 2016; Witek *et al.*, 2016; Arora *et al.*, 2018) **(Table 4)** (**Chapter 3** and **Chapter 5**).

Table 4. NGS-based methods for identification of plant resistance genes.

Methods	Purpose	Crop	Example	Population type	Population size	Reference
Whole genome/ transcriptome						
Genomic reseuencing	Fast mapping	Rice	Rice blast resistant gene	F2 segregating population	40	(Takagi, 2013)
Indel-Deq	Fast mapping	Pigeopea	fusarium wilt and sterility mosaic disease	F2/ RIL	40	(Singh, 2017)
BSA-RNAseq	Fast mapping	Wheat	yellow rust	F2 segregating population	100	(RamirezGonzalez, 2015)
BSA-RNAseq	Fast mapping	Potato	Pep-25 receptor	F1 segregating population	68	**(Chapter 5, this thesis)**
Complexity reduction approach						
MutChromSeq	Clone *NLR* genes	Wheat	powdery mildew resistance gene 2 (*Pm2*)	M3 plants	6	(SanchezMartin, 2016)
TACCA	Clone *NLR* genes	Wheat	leaf-rust resistance gene Lr22a	F2 segregating population and EMS mutants	1,656	(Thind, 2017)
RenSeq	Fast mapping NLR genes	Potato	late blight *Rpi-ber2* and *Rpi-rzc1*	F1 segregating population	50-100	(Jupe, 2013)
dRenSeq	Diagnostic known *R* genes	Potato	functional alleles of *Rpi-vnt1*	n/a	n/a	(VanWeymers, 2016)
AgRenSeq	Associate mapping	Wheat	Sr33, Sr45 and SrTA1662	Diversity panel	174	(Arora, 2018)
RLP/KSeq	Fast mapping	Potato	ELR, SCR74 receptor	F1 segregating population	50	**(Chapter 3, this thesis)**
SMRT-RenSeq	Clone NLR genes	Potato	*Rpi-amr3i*	F2	100	(Witek, 2016)
MutRenSeq	Clone NLR genes	Wheat	*Sr22* and *Sr45*	M2	1,300	(Steuernagei, 2016)

Here, we compare the different methods, and describe four aspects towards cloning new *R* genes. 1). Identification of a resistance source; 2). Methods for cloning the *R* gene without generating segregating populations; 3). Fast mapping the *R* gene in a segregating population by NGS; 4). Fine mapping and cloning *R* genes by long-reads sequencing.

The first step is to find the source of resistance **(Figure 5A)**. Traditionally, the resistant plants can be found in field trials, or by disease testing in the laboratory. However, in many cases, the disease test can only detect the *R* genes that confer a strong, qualitative resistance. And, if multiple *R* genes are present in the same plant, or if the *R* genes confer mild, quantitative levels of resistance, multiple crossings need to be performed to genetically dissect the resistance. In our case, effectoromics-based methods offer another route to detect the *R* genes or surface immune receptors (Vleeshouwers *et al.*, 2008), regardless whether they confer high of low levels of resistance, the qualitative phenotyping for effector responses enables identifying the matching genes from one single plant genotype in a single step. In **Chapter 3**, we successfully mapped two surface immune receptors in one single population. Up to now, many plant pathogen genomes have been sequenced, and the effectors can be predicted

(Sperschneider *et al.*, 2017; Sperschneider *et al.*, 2018). Therefore effectoromics can offer a superior option to detect immune receptors in plants **(Figure 5B)**.

Once the resistance/susceptible genotypes have been detected, mutagenesis can be considered, provided that the *R* genes are from homozygous genotypes **(Figure 5B)**. Homozygous loss-of-function mutations can be selected in M2 generations and the causal genes can be detected directly by MutRenSeq or MutChromSeq, even in plants with relatively big and complex genomes like wheat (Sánchez-Martín *et al.*, 2016; Steuernagel *et al.*, 2016). The MutChromSeq method depends on chromosome flow sorting, however, optimization is needed when applied to other crops. MutRenSeq is more simple, it can dramatically reduce the genomic complexity, and the generating of RNA bait library can be outsourced.

A recently developed association genetics with *R* gene enrichment sequencing (AgRenSeq) allows to identify the *R* genes directly from the diversity panel, without generating any segregating populations or mutants. Unlike other methods, AgRenSeq used k-mers instead of mapping all the reads to reference genome, as a proof of concept, this method was used to clone four stem rust *R* genes within six months (Arora *et al.*, 2018) **(Table 4, Figure 5)**.

Diagnostic RenSeq (dRenSeq) can be applied if the identified resistance is based on a known *R* gene or homolog **(Figure 5B)** (Van Weymers *et al.*, 2016). Traditional map-based cloning is still powerful to generate a refined genetic map. Once the segregating population is ready, RenSeq, RLP/KSeq, BSA-RNAseq or genomic resequencing can be used for fast mapping of the candidate genes. If the phenotype is determined by effectoromics, the subsequent enrichment sequencing methods can be adapted based on the type of effectors: for cytoplasmic RXLR effectors, normal RenSeq can be applied to detect *NLR* genes, but for apoplastic effectors, RLP/KSeq is adapted to detect the *RLP/RLK* genes that represent the surface immune receptors **(Figure 5C)**. The fast mapping can also be done by BSA-RNAseq, by comparing the infected/un-infected samples/pools, the expression information can further facilitate to detect the candidate genes (**Chapter 5**). The latter strategy is slightly less biased, since also RLP/RLK that do not contain an eLRR are more likely to be detected, than with the current *RLP/RLK* library (**Chapter 3**).

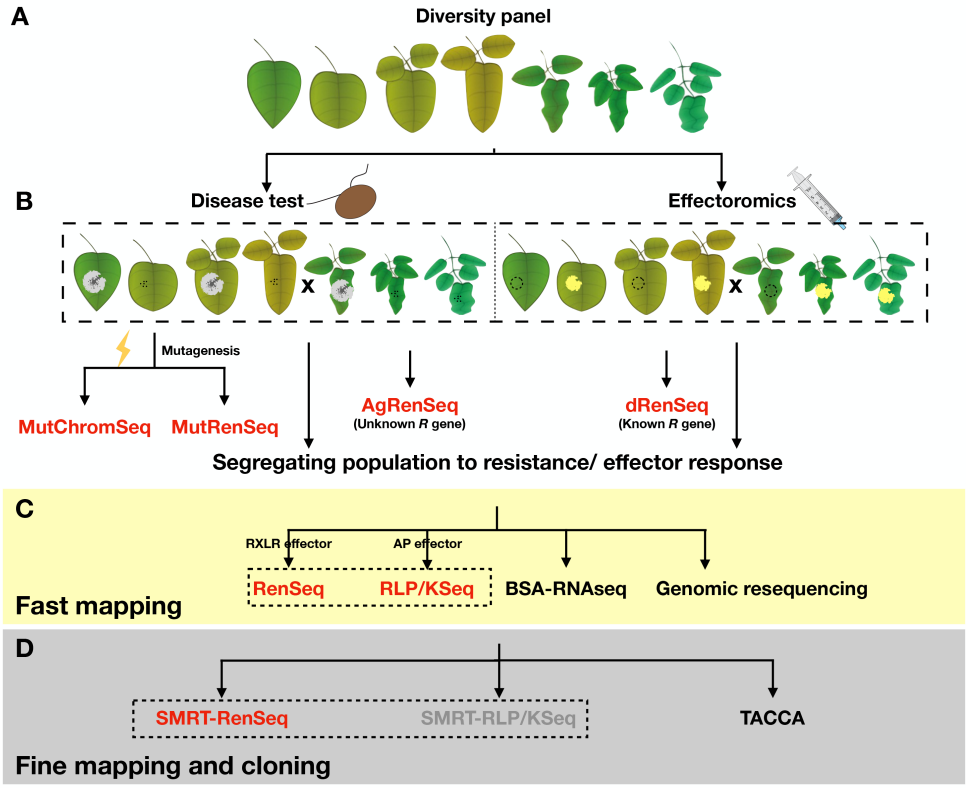


Figure 5. Decision tree of the cutting edge mapping by sequencing methods for cloning plant *R* genes.

The markers obtained from those methods can only be used for fast mapping the candidate gene onto a chromosome. However, for fine mapping and cloning the candidate gene, expanding the mapping population, generating a BAC library and isolating positive clones is still necessary, which is normally expensive and time-consuming. SMRT-RenSeq was developed to speed up fine mapping without generating a large population or generate a BAC library **(Figure 5D)**. By combing with single-molecule real-time (SMRT) sequencing, the full *NLR* genes can be assembled directly, and other short RenSeq reads from the R/ S parents or bulked progenies can be mapped to the *NLR* assemblies. Using this strategy, a late blight resistance gene *Rpi-amr3i* was successfully cloned from *Solanum americanum* (Witek *et al.*, 2016).

The traditional RenSeq method can only enrich the *NLR* genes, but since surface immune receptors also confer resistance, we suggest combining the *NLR* gene library and the *RLP/RLK* library preparation when the candidate gene is unknown **(Chapter 3) (Figure 5C and 5D)**.

In summary, with the fast development of the next and third generation sequencing, genotyping-by-sequencing (GBS) has already become a routine method in gene cloning projects, here we reviewed the current methods and provide a practical guide for selecting the best method for different projects.

Final remarks

To feed the increasing world population in a sustainable manner, it is important to develop crops with durable resistance. In this thesis, we explored the first layer of plant immune system by studying the MAMPs/PAMPs and apoplastic effectors from *Phytophthora* and their receptors in potato. Methodologically, we showed that the combination of functional genomics (effectoromics) and NGS-based methods (RLP/Kseq and BSA-RNAseq) enables fast identifying and mapping of potato surface immune receptors. The methodology can be applied on other crops, and is especially advised for plants with large and complex genomes. Scientifically, our findings expand current understanding of effector biology and their evolution. We reveal that the members of a same effector family can evolve to distinct directions, and act as MAMPs/PAMPs or effectors, during the co-evolution with their hosts. Furthermore, our work is expected to lead to identification of two new surface immune receptors in potato. This will deepen our knowledge about the molecular interaction between oomycetes and plants in the apoplastic space, and potentially contribute to achieve more durable resistance to potato late blight. In the end, we discussed the similarity between the plant immune and fertilization systems, we reviewed the latest knowledge of both these systems and propose that the plant fertilization system has evolved from the ancient plant immune system.

References

Albert I, Böhm H, Albert M, Feiler CE, Imkampe J, Wallmeroth N, Brancato C, Raaymakers TM, Oome S, Zhang H, et al (2015) An RLP23–SOBIR1–BAK1 complex mediates NLP-triggered immunity. NPLANTS **1**: 15140

Alché J de D, M'rani-Alaoui M, Castro AJ, Rodríguez-García MI (2004) Ole e 1, the Major Allergen from Olive (*Olea europaea* L.) Pollen, Increases its Expression and is Released to the Culture Medium during in vitro Germination. Plant and Cell Physiology **45**: 1149–1157

Amien S, Kliwer I, Márton ML, Debener T, Geiger D, Becker D, Dresselhaus T (2010) Defensin-like ZmES4 mediates pollen tube burst in maize via opening of the potassium channel KZM1. PLOS Biology **8**: e1000388

Arora S, Steuernagel B, Chandramohan S, Long Y, Matny O, Johnson R, Enk J, Periyannan S, Hatta MAM, Athiyannan N, et al (2018) Resistance gene discovery and cloning by sequence capture and association genetics. bioRxiv 248146

Bai Y, Pavan S, Zheng Z, Zappel NF, REINSTÄDLER A, Lotti C, De Giovanni C, Ricciardi L, Lindhout P, Visser R, et al (2007) Naturally occurring broad-spectrum powdery mildew resistance in a central american tomato accession is caused by loss of *Mlo* function. http://dxdoiorg/101094/MPMI-21-1-0030. doi: 10.1094/MPMI-21-1-0030

Balandín M, Royo J, Gómez E, Muniz LM, Molina A, Hueros G (2005) A protective role for the embryo surrounding region of the maize endosperm, as evidenced by the characterisation of *ZmESR-6*, a defensin gene specifically expressed in this region. Plant Mol Biol **58**: 269–282

Boyes DC, Nasrallah ME, Vrebalov J, Nasrallah JB (1997) The self-incompatibility (S) haplotypes of *Brassica* contain highly divergent and rearranged sequences of ancient origin. The Plant Cell **9**: 237–247

Büschges R, Hollricher K, Panstruga R, Simons G, Wolter M, Frijters A, van Daelen R, van der Lee T, Diergaarde P, Groenendijk J, et al (1997) The barley *Mlo* gene: a novel control element of plant pathogen resistance. Cell **88**: 695–705

Cai R, Lewis J, Yan S, Liu H, Clarke CR, Campanile F, Almeida NF, Studholme DJ, Lindeberg M, Schneider D, et al (2011) The plant pathogen P*seudomonas syringae* pv. *tomato* is genetically monomorphic and under strong selection to evade tomato immunity. PLoS Pathog **7**: e1002130

Cai Q, Qiao L, Wang M, He B, Lin F-M, Palmquist J, Huang S-D, Jin H (2018) Plants send small RNAs in extracellular vesicles to fungal pathogen to silence virulence genes. Science **360**: 1126–1129

Chae K, Gonong BJ, Kim S-C, Kieslich CA, Morikis D, Balasubramanian S, Lord EM (2010) A multifaceted study of stigma/style cysteine-rich adhesin (SCA)-like *Arabidopsis* lipid transfer proteins (LTPs) suggests diversified roles for these LTPs in plant growth and reproduction. Journal of Experimental Botany **61**: 4277–4290

Chookajorn T, Kachroo A, Ripoll DR, Clark AG, Nasrallah JB (2004) Specificity determinants and diversification of the *Brassica* self-incompatibility pollen ligand. Proceedings of the National Academy of Sciences **101**: 911–917

Cook DE, Mesarich CH, Thomma BPHJ (2015) Understanding plant immunity as a surveillance system to detect invasion. Annu Rev Phytopathol **53**: 541–563

de Jonge R, van Esse HP, Kombrink A, Shinya T, Desaki Y, Bours R, van der Krol S, Shibuya N, Joosten MHAJ, Thomma BPHJ (2010) Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. Science **329**: 953–955

Dickinson H (1994) Simply a social disease? Nature **367**: 517–518

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

Dixon MS, Jones DA, Keddie JS, Thomas CM, Harrison K, Jones JD (1996) The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. Cell **84**: 451–459

Doehlemann G, van der Linde K, Aßmann D, Schwammbach D, Hof A, Mohanty A, Jackson D, Kahmann R (2009) Pep1, a secreted effector protein of *Ustilago maydis*, is required for successful invasion of plant cells. PLoS Pathog **5**: e1000290

Domazakis E, Wouters D, Visser RGF, Kamoun S, Joosten MHAJ, Vleeshouwers VGAA (2018) The ELR-SOBIR1 complex functions as a two-component receptor-like kinase to mount defense against *Phytophthora infestans*. MPMI. doi: 10.1094/MPMI-09-17-0217-R

Du J, Verzaux E, Chaparro-Garcia A, Bijsterbosch G, Keizer LCP, Zhou J, Liebrand TWH, Xie C, Govers F, Robatzek S, et al (2015) Elicitin recognition confers enhanced resistance to *Phytophthora infestans* in potato. NPLANTS 10.1038–nplants.2015.34

Escobar-Restrepo J-M, Huck N, Kessler S, Gagliardini V, Gheyselinck J, Yang W-C, Grossniklaus U (2007) The FERONIA Receptor-like kinase mediates male-female interactions during pollen tube reception. Science **317**: 656–660

Foote HC, Ride JP, Franklin-Tong VE, Walker EA, Lawrence MJ, Franklin FC (1994) Cloning and expression of a distinctive class of self-incompatibility (S) gene from *Papaver rhoeas* L. Proceedings of the National Academy of Sciences **91**: 2265–2269

Fujii S, Kubo K-I, Takayama S (2016) Non-self- and self-recognition models in plant self-incompatibility. NPLANTS **2**: 16130

Galiana E, Bonnet P, Conrod S, Keller H, Panabieres F, Ponchet M, Poupet A, Ricci P (1997) RNase activity prevents the growth of a fungal pathogen in Tobacco leaves and increases upon induction of systemic acquired resistance with Elicitin. PLANT PHYSIOLOGY **115**: 1557–1567

Gassmann W, Dahlbeck D, Chesnokova O, Minsavage GV, Jones JB, Staskawicz BJ (2000) Molecular evolution of virulence in natural field strains of *Xanthomonas campestris* pv. *vesicatoria*. J Bacteriol **182**: 7053–7059

Gómez-Gómez L, Boller T (2000) FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. Molecular Cell **5**: 1003–1011

Haruta M, Sabat G, Stecker K, Minkoff BB, Sussman MR (2014) A peptide hormone and its receptor protein kinase regulate plant cell expansion. Science **343**: 408–411

Hind SR, Strickler SR, Boyle PC, Dunham DM, Bao Z, O'Doherty IM, Baccile JA, Hoki JS, Viox EG, Clarke CR, et al (2016) Tomato receptor FLAGELLIN-SENSING 3 binds flgII-28 and activates the plant immune system. NPLANTS **2**: 16128

Hodgkin T, Lyon GD, Dickinson HG (1988) Recognition in flowering plants: A comparison of the Brassica self-incompatibility system and plant pathogen interactions. New Phytol **110**: 557–569

Houterman PM, Ben J C Cornelissen, Rep M (2008) Suppression of plant resistance gene-based immunity by a fungal effector. PLoS Pathog **4**: e1000061

Hugot K, Ponchet M, Marais A, Ricci P, Galiana E (2007) A Tobacco S-like RNase inhibits hyphal elongation of plant pathogens. MPMI. doi: 10.1094/MPMI.2002.15.3.243

HUMPHRY M, REINSTÄDLER A, IVANOV S, BISSELING T, Panstruga R (2011) Durable broad-spectrum powdery mildew resistance in pea er1 plants is conferred by natural loss-of-function mutations in PsMLO1. Molecular Plant Pathology **12**: 866–878

Igic B, Lande R, Kohn JR (2008) Loss of self-incompatibility and its evolutionary consequences. Int J Plant Sci **169**: 93–104

Jo K-R, Kim C-J, Kim S-J, Kim T-Y, Bergervoet M, Jongsma MA, Visser RG, Jacobsen E, Vossen JH (2014) Development of late blight resistant potatoes by cisgene stacking. BMC Biotechnology **14**: 50

Jones DA, Thomas CM, Hammond-Kosack KE, Balint-Kurti PJ, Jones JD (1994) Isolation of the tomato *Cf-9* gene for resistance to C*ladosporium fulvum* by transposon tagging. Science **266**: 789–793

Jones JDG, Dangl JL (2006) The plant immune system. Nature **444**: 323–329

Joosten MHAJ, Cozijnsen TJ, de Wit PJGM (1994) Host resistance to a fungal tomato pathogen lost by a single base-pair change in an avirulence gene. Nature **367**: 384

Jupe F, Witek K, Verweij W, Śliwka 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

Jørgensen IH (1992) Discovery, characterization and exploitation of Mlo powdery mildew resistance in barley. Euphytica **63**: 141–152

Kachroo A, Schopfer CR, Nasrallah ME, Nasrallah JB (2001) Allele-specific receptor-ligand interactions in *Brassica* self-incompatibility. Science **293**: 1824–1826

Kamoun S (2001) Nonhost resistance to *Phytophthora*: novel prospects for a classical problem. Current Opinion in Plant Biology **4**: 295–300

Kamoun S (2006) A catalogue of the effector secretome of plant pathogenic oomycetes. Annu Rev Phytopathol **44**: 41–60

Kessler SA, Shimosato-Asano H, Keinath NF, Wuest SE, Ingram G, Panstruga R, Grossniklaus U (2010) Conserved molecular components for pollen tube reception and fungal invasion. Science **330**: 968–971

Kubo K-I, Entani T, Takara A, Wang N, Fields AM, Hua Z, Toyoda M, Kawashima S-I, Ando T, Isogai A, et al (2010) Collaborative non-self recognition system in S-RNase–based self-incompatibility. Science **330**: 796–799

Kunwar S, Iriarte F, Fan Q, da Silva EE, Ritchie L, Nguyen NS, Freeman JH, Stall RE, Jones JB, Minsavage GV, et al (2018) Transgenic expression of *EFR* and *Bs2* genes for field management of bacterial wilt and bacterial spot of tomato. Phytopathology PHYTO–12–17–0424–R–55

Kusch S, Pesch L, Panstruga R (2016) Comprehensive phylogenetic analysis sheds light on the diversity and origin of the MLO family of integral membrane proteins. Genome Biology and Evolution **8**: 878–895

Lacombe S, Rougon-Cardoso A, Sherwood E, Peeters N, Dahlbeck D, van Esse HP, Smoker M, Rallapalli G, Thomma BPHJ, Staskawicz B, et al (2010) Interfamily transfer of a plant pattern-recognition receptor confers broad-spectrum bacterial resistance. Nature Biotechnology **28**: 365

Lai Z, Ma W, Bin Han, Liang L, Zhang Y, Hong G, Xue Y (2002) An F-box gene linked to the self-incompatibility *S* locus of *Antirrhinum* is expressed specifically in pollen and tapetum. Plant Mol Biol **50**: 29–41

Liu Y, Wu H, Chen H, Liu Y, He J, Kang H, Sun Z, Pan G, Wang Q, Hu J, et al (2015) A gene cluster encoding lectin receptor kinases confers broad-spectrum and durable insect resistance in rice. Nature Biotechnology **33**: 301–216

Liu Z, 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 *scr74* gene family of *Phytophthora infestans*. Mol Biol Evol **22**: 659–672

Luderer R, Takken FLW, de Wit PJGM, Joosten MHAJ (2002) *Cladosporium fulvum* overcomes *Cf-2*-mediated resistance by producing truncated AVR2 elicitor proteins. Molecular Microbiology **45**: 875–884

Marshall E, Costa LM, Gutierrez-Marcos J (2011) Cysteine-Rich Peptides (CRPs) mediate diverse aspects of cell–cell communication in plant reproduction and development. Journal of Experimental Botany **62**: 1677–1686

Masachis S, Segorbe D, Turrà D, Leon-Ruiz M, Fürst U, Ghalid El M, Leonard G, López-Berges MS, Richards TA, Felix G, et al (2016) A fungal pathogen secretes plant alkalinizing peptides to increase infection. Nature Microbiology 2016 1:6 **1**: 16043

Mesarich CH, Griffiths SA, van der Burgt A, Ökmen B, Beenen HG, Etalo DW, Joosten MHAJ, de Wit PJGM (2014) Transcriptome sequencing uncovers the *Avr5* avirulence gene of the tomato leaf mold pathogen *Cladosporium fulvum*. MPMI. doi: 10.1094/MPMI-02-14-0050-R

Mueller AN, Ziemann S, Treitschke S, Aßmann D, Doehlemann G (2013) Compatibility in the *Ustilago maydis*–Maize interaction requires inhibition of host cysteine proteases by the fungal effector Pit2. PLoS Pathog **9**: e1003177

Murfett J, Atherton TL, Mou B, Gassert CS, McClure BA (1994) S-RNase expressed in transgenic *Nicotiana* causes S-allele-specific pollen rejection. Nature **367**: 563

Muschietti J, Dircks L, Vancanneyt G, McCormick S (1994) LAT52 protein is essential for tomato pollen development: pollen expressing antisense LAT52 RNA hydrates and germinates abnormally and cannot achieve fertilization. The Plant Journal **6**: 321–338

Navarro L (2004) The Transcriptional Innate Immune Response to flg22. Interplay and Overlap with Avr Gene-Dependent Defense Responses and Bacterial Pathogenesis. PLANT PHYSIOLOGY **135**: 1113–1128

Nicastro G, Orsomando G, Ferrari E, Manconi L, Desario F, Amici A, Naso A, Carpaneto A, Pertinhez TA, Ruggieri S, et al (2009) Solution structure of the phytotoxic protein PcF: The first characterized member of the *Phytophthora* PcF toxin family. Protein Sci **18**: 1786–1791

Okuda S, Tsutsui H, Shiina K, Sprunck S, Takeuchi H, Yui R, Kasahara RD, Hamamura Y, Mizukami A, Susaki D, et al (2009) Defensin-like polypeptide LUREs are pollen tube attractants secreted from synergid cells. Nature **458**: 357

Orsomando G, Lorenzi M, Ferrari E, de Chiara C, Spisni A, Ruggieri S (2003) PcF protein from *Phytophthora cactorum* and its recombinant homologue elicit phenylalanine ammonia lyase activation in tomato. CMLS, Cell Mol Life Sci **60**: 1470–1476

Park S-Y, Jauh G-Y, Mollet J-C, Eckard KJ, Nothnagel EA, Walling LL, Lord EM (2000) A lipid transfer–like protein Is necessary for Lily pollen tube adhesion to an in vitro stylar matrix. The Plant Cell **12**: 151–163

Pearce G, Moura DS, Stratmann J, Ryan CA (2001) RALF, a 5-kDa ubiquitous polypeptide in plants, arrests root growth and development. Proceedings of the National Academy of Sciences **98**: 12843–12847

Pessina S, Lenzi L, Perazzolli M, Campa M, Costa LD, Urso S, Valè G, Salamini F, Velasco R, Malnoy M (2016) Knockdown of *MLO* genes reduces susceptibility to powdery mildew in grapevine. Horticulture Research, Published online: 29 October 2014; | doi:101038/hortres201454 **3**: 16016

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 J **13**: 613–624

Rep M, van der Does HC, Meijer M, Van Wijk R, Houterman PM, Dekker HL, De Koster CG, Cornelissen BJC (2004) A small, cysteine-rich protein secreted by *Fusarium oxysporum* during colonization of xylem vessels is required for *I-3*-mediated resistance in tomato. Molecular Microbiology **53**: 1373–1383

Sagaram US, Pandurangi R, Kaur J, Smith TJ, Shah DM (2011) Structure-activity determinants in antifungal plant defensins MsDef1 and MtDef4 with different modes of action against *Fusarium graminearum*. PLoS ONE **6**: e18550

Sato K, Nishio T, Kimura R, Kusaba M, Suzuki T, Hatakeyama K, Ockendon DJ, Satta Y (2002) Coevolution of the S-locus genes *SRK*, *SLG* and *SP11/SCR* in *Brassica oleracea* and *B. rapa*. Genetics **162**: 931–940

Sánchez-Martín J, Steuernagel B, Ghosh S, Herren G, Hurni S, Adamski N, Vrána J, Kubaláková M, Krattinger SG, Wicker T, et al (2016) Rapid gene isolation in barley and wheat by mutant chromosome sequencing. Genome Biology **17**: 221

Schopfer CR, Nasrallah ME, Nasrallah JB (1999) The male determinant of self-incompatibility in *Brassica*. Science **286**: 1697–1700

Shiba H, Takayama S, Iwano M, Shimosato H, Funato M, Nakagawa T, Che FS, Suzuki G, Watanabe M, Hinata K, et al (2001) A pollen coat protein, SP11/SCR, determines the pollen S-specificity in the self-incompatibility of *Brassica* species. PLANT PHYSIOLOGY **125**: 2095–2103

Singh VK, Khan AW, Saxena RK, Sinha P, Kale SM, Parupalli S, Kumar V, Chitikineni A, Vechalapu S, Sameer Kumar CV, 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 J **15**: 906–914

Sperschneider J, Dodds PN, Gardiner DM, Singh KB, Taylor JM (2018) Improved prediction of fungal effector proteins from secretomes with EffectorP 2.0. Molecular Plant Pathology **5**: e1000261–27

Sperschneider J, Dodds PN, Singh KB, Taylor JM (2017) ApoplastP: prediction of effectors and plant proteins in the apoplast using machine learning. New Phytol **217**: 1764–1778

Sprunck S, Rademacher S, Vogler F, Gheyselinck J, Grossniklaus U, Dresselhaus T (2012) Egg cell–secreted EC1 triggers sperm cell activation during double fertilization. Science **338**: 1093–1097

Stegmann M, Monaghan J, Smakowska-Luzan E, Rovenich H, Lehner A, Holton N, Belkhadir Y, Zipfel C (2017) The receptor kinase FER is a RALF-regulated scaffold controlling plant immune signaling. Science **355**: 287–289

Steuernagel B, Periyannan SK, Hernández-Pinzón I, Witek K, Rouse MN, Yu G, 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

Tai TH, Dahlbeck D, Clark ET, Gajiwala P, Pasion R, Whalen MC, Stall RE, Staskawicz BJ (1999) Expression of the *Bs2* pepper gene confers resistance to bacterial spot disease in tomato. Proceedings of the National Academy of Sciences **96**: 14153–14158

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

Takasaki T, Hatakeyama K, Suzuki G, Watanabe M, Isogai A, Hinata K (2000) The S receptor kinase determines self-incompatibility in *Brassica* stigma. Nature **403**: 913–916

Takayama S, Isogai A (2005) Self-incompatibility in plants. Annu Rev Plant Biol **56**: 467–489

Takayama S, Shiba H, Iwano M, Shimosato H, Che FS, Kai N, Watanabe M, Suzuki G, Hinata K, Isogai A (2000) The pollen determinant of self-incompatibility in *Brassica campestris*. Proceedings of the National Academy of Sciences **97**: 1920–1925

Takayama S, Shimosato H, Shiba H, Funato M, Che F-S, Watanabe M, Iwano M, Isogai A (2001) Direct ligand-receptor complex interaction controls *Brassica* self-incompatibility. Nature **413**: 534–538

Takebayashi N, Morrell PL (2001) Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. American Journal of Botany **88**: 1143–1150

Takeuchi H, Higashiyama T (2016) Tip-localized receptors control pollen tube growth and LURE sensing in *Arabidopsis*. Nature **531**: 245–248

Takken FLW, Thomas CM, Joosten MHAJ, Golstein C, Westerink N, Hille J, Nijkamp HJJ, de Wit PJGM, Jones JDG (1999) A second gene at the tomato *Cf-4* locus confers resistance to *Cladosporium fulvum* through recognition of a novel avirulence determinant. The Plant Journal **20**: 279–288

Tang W, Ezcurra I, Muschietti J, McCormick S (2002) A cysteine-rich extracellular protein, LAT52, interacts with the extracellular domain of the pollen receptor kinase LePRK2. The Plant Cell **14**: 2277–2287

Tang W, Kelley D, Ezcurra I, Cotter R, McCormick S (2004) LeSTIG1, an extracellular binding partner for the pollen receptor kinases LePRK1 and LePRK2, promotes pollen tube growth in vitro. The Plant Journal **39**: 343–353

Terras FR, Schoofs HM, De Bolle MF, Van Leuven F, Rees SB, Vanderleyden J, Cammue BP, Broekaert WF (1992) Analysis of two novel classes of plant antifungal proteins from radish (*Raphanus sativus* L.) seeds. Journal of Biological Chemistry **267**: 15301–15309

Thind AK, Wicker T, Šimková H, Fossati D, Moullet O, Brabant C, Vrána J, Doležel J, Krattinger SG (2017) Rapid cloning of genes in hexaploid wheat using cultivar-specific long-range chromosome assembly. Nature Biotechnology **35**: 793–796

Thomas CM, Jones DA, Parniske M, Harrison K, Balint-Kurti PJ, Hatzixanthis K, Jones JD (1997) Characterization of the tomato Cf-4 gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in Cf-4 and Cf-9. The Plant Cell **9**: 2209–2224

Thomma BPHJ, Nürnberger T, Joosten MHAJ (2011) Of PAMPs and Effectors: the blurred PTI-ETI dichotomy. The Plant Cell **23**: 4–15

Tian D, Traw MB, Chen JQ, Kreitman M, Bergelson J (2003) Fitness costs of *R*-gene-mediated resistance in *Arabidopsis thaliana*. Nature **423**: 74–77

Tsuda K, Katagiri F (2010) Comparing signaling mechanisms engaged in pattern-triggered and effector-triggered immunity. Current Opinion in Plant Biology **13**: 459–465

Tu J, Datta K, Khush GS, Zhang Q, Datta SK (2000) Field performance of *Xa21* transgenic indica rice (*Oryza sativa* L.), IR72. Theor Appl Genet **101**: 15–20

Van den Ackerveken GFJM, Dunn RM, Cozijnsen AJ, Vossen JPMJ, Van den Broek HWJ, de Wit PJGM (1994) Nitrogen limitation induces expression of the avirulence gene *avr9* in the tomato pathogen *Cladosporium fulvum*. Mol Gen Genet **243**: 277–285

Van Der Weerden NL, Reviews MAFB, 2013 (2012) Plant defensins: common fold, multiple functions. TRENDS in Plant Science. doi: 10.1016/j.fbr.2012.08.004

Van Weymers PSM, Baker K, Chen X, 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**: 556–11

Vleeshouwers VGAA, Oliver RP (2014) Effectors as tools in disease resistance breeding against biotrophic, hemibiotrophic, and necrotrophic plant pathogens. MPMI **27**: 196–206

Vleeshouwers VGAA, Rietman H, Krenek P, Champouret N, Young C, Oh S-K, Wang M, Bouwmeester K, Ben Vosman, Visser RGF, et al (2008) Effector genomics accelerates discovery and functional profiling of potato disease resistance and *Phytophthora infestans* avirulence Genes. PLoS ONE **3**: e2875

Wang T, Liang L, Xue Y, Jia P-F, Chen W, Zhang M-X, Wang Y-C, Li H-J, Yang W-C (2016) A receptor heteromer mediates the male perception of female attractants in plants. Nature **531**: 241–244

Westerink N, Brandwagt BF, de Wit PJGM, Joosten MHAJ (2004) *Cladosporium fulvum* circumvents the second functional resistance gene homologue at the *Cf-4* locus (*Hcr9-4E*) by secretion of a stable *avr4E* isoform. Molecular Microbiology **54**: 533–545

Wheeler MJ, de Graaf BHJ, Hadjiosif N, Perry RM, Poulter NS, Osman K, Vatovec S, Harper A, Franklin FCH, Franklin-Tong VE (2009) Identification of the pollen self-incompatibility determinant in *Papaver rhoeas*. Nature **459**: 992

Whitehouse HLK (1950) Multiple-allelomorph incompatibility of pollen and style in the evolution of the angiosperms. JSTOR. doi: 10.2307/42907111

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

Woriedh M, Wolf S, Márton ML, Hinze A, Gahrtz M, Becker D, Dresselhaus T (2013) External application of gametophyte-specific ZmPMEI1 induces pollen tube burst in maize. Plant Reprod **26**: 255–266

Wright SI, Kalisz S, Slotte T (2013) Evolutionary consequences of self-fertilization in plants. Proceedings of the Royal Society of London B: Biological Sciences **280**: 20130133

Yue J-X, Meyers BC, Chen J-Q, Tian D, Yang S (2012) Tracing the origin and evolutionary history of plant nucleotide-binding site–leucine-rich repeat (NBS-LRR) genes. New Phytol **193**: 1049–1063

Zipfel C (2009) Early molecular events in PAMP-triggered immunity. Current Opinion in Plant Biology **12**: 414–420

Zipfel C, Kunze G, Chinchilla D, Caniard A, Jones JDG, Boller T, Felix G (2006) Perception of the Bacterial PAMP EF-Tu by the Receptor EFR Restricts Agrobacterium-Mediated Transformation. Cell **125**: 749–760

Summary

How to feed the world has become a huge challenge with the increasing world population. Potato, the most important non-cereal crop, is widely consumed as staple food in the Western world, and recently it has proposed to be also promoted as a staple food in developing countries like China. However, the global production of potato has greatly been hampered by a disastrous disease, called potato late blight. This disease is caused by the oomycete pathogen *Phytophthora infestans*, which triggered the great Irish famine in the 1840s. Besides chemical control, resistance (*R*) genes from wild relatives of potato have been introduced into modern cultivars by traditional breeding as well as by transgenic technology (GMO). However, most *R* genes that belong to the nucleotide-binding domain and leucine-rich repeat containing (*NLR*) class, which are generally rapidly defeated by the fast-evolving *P. infestans* population in the field. On the other hand, another form of disease resistance that constitutes the first layer of plant defense in the apoplastic space, has not received enough attention by breeders. This apoplastic defense is mediated by the plant surface immune receptors. These receptors can recognize microbe- or pathogen-associated molecular patterns (MAMPs/PAMPs) or apoplastic effectors of pathogens and trigger a defense response. MAMP/PAMP-triggered immunity (MTI/PTI) is believed to be more wide spectrum and more durable than the typical effector triggered immunity (ETI) that is mediated by cytoplasmic NLR receptors. In this thesis, I studied various apoplastic effectors from *P. infestans* and their receptors in the host, with the ultimate aim to achieve a broader and more durable resistance to late blight.

In **chapter 1**, I summarized the history of potato late blight resistance breeding, the current knowledge of the plant immune system, particularly the plant surface immune receptors and the effector biology in the age of genome sequencing.

To understand the surface immunity against *P. infestans* in potato, I first studied an ubiquitous but functionally unknown apoplastic effector family in *Phytophthora*- the PcF/SCR effectors. PcF (*Phytophthora cactorum-Fragaria*) was identified from *P. cactorum*, and the related effector SCR74 from *Phytophthora infestans* belongs to a highly diverse gene family. They represent small cysteine-rich (SCR) proteins which are normally up-regulated during infection. In **Chapter 2**, I collected all the annotated PcF/SCR proteins that share a PcF domain from

public database. Sequence analysis, phylogenetic, genomic analysis as well as sub-cellular localization, mutagenesis, functional screening and disease tests were performed in this study. PcF genes are conserved in all tested isolates of *P. cactorum* and its orthologs from different *Phytophthora* species are sharing a co-linear genomic architecture. PcF can be recognized by a broad spectrum of different host plants, including strawberry, tomato and potato. In contrast, the *SCR74* genes are exclusively present and expanded in *P. infestans* and under positive selection. They are secreted from haustoria and the cysteine residues are important for maintaining their function in the apoplast. Our effectoromics screening indicated that the SCR74 recognition is confined to wild potatoes. Collectively, this study provides a good example of the effectors from the same family possessing both MAMP/PAMP and effector features. This may lead to identification of multiple host surface immune receptors against these PcF/SCR effectors.

To accelerate the map-based cloning of plant surface immune receptors that perceive apoplastic effectors, such as SCR74 as described in **Chapter 2**, I developed a receptor-like protein (RLP) and receptor-like kinase (RLK) enrichment sequencing (RLP/Kseq) method (**Chapter 3)**. Two diploid *Solanum microdontum* genotypes which respond to INF1 or SCR74-B3b, respectively, were crossed. The F1 population segregates for responses to INF1 and/or SCR74-B3b independently. I designed an enrichment bait library representing the *RLP/RLK* genes predicted from the potato reference genome *S. tuberosum* Group Phureja clone DM1-3 (DM). RLP/KSeq confirmed the localization of the INF1 receptor ELR on chromosome 12, and lead to quick mapping of the putative SCR74 receptor on chromosome 9. Our findings show that RLP/Kseq enabled rapid mapping of plant surface immune receptors and it is especially useful for crop plants with large and complex genomes.

To fine map the SCR74 receptor, in **Chapter 4**, I expanded the mapping population and developed more markers. The SCR74-B3b receptor was mapped to a 74kb interval based on the DM genome. The candidate genes include 3 G-type LecRK (*G-LecRK*) genes. To functional study the putative SCR74 receptor, homology-based cloning was firstly deployed to isolate candidate gene(s), from the SCR74-B3b responsiveness *S. microdontum* spp. *gigantophyllum* genotype GIG362-6. Later, I generated a bacterial artificial chromosome (BAC) library for GIG362-6. Three BAC clones covering the mapping interval were isolated and sequenced. I also used RNAseq data to provide expression data of these candidate genes.

Attempts were made to perform complementation tests of these candidate *G-LecRK* genes by co-expression with the matching SCR74-b3b effector, however, I have not been able to confirm the identification of the SCR74-B3b receptor yet.

Chapter 5 is dedicated to the Pep-13/25, which are typical MAMPs/PAMPs from the transglutaminase GP42 of *Phytophthora*. The Pep-13/25 peptides are known to induce defense responses in parsley and potato, and were studied decades ago, but the receptor has remained undiscovered. Here, MAMP Pep-13/25 were functionally screened on wild *Solanaceae* species for cell death response. Various wild potatoes were found responsive to Pep-13/25, including GIG362-6 that was previously used for mapping of the SCR74-B3b receptor (**Chapter 3** and **4**). A population that is segregating for the response to Pep-25 was generated, and a bulked segregant RNAseq (BSA-RNAseq) method was successfully applied to locate the candidate Pep-25 receptor on chromosome 3 of the DM reference genome. Furthermore, I fine-mapped the Pep-25 receptor to a *RLK* locus. One BAC clone from the Pep-25 responding genotype GIG362-6 was isolated and sequenced. In addition, RNAseq was used for characterizing the gene expression level in the presence or absence of treatment with *P. infestans*. This work provides insight into the Pep-13/25 recognition in *Solanaceae* species and will most likely lead to the identification of a new MAMP receptor in potato.

In **Chapter 6**, I discuss the new findings of this thesis and put them in a broader picture. Four main topics are discussed: **1. What's the boundary between MAMP and effectors, as well as between MPI and ETI? And do the surface immune receptors contribute to resistance breeding?** I believe that the first layer of defense is definitely crucial for engineering more durable resistance to plant diseases, however more studies are needed to understand the whole picture of the plant immunity network; **2. Is the reproduction system of flowering plants evolve from the plant immune system?** The latest knowledge of both systems was reviewed and compared, and based on current data, I suggest that SCR proteins and their receptors might be a "missing link" between the plant reproduction system and the immune system. **3. How to clone plant *R* genes in the era of genome sequencing?** With the advancement of sequencing technologies, many genotyping-by-sequencing methods have emerged, as well as enrichment-based methods like RenSeq. I compare the pros and cons of each method and provide a roadmap to select the best methods to clone plant *R* genes.

Overall, this thesis provides: 1) new insights in the plant-microbe interaction in the extracellular space, 2) our newly developed RLP/KSeq methodology will accelerate the mapping and cloning of novel plant surface immune receptors, and 3) our findings will lead to identification of at least two novel surface immune receptors, which might contribute to more durable late blight resistance in potato.

中文摘要

随着全球人口的快速增长，如何解决其温饱问题成为了一个巨大的挑战。作为最重要的粮食类作物之一，马铃薯在西方一直被作为主粮消费。如今，在很多发展中国家，马铃薯同样被用来作为主粮，比如中国最近就提出了「土豆主粮化」战略。然而，在全球范围内，马铃薯晚疫病(potato late blight)严重影响着马铃薯产量。马铃薯晚疫病是由一种卵菌(oomycete)病原微生物-致病疫霉(*Phytophthora infestans*)引起的病害，这种病原菌曾经在1845年造成过爱尔兰大饥荒。目前，为了防治马铃薯晚疫病，需要使用大量的化学农药。除此之外，在大饥荒之后，许多来源于野生马铃薯中的抗性基因(*R* genes)被引入现代栽培马铃薯。目前大部分已克隆的抗性基因属于一类*NLR* (nucleotide-binding domain and leucine-rich repeat containing) 基因，但是大部分已发掘的抗性基因已经被快速进化的致病疫霉克服，与此同时，植物的第一层免疫系统却常常被忽视。第一层免疫系统是通过膜上的免疫受体识别微生物保守的分子模式(MAMPs/PAMPs)，以及胞外效应子(Apoplastic effectors)，从而触发的免疫反应。比起胞内的免疫系统，这层植物免疫系统能够广泛识别不同的微生物，并且具有相对持久的抗性。在本论文中，作者研究了卵菌中多个胞外效应子以及MAMPs，同时通过遗传学和生物信息学分析定位了它们在植物中的免疫受体。本工作的最终目标是探索马铃薯晚疫病持久抗性的分子机理及其在抗病育种的应用。

在本论文的**第一章**，作者总结了马铃薯晚疫病抗性育种的历史，植物免疫系统，植物膜上免疫受体以及基因组时代的效应子组学。

为了理解马铃薯对抗致病疫霉的膜上免疫系统，作者首先研究了一类在广泛存在于疫霉属中，但功能仍然未知的胞外效应子PcF/SCR家族，这类效应子通常在病原侵染植物时上调表达。第一个家族成员PcF首先发现于*P. cactorum*，而*SCR74*家族仅存在于致病疫霉(*P. infestans*)中，具有非常高的多态性。在**第二章**，作者从公共数据库中收集了所有已知的PcF/SCR效应子的序列，它们都具有共同的PcF结构域。通过对这个效应子家族进行序列分析，系统发生分析，基因组分析，亚细胞定位，突变分析，功能基因组筛选以及致病分析，作者发现*PcF*基因在所有*P. cactorum*菌株中都很保守，并且它们在其他疫霉属物种中的直系同源基因都很保守，同时它们在其基因组上的位置具有共线性。通过效应子组学在茄科的多种植物中进行功能筛选，作者发现PcF可以广泛的被不

同的宿主识别，包括番茄和马铃薯，还有之前报道过的草莓。然而*Scr74*家族具有完全不同的特性，首先，*Scr74*基因只存在于致病疫霉中，该家族在致病疫霉中大量扩增并且受到正向选择，它们的半胱氨酸非常保守，并且对其功能非常重要。它们在致病疫霉的侵染过程中从吸器中分泌。通过效应子组学的筛选，作者发现SCR74的识别只存在于野生马铃薯中。综上所述，本研究发现一类效应子家族的成员可以既具备MAMP/PAMP的特征，也具备效应子的特征。并且本章的工作为克隆多个植物免疫受体打下了基础。

为了加速克隆植物膜上免疫受体，比如在**第二章**中研究的SCR74受体，在**第三章**，作者开发了一种新的方法-受体蛋白(RLP)以及受体激酶(RLK)富集测序(RLP/KSeq)。作者将能够分别识别两个不同的胞外效应子，INF1和SCR74的两个二倍体野生马铃薯 *Solanum microdontum* 进行杂交，得到了对INF1和SCR74的识别独立分离的F1群体。为了设计用于富集测序的诱饵文库(bait library)，作者根据马铃薯参考基因组 *S. tuberosum* Group *Phureja* clone DM1-3 (DM) 预测了其中全部免疫受体蛋白及免疫受体激酶，设计并合成了诱饵文库。通过新开发的RLP/KSeq技术，我们证实了INF1受体位于第12号染色体，同时将未知的SCR74受体定位到了9号染色体。本章的内容证明了RLP/Kseq技术可以用于快速克隆植物免疫受体，尤其适合应用于基因组较大且复杂的物种。

为了对SCR74受体进行精细遗传定位以及克隆，在**第四章**，作者扩大了作图群体并开发了更多分子标记。根据参考基因组，SCR74-B3b的受体被精细定位到了一个74kb的区间，在这个区间里包含三个*G-LecRK*基因。为了对候选基因进行功能分析，作者首先通过同源克隆的方法，从能够识别SCR74-B3b的马铃薯野生种GIG362-6（*S. microdontum* spp. *gigantophyllum*）中对候选基因进行扩增。之后，作者构建了GIG362-6的细菌人工染色体文库(BAC library)，并筛选出三个覆盖定位区间的BAC克隆，同时通过RNAseq缩小候选基因的范围。已将几个候选基因克隆进表达载体，并通过和SCR74共表达进行功能验证。然而，目前还未证实哪个候选基因是SCR74-B3b的受体。

在**第五章**，作者研究了一个卵菌中典型的MAMPs/PAMPs，Pep-13/25。Pep-13/25是疫霉属谷氨酰胺转移酶蛋白中的一段短肽，它们曾被报导可以在马铃薯和欧芹中触发防

御反应，但其受体至今没有被发现。在本章中，作者通过注射Pep-13/25蛋白，在多个茄科物种中筛选了能够识别Pep-13/25的基因型。作者发现了多个能够识别Pep-13/25的马铃薯基因型，其中包括在**第三章**和**第四章**被用来进行SCR74受体定位的亲本GIG362-6。在本章中，作者构建了一个对Pep-25响应分离的群体，并通过混池RNAseq对Pep-25受体进行了快速遗传定位，根据参考基因组，成功的将Pep-25受体定位到了第三号染色体。通过扩大群体精细定位，最终将Pep-25受体定位到了一个受体激酶富集的位点。另外，通过RNAseq得到的转录组数据，作者发现一些受体激酶基因在致病疫霉的侵染的过程中会上调表达。这章的工作首次鉴定了在茄科不同物种对Pep-13/25的识别模式，并且为Pep-13/25受体的克隆打下了坚实的基础。

在**第六章**，作者在更宏观的视角下讨论了本论文中的发现，主要探讨了以下四个问题：**1. 什么是MAMP和效应子的界限？什么是MTI和ETI免疫系统的界限？膜上的免疫受体是否能用于抗病育种？**作者认为植物的第一层免疫系统对于植物的抗性至关重要，但仍然需要更多的研究去理解整个植物免疫网络；**2、开花植物的繁殖系统是否从植物免疫系统中演化而来？**作者对两个系统的最新进展进行了综述和比较，根据目前的数据，作者认为病原微生物中的SCR蛋白和其对应的植物免疫受体可能是这两个系统中「缺失的一环」。**3、如何在基因组时代快速克隆植物抗病基因？**随着测序技术的快速发展，很多genotyping-by-sequencing的方法被开发出来，以及基于富集测序的技术如RenSeq。作者比较了各种方法的优缺点以及如何选择最好的方法快速克隆植物抗病基因。

综上所述，本论文中的工作做出了如下贡献：1、为植物-病原微生物的胞外互作提供了新的视角；2、本论文中开发的RLP/KSeq可以帮助快速克隆植物免疫受体。3、本论文的工作将会引领两个马铃薯中未知免疫受体的克隆，并最终为马铃薯持久晚疫病抗性做出贡献。

Acknowledgements

The adventure is finally coming to an end, I've got a thousand words floating around in my mind.

Vivianne, rather than a supervisor, you are more like an older sister to me. "Are you happy?" is your mantra. Our discussion always brought me tons of positive energy and driving force to keep me moving forward. I am so appreciate the knowledge, skills and networks you gave to me. You always indulge and encourage my unrestrained ideas and allow me to explore the possibilities outside the comfort zone. We've also shared many valuable memories like the backyard BBQ, the trips to Wuhan, Zakopane, Banyuls and Portland. Since I will continue to work in this field, I am looking forward to our collaboration in the future! Wish you, Leonne, Celine, and Caro, Peggy, Hartje, the rabbits all the best! ☺

Richard, thanks for offering me the opportunity to do the PhD in Plant Breeding group, and being my promoter. In the last stage of my PhD, I appreciate your time for reviewing and commenting on my PhD thesis.

Evert, although you are not my promoter officially, but I always like to share ideas and thoughts with you, your feedback and encouragement always gave me lots of strength. Actually you are the first one to read all my chapters, many thanks for everything!

Manos, **Carolina**, **Jaap**, **Doret**, **Juan**, **Aviv**, **Yerisf**, **Xinfang** and **Tatiana**, It's great honor and luck to have colleagues like you, we are always the most open-minded, selfless and collaborative team, everybody likes to share and willing to help. I also enjoy the time we've been shared, the BBQs, beers, travels and the discussions. **Manos**, I was happy to be your paranymph and witness your wedding in Greece, wish you all the best for your future career. **Carolina**, many thanks for being my paranymph! And I am honored to be the witness of your (secret ☺) wedding. Hope you find the ideal job very soon and be happy with Mario! **Doret**, thanks for all the helps in the lab, your concerns and encouragement, and thanks to arrange my (sometimes too many) plants in greenhouse. **Aviv**, **Yerisf**, **Xinfang** and **Tatiana,** hope you

all have a happy and productive PhD time in Vivi's group, **Yerisf** and **Tatiana,** thanks for reviewing my thesis in the last moment.

Thanks to all the collaborators, **Sophien Kamoun**, **Joe Win**, **Suomeng Dong**, **Ingo Hein**, **Katie Baker**, **Miles Armstrong, Paul Birch**, **Shumei Wang** and **Thorsten Nürnberger. Sophien**, I've learnt so much from you, from your YouTube channel, twitter, or when we met in TSL or conferences, I never saw a scientist like you, always willing to share and help the young people in their career. **Ingo**, I still remember my visiting in JHI, the Friday afternoon drinks, the lobster sandwich, the ice cream and the beach with many fossils. Hope all the best to you, Eleanor, Isla and the new baby.

Thanks to the master students worked with me, **Fokion Chatziavgerinos**, **Roland Wouters**, **Laura de Rond, Happyka Fillianti**, **Islam Ahmed Abdelalim Darwish** and **Nicoletta Bertolin**. Thanks for your contribution on this thesis and I really enjoy the time to work with you. **Roland**, I really appreciate your critical thinking and broad knowledge about the nature. **Laura**, you are the most patient, careful, organized and thoughtful student I've ever met. Hope you all the best in your professional career. **Happyka**, you took a whole year with me for your major and minor thesis, you are so efficient, independent and productive at work. It's really funny to hear your saying: "Yeah, it's hard to beat me, because I am a Mom". **Nicoletta**, your signature, infectious smile always brought positive energy to people around you. And you are a very fast learner, hi, should try to be a PhD!

Thanks to the researchers in Plant Breeding and Phytopathology who provide help and discussions, especially **Dr. Herman van Eck**, **Dr. Jack Vossen**, **Prof. Yuling Bai**, **Prof. Bart Thomma** and **Dr. Klaas Bouwmeester**.

Thanks to all the staffs in Plant Breeding who helped me, **Dirkjan Huigen**, **Gerard Bijsterbosch**, **Gert van Arkel**, **Dr. Anne-marie Wolters**, **Isolde Bertram**, **Marjan Bergervoet**, **Iris Capel**, **Bernadette van Kronenburg**, **Fien Meijer-Dekens**, **Irma Straatman**, **Annelies Loonen**, **Linda Kodde**, **Jos Molthoff**, **Johan Bucher**, **Dr. Ronald Hutten**, **Patrick Hendrickx**, **Brian Lavrijssen** and **Danny Esselink**. Especially thanks to **Isolde Bertram**, I already can't remember how many plants you propagated for me, thanks for

taking care of all the potato materials in our collection. Thanks to **Dirkjan Huigen**, you are always helpful to my request, for taking care of our seeds collection, and help to arrange the experiments and crossing in greenhouse. Thanks to **Gerard Bijsterbosch**, in my starting phase, your help make me quickly adapted to the environment. Thanks to the greenhouse staffs **Harm Wiegersma**, **Henk Smid**, **Bertus van der Laan**, **Michel Hagendoorn** and **Bert Essense,** I cannot made so many greenhouse works without your assist, many thanks for taking care of my plants! Thanks to our secretaries **Nicole Trefflich**, **Letty Dijker**, **Janneke van Deursen** and **Daniëlle van der Wee**, for all the helps.

Thanks to all the fellow PhDs at Plant Breeding: **Dr. Anne Giesbers**, **Dr. Chang-Lin Chen**, **Dr. Cynara Romero**, **Charlotte Prodhomme**, **Dr. Geert van Geest**, **Pauline van Haperen**, **Jasper Vermeulen**, **Dr. Rafael Chan**, **Engin Yucel**, **Jeroen Berg**, **Johan Willemsen**, **Katharina Hanika**, **Sri Sunarti**, **Mas Muniroh**, **Jarst van Belle**, **Mathilde Daniau**, **Dr. Michela Appiano**, **Dr. Myluska Caro**, **Peter Dinh**, **Dr. Valentina Bracuto, Atiyeh Kashaninia, Ashikin Nurashikin Binti Kemat.**

Thanks to my Chinese colleagues and friends in Wageningen, **Juan Du, Xi Cheng, Kaile Sun, Chunzhi Zhang, Yiqian Fu, Dongli Gao, Xiaoxue Sun, Xuan Xu, Yajun Wang, Xuexue Shen, Mengjing Sun, Mengyang Liu, Peirong Li, Hao Hu, Huayi Li, Zhe Yan, Lei Cui, Ying Liu, Xing Wang, Yi Wu, Zilin Liao, Li Du, Miru Du, Lina Lou, Ning Guo, Hongmei Huang, Xueping Chen, Lujun Zhang, Xu Cheng, Feng Zhu, Yongran Ji, Wei Qin, Chunxu Song, Jinling Li, Yin Song, Jinbin Wu, Fengjiao Bu, Tao Zhao, Wenbiao Shi, Defeng Shen, Xiaoning Zhang. Xuexue,** thanks for being my paranymph in such a busy period, wish you can graduate smoothly and achieve a great career!

I would like to thank the teachers in HuaZhong Agricultural University, **Prof. Hanhui Kuang**, you led me into this field and taught me how important to always keep integrity in science, and your advise enlighten me when I made very difficult decisions. **Prof. Conghua Xie,** I started to learn the knowledge of potato in your group meeting, and I really appreciate your comments and suggestions with my popular science articles.

Finally, I would like to thank my family, I can't make all of it happen without your unconditional support.

感谢我的父亲母亲，林小叶，杨眉。自从大学以后我就一直在外，回家的时间太少，接下来还要在外面几年，希望能早日回国多陪陪您们。感谢这么多年来对我无条件的支持。感谢我的爷爷，林祉成，从小就最喜欢跟您一起玩，在外的时候经常感到愧疚，没法多回国陪陪您，不过如今我们还可以通过邮件交流。也感谢我的姥爷姥姥：郑新如，杨挺，谢谢你们对我的言传身教。

Jia, I feel grateful for all the minutes we've been through together. Thanks for your tolerance about my sometimes workaholic behavior and thanks for all the encouragement and supporting. I am looking forward to our coming journey.

About the author

Xiao Lin（林啸）was born on 22nd of Apr, 1988, in Beijing, China. Biology has always been his favorite subject since high school. In 2006, he started a Seven-Year Combined Bachelor's and Master's Degree Program at HuaZhong Agricultural University, in Wuhan, China. After 4 years bachelor study on biology science, he spent 3 years to study the genomic evolution of *NBS-LRR* genes in the genomes of *Cucurbitaceae* plants, under the supervision of Professor Hanhui Kuang. In 2013, he got a PhD fellowship from China Scholarship Council (CSC), and started his PhD study with Dr. Vivianne G. A. A Vleeshouwers and Prof. Richard G. F. Visser, in Plant Breeding group of Wageningen University, the Netherlands. During the PhD, he dedicated the research on potato late blight and he applied the genetic and genomic tools to explore the surface immune receptors in potato against the apoplastic effectors or MAMPs from *Phytophthora infestans*. Besides, he likes rock climbing and beer brewing. He also wrote many popular science articles in Chinese on plant breeding technologies, plant immunity and GMO, and some articles were published on newspapers and magazines, including a ebook「土豆驯化记」(*Story of Potato Domestication*). After his PhD, he will continue the journey in science as a post-doctoral researcher in Professor Jonathon D. G. Jones' group at The Sainsbury Lab, Norwich, UK.

List of publications

Lin X, Baker K, Armstrong M, Wouters D, Visser RGF, Wolters PJ, Hein I and Vleeshouwers VGAA(2018), RLP/K enrichment sequencing; a novel method to identify receptor-like protein (RLP) and receptor-like kinase (RLK) genes. (Submitted)

Aguilera-Galvez C, Champouret N, Rietman H, **Lin X**, Wouters D, Chu Z, Jones JDG, Vossen JH, Visser RGF, Wolters PJ, Vleeshouwers VGAA (2018) Two different *R* gene loci co-evolved with Avr2 of *Phytophthora infestans* and confer distinct resistance specificities in potato. Stud Mycol. 89:105-115

Derevnina L, Dagdas YF, De la Concepcion JC, Bialas A, Kellner R, Petre B, Domazakis E, Du J, Wu CH, **Lin X**, Aguilera-Galvez C, Cruz-Mireles N, Vleeshouwers VGAA, Kamoun S (2016) Nine things to know about elicitins. N*ew Phytologist* 1-9

Domazakis E, **Lin X**, Aguilera-Galvez C, Wouters D, Bijsterbosch G, Wolters PJ and Vleeshouwers VGAA (2016) Effectoromics-based identification of cell surface receptors in potato. *Methods in Molecular Biology book chapter* Plant Pattern Recognition Receptors pp 337-353

Lin X, Zhang Y, Kuang H and Chen J (2013) Frequent loss of lineages and deficient duplications accounted for low copy number of disease resistance genes in *Cucurbitaceae*. *BMC genomics* 14:335

Education Statement of the Graduate School

Experimental Plant Sciences

Issued to: Xiao Lin
Date: 31 October 2018
Group: Laboratory of Plant Breeding
University: Wageningen University & Research

1) Start-Up Phase	*date*
► **First presentation of your project**	
Map-Based Cloning of Novel Immune Receptors Recognizing *P. infestans* Apoplastic Effectors in Potato	27 Oct 2014
► **Writing or rewriting a project proposal**	
Isolation and characterization of apoplastic immune receptors in potato: towards a novel type of durable resistance against late blight	May 2014
► **Writing a review or book chapter**	
► **MSc courses**	
► **Laboratory use of isotopes**	
Subtotal Start-Up Phase	*7.5 **

2) Scientific Exposure	*date*
► **EPS PhD student days**	
EPS PhD student day, Leiden, NL	29 Nov 2013
EPS PhD student days 'Get2Gether', Soest, NL	29-30 Jan 2015
EPS PhD student days 'Get2Gether', Soest, NL	28-29 Jan 2016
► **EPS theme symposia**	
EPS Theme 2 Symposium 'Interactions between Plants and Biotic Agents', together with Willie Commelin Scholten Day, Amsterdam, NL	25 Feb 2014
EPS Theme 2 Symposium 'Interactions between Plants and Biotic Agents' together with Willie Commelin Scholten Day, Amsterdam, NL	24 Jan 2018
► **National meetings (e.g. Lunteren days) and other National Platforms**	
Annual meeting 'Experimental Plant Sciences', Lunteren, NL	14-15 Apr 2014
Annual meeting 'Experimental Plant Sciences', Lunteren, NL	09-10 Apr 2018
► **Seminars (series), workshops and symposia**	
Symposium: CALN 2013 Annual Meeting 'Innovative Horticultural Industry and Food Safety in the Netherlands: Past-Present-Future', Wageningen, NL	9 Nov 2013
Symposium: Plant Breeding Research Day 2014, Wageningen, NL	30 Sep 2014
EPS Symposium: 'Omics Advances for Academia and Industry - Towards True Molecular Plant Breeding', Wageningen, NL	11 Dec 2014
Mini-symposium: 'Rewriting our genes' - Wageningen Young Academy, Wageningen, NL	30 Sep 2016
Seminar: 'Using the Nicotiana-TMV system to study resistance gene evolution and plant genome stability' - Hanhui Kuang	11 Sep 2013
Seminar: 'Understanding the molecular mechanisms underlying rice tillering' - Jiayang Li	15 Nov 2013
Seminar: 'Dissecting the interaction between Phytophthora sojae and soybean' - Yuanchao Wang	16 Jul 2014
Seminar: 'Helper NLR proteins of the NRC family in solanaceous plant' - Chih-Hang Wu	5 Mar 2015
Seminar: 'Effectors as molecular probes to understand pathogenesis' - Wenbo Ma	20 Jun 2016
► **Seminar plus**	
► **International symposia and congresses**	
2nd Annual COST SUSTAIN Conference, Zakopane, Poland	15-17 Oct 2014
4th International Conference on Biotic Plant Interactions (ICBPI), Nanjing, China	01-03 Aug 2015
2016 XVII International Congress of the International Society for Molecular Plant-Microbe Interactions (IS-MPMI), Portland, Oregon, USA	17-21 Jul 2016
3rd Annual COST SUSTAIN Conference, Banyuls-sur-Mer, France	17-19 Feb 2016
International Symposium on Potato Late Blight Resistance, Wuhan, China	10-14 Oct 2017
► **Presentations**	
Poster: Spring school 'Host-microbe interactomics'	02-04 Jun 2014
Poster: 4th International Conference on Biotic Plant Interactions (ICBPI)	01-03 Aug 2015
Poster: 2016 XVII International Congress of the International Society for Molecular Plant-Microbe Interactions (IS-MPMI)	17-21 Jul 2016
Talk: COST SUSTAIN 3rd Annual Conference	17-19 Feb 2016
Talk: The International Symposium on Potato Late Blight Resistance	10-14 Oct 2017
Talk: EPS Theme 2 Symposium 'Interactions between Plants and Biotic Agents'	24 Jan 2018
Talk: Users committee meeting for VIDI STW project (first presentation)	18 Nov 2014

	Talk: Users committee meeting for VIDI STW project (last presentation)	19 Jun 2018
►	**IAB interview**	
►	**Excursions**	
	Excursion to the company Rijk Zwaan, De Lier, NL	27 Sep 2013
	Excursion to the companies Genetwister and In2Care, Wageningen, NL	19 Sep 2014
	Excursion to the company Enza Zaden, Enkhuizen, NL	12 Jun 2015
	Excursion to the company J.R.Simplot, Boise, USA	15 Jul 2016
	Excursion to the company Tomato World, Honselerdijk, NL	14 Oct 2016
	Excursion to Flower Trials and the company Dümmen Orange, De Lier, NL	15 Jun 2018
	Subtotal Scientific Exposure	*19.2 **

3) In-Depth Studies		*date*
►	**EPS courses or other PhD courses**	
	Current Trends in Phylogenetics, Wageningen, NL	14-18 Oct 2013
	The Power of RNA-seq, Wageningen, NL	16-18 Dec 2013
	Statistics for the Life Sciences, Wageningen, NL	21-28 May 2014
	Plant Pathogenomics - Training School of the COST Action SUSTAIN on genomics of plant pathogens, Norwich, UK	03-07 Apr 2016
	BASH course, Wageningen, NL	18-20 Apr 2016
	JoinMap course, Wageningen, NL	18 May 2016
►	**Journal club**	
	Member and organizer of Plant Breeding literature discussion club	2015-2017
►	**Individual research training**	
	Short Term Scientific Mission (STSM), James Hutton Institute, Dundee, UK	04-26 Apr 2015
	Subtotal In-Depth Studies	*11.0 **

4) Personal Development		*date*
►	**Skill training courses**	
	English for IELTS Summer Course, Wageningen, NL	18-29 Aug 2014
	Scientific writing, Wageningen, NL	Feb-Apr 2017
	Entrepreneurship in and outside Science, Wageningen, NL	Dec 2014 - Jan 2015
►	**Organisation of PhD students day, course or conference**	
►	**Membership of Board, Committee or PhD council**	
	Subtotal Personal Development	*5.6 **

TOTAL NUMBER OF CREDIT POINTS	**43.3***
Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which comprises of a minimum total of 30 ECTS credits.	
** A credit represents a normative study load of 28 hours of study.*	

This research was performed at Plant Breeding, Wageningen University & Research. This research was supported by the VIDI grant (no. 12378), awarded to Dr Vivianne G. A. A. Vleeshouwers from the Netherlands Organization for Scientific Research (NWO). Xiao Lin was sponsored by a PhD fellowship from the China Scholarship Council (CSC, no. 201306760013).

Cover illustration: @鲸鱼老爸 & @蕾蕾的白萝卜

Cover layout Design: Xiao Lin

Printed by: Ridderprint B.V. | Ridderkerk