# Suppression and triggering of Arabidopsis immunity by Albugo species

**TORSTEN SCHULTZ-LARSEN**

**Thesis submitted to the University of East Anglia**

**for the degree of Doctor of Philosophy**

**September 2012**

# Table of Content

## Abstract

*Albugo* species are obligate biotrophic phytopathogens. Like other biotrophs, they are anticipated to secrete effectors that can suppress or trigger plant defenses; the nature of *Albugo* effectors is currently unknown.

Sequencing of *A. laibachii* isolate Nc14 (*AlNc14*) genome reveals 13032 genes encoded in a ~37 Mb genome. We analyze the effector complement of *AlNc14* and find known effector classes but also classes unique to *A. laibachii*. Experiments reveal that CHXCs are a novel class of effectors that suppress host defense.

We functionally characterize two predicted *AlNc14* effectors in detail; CHXC1 a potential core effector conserved in other oomycete species, and SSP6, a fast-evolving effector specific to *A. laibachii*.

CHXC1 encodes a nuclear localized HECT E3 ligase homolog, which suppresses host defenses dependent on cys651.

We find 7 variants of SSP6 that are under diversifying selection. Two highly expressed variants SSP6-2c and SSP6-A are plasma membrane localized when expressed *in planta*. Interestingly, SSP6-2c but not SSP6-A, is able to enhance growth of *P. infestans* race blue 13 and suppress flg22-dependent ROS production. In *Arabidopsis* cells we find SSP6-2c localizes around *AlNc14* haustoria. We propose that *AlNc14* secretes the effectors SSP6-2c and CHXC1 into the plant cell to suppress defense and promote infection.

Current methods to screen for virulence of effector candidates predominantly rely on measuring growth of bacterial pathogens. Quantitative assessment of resistance and susceptibility to eukaryotic pathogens is more difficult. We develop a semi-automated high-throughput system for assaying *Hpa* growth.

We investigate the genetic basis of resistance to *Albugo* in *Arabidopsis*. We find that resistance to *AlNc14* is linked to *RAC1* and *RAC3* in Ksk-1. In contrast, resistance to *A. candida* Nc2 (*AcNc2*) is linked to *WRR4* in Col-0, Col-5 and Ksk-1. A second dominant locus, *WRR5a/b* in Col-5 also confers resistance to *AlNc2*. Thus, different R-genes and presumably different effectors govern resistance to *AlNc14* and *AcNc2*.

## Acknowledgements

First of all I would like to really thank Jonathan Jones, aka JJ, for giving me this opportunity to pursue my studies in his lab resulting in this PhD thesis. I am grateful for the supervision and insight that Eric Kemen has provided me with from the start of this project. I would also like to thank Sophien Kamoun for being on my supervisory committee and providing useful input. This PhD would have been a financial disaster without the funding that both the Danish agency for science, technology and Innovation, and JJ provided.

I have written the thesis in plural, as this thesis would not have been in its current state without the support and advice from Eric, Ariane, Kate, Lennart, Alexandre, Volkan, Anastasia, Oliver, and Jonathan.

I would like to thank the Albugans of all ages - you have all provided directions, shared exciting new data, and valuable discussions. The rest of the JJ lab I would like to thank for friendships, memorable boat trips, and support. The lab would not have been the same without you. Of course I would like to thank everybody at the TSL and JIC past and present who've made these last four years fun, memorable and Norwich a fine place to be (you know who you are).

I would also like to thank my family for support and love. Ida I thank you for being who you are.

And… to all who I have mentioned and everyone that I forgot, I thank you all once again.

## Publications arising from this work

Eric Kemen, Anastasia Gardiner, **Torsten Schultz-Larsen**, Ariane C Kemen, Alexi L Balmuth, Alexandre Robert-Seilaniantz, Kate Bailey, Eric Holub, David J Studholme, Dan Maclean, and Jonathan D G Jones, **2011**. Gene Gain and Loss during Evolution of Obligate Parasitism in the White Rust Pathogen of Arabidopsis thaliana. *PLoS Biology*, 9(7), p.e1001094.

Anastasia Gardiner, Eric Kemen, Kate Bailey, Mark McMullan, Alexandre Robert-Seilaniantz, **Torsten Schultz-Larsen**, Cock van Oosterhout, Jonathan D. G. Jones. **2012.** Widespread Introgression in Mosaic Genomes of *Albugo candida* Pathotypes, *Manuscript in preparation*

**Torsten Schultz-Larsen**, Eric Kemen, and Jonathan Jones. **2012.** A method for automated quantification of *Hyaloperonospora arabidopsidis* growth using Uvitex2B. *Manuscript in preparation*

## List of abbreviations

| | |
|---|---|
| *Ac* | Albugo candida |
| acd11 | Accelerated cell death 11 |
| *Acem1* | A. laibachii strain Acem1 |
| *Acem2* | *Albugo candida* isolate Em2 |
| *AcNc2* | *Albugo candida* isolate Nc2 |
| AcNc2 | A. candida Norwich 2 |
| *AcNc2* | A. candida Nc2 |
| ADR1 | ACTIVATED DISEASE RESISTANCE 1 |
| AFB1 | auxin signaling F-box protein 1 |
| *Alem1* | *Albugo laibachii* isolate Em1 |
| AlNc14 | *Albugo laibachii* isolate *Nc14* |
| *arf9*-1 | auxin response factor 9-1 |
| AtNup88/MOS7 | modifier of SNC1 |
| ATR | *Arabidopsis thaliana* recognized |
| ATR1 | Arabidopsis thaliana recognized 1 |
| ATR13 | Arabidopsis thaliana recognized 13 |
| Ave1 | Avirulence on Ve1 tomato |
| Avr | Avirulence |
| AX21 | activator of XA21-mediated immunity |
| BAK1 | BRASSINOSTEROID INSENSITIVE 1-ASSOCIATED KINASE 1 |
| Bgh | Blumeria graminis f.sp. hordei |
| BIK1 | BOTRYTIS-INDUCED KINASE 1 |
| bp | base pairs |
| bri1 | brassinosteroid-insensitive-1 |
| BSMT1 | SA methyl transferase 1 |
| BTH | Benzo-thiadizole |
| CC-NB-LRR | coiled-coil NB-LRR |
| CEBiP | chitin oligosaccharide elicitor- binding protein |
| CERK1 | Chitin Elicitor Receptor Kinase 1 |
| CHXC | Cys-His-X-Cys (motif of *Albugo* sp. effectors, X: any amino acid) |
| Col-0 | Arabidopsis Columbia ecotype |
| Col-ein5-1 | ethylen insensitive 5 |
| Con-A | Concanavalin A |
| CRN | Crinkling and Necrosis |
| CSP22 | cold shock protein 22 |
| DAMPs | Danger Associated Molecular Patterns |
| dmr | downy mildew resistant |
| dpi | Days post infection |
| DTT | Dithiothreitol |
| Ecp6 | extracellular protein 6 |
| EDS1 | enhanced disease susceptibility 1 |
| EDV | effector-detector-vector |
| Ef-Tu | elongation factor Tu |

| | |
|---|---|
| EHM | Extra-haustorial membrane |
| EHMx | Extra-haustorial matrix |
| EOD1/BB | Big brother |
| EPIC1 | extracellular protease inhibitor with cystatin- like domain 1 |
| ETI | Effector-Triggered Immunity |
| ETS | Effector-Triggered susceptibility |
| flg22 | N-terminal 22 amino acid peptide from bacterial flagellin |
| FLS2 | FLAGELLLIN INSENSITIVE 2 |
| HDP | heme detoxification protein |
| HECT | homologous to the E6-AP carboxyl terminus |
| *Hpa* | *Hyaloperonospora arabidopsidis* |
| Hpi | Hours post infection |
| HR | Hypersensitive response |
| hrc | hrp genes conserved |
| hrp | hypersensitive response and pathogenicity |
| HSP90 | heat shock protein 90 |
| ICS1 | Isochorismate Synthase |
| INDEL | insertions and deletions |
| INV1p | Invertase |
| JAZ | jasmonate ZIM domain |
| kDa | kilo daltons |
| LAZ2/SDG8 | Lazarus2/SET(Su(var)3-9, E(z) and Trithorax- conserved) DOMAIN GROUP 8 |
| LRR | Leucine–Rich Repeat |
| LRR-RLK | leucine-rich repeat receptor-like kinase |
| lsd1 | lesions simulating disease 1 |
| LysM | lysin-motif |
| MAHRP1 | membrane associated histidine-rich protein 1 |
| MAMP | Microbe-Associated Molecular Patterns |
| MBP | maltose binding protein |
| MOS6 | MODIFIER OF SNC1, 6 |
| NahG | salicylate hydroxylase |
| NB-LRR | nucleotide-binding-site and leucine rich repeat |
| NDR1 | NON RACE-SPECIFIC DISEASE RESISTANCE |
| NleL | Non-Lee-Encoded effector Ligase |
| NLPs | Nep1-like proteins |
| NLS | nuclear localization signal |
| NopM | nodulation outer protein M |
| NPC | nuclear pore complex |
| PAD4 | phytoalexin deficient 4 |
| PAMPs | Pathogen-Associated Molecular Patterns |
| PCD | programmed cell death |
| pEDV6 | effector detector vector version 6 |
| PEN1 | penetration 1 |
| PEXEL/VTS | Plasmodium export element/vacuolar targeting signal |
| PexRD2 | putative extracellular protein RD2 |
| PGNs | peptidoglycans |

| | |
|---|---|
| PIP1 | Phytophthora inhibited protease 1 |
| PIPs | phosphatidyl inositol phosphates |
| pmr | powdery mildew resistance |
| PNEP | PEXEL-negative |
| PRRs | Pattern Recognition Receptors |
| *Pst* | *Pseudomonas syringae* pv. tomato |
| Pst DC3000 | Pseudomonas syringae pv. tomato DC3000 |
| PTEX | PEXEL/VTS motif containing protein |
| PTI | PAMP/Pattern Triggered Immunity |
| *R*-genes | resistance genes |
| RAC | resistance to Albugo candida |
| RCR3 | required for Cladosporium fulvum resistance 3 |
| REX2 | ring-exported protein 2 |
| RIN4 | RPM1 interacting protein 4 |
| RING | really interesting new gene |
| ROS | Reactive Oxygen Species |
| RPM1 | resistance to *Pseudomonas syringae pv. maculicola* 1 |
| RPP | Resistance to *Peronospora parasitica* (ex-*Hpa*) |
| RPS2 | *Pseudomonas syringae pv. syringae 2* |
| Rps3b | resistance to Phytophthora sojae 3b |
| RPS4 | P. syringae resistance 4 |
| RTP1p | rust transferred protein 1 |
| RxLR | Arg-X-Leu-Arg (motif of oomycete effectors, X: any amino acid) |
| SA | Salicylic acid |
| SAG101 | senescence-associated protein 101 |
| SAR | systemic acquired resistance |
| SBP1 | skeleton-binding protein 1 |
| SERKs | Somatic Embryo Receptor kinases |
| SGT1 | Suppressor of G-two allele of skp1 |
| SmCHS1 | Saprolegnia monoica chitin synthase 1 |
| SNC1 | suppressor of NPR1 constitutive 1 |
| SNP | single nucleotide polymorphism |
| SP | signal peptide |
| SSPs | for short secreted protein |
| STAND | Signal Transduction ATPases with Numerous Domains |
| T3SS | Type III secretion system |
| TIR-NB-LRR | Toll and mammalian interleukin (IL-1) receptor NB-LRR |
| TIR1 | Transport Inhibitor Response 1 |
| VAMPs | vesicle-associated membrane proteins |
| WGA | wheatgerm agglutinin |
| WRR4 | White rust resistance 4 |
| Xa21 | Xanthomonas resistance 21 |
| Y2H | Yeast-two-hybrid |
| µg | micro gram |
| µl | micro liter |
| µM | micro molar |

# 1. INTRODUCTION

## 1.1 Interactions between plants and microbes

Plants are sessile organisms that continuously are exposed to microorganisms that can colonize them as hosts for proliferation and reproduction. However, this rarely results in disease. Plants have evolved means to resist invading pathogens with physical barriers and synthesis of antimicrobial phytoalexins conferring broad resistance to microbes. Most pathogens coevolve with their host plants and selection favors pathogens that can evade or overcome resistance.

Pathogenic microbes often inhabit the extracellular space between plant cells, but many also obtain nutrients from the plant cells via specialized feeding structures. In order to block defense responses, pathogens secrete effector molecules into plant cells. For this reason plants have evolved two strategies to detect pathogen presence in intracellular and extracellular compartments (Jones & Dangl 2006; Boller & Felix 2009).

The first layer consists of plasma membrane-located extracellular receptors termed pattern recognition receptors (PRRs) that recognize pathogen associated molecular patterns (PAMPs), or damage-associated molecular patterns DAMPs (Boller & Felix 2009). DAMPs are degradation products of endogenous plant components, such as cutin or cell wall components degraded by the activity of the pathogen. In contrast PAMPs are molecules that are often relatively evolutionarily conserved within a class of microbes regardless of their pathogenicity. The best characterized example of a PAMP is flagellin, which is a core component of the flagellar motor complex. Recognition by PRRs leads to induction of PAMP-triggered immunity (PTI) resulting in the restriction of pathogen growth.

The second layer of defense is primarily active inside the plant cell and is governed by the action of intracellular receptors encoded by resistance (*R*) genes that respond to pathogen-secreted effectors to trigger effector triggered immunity (ETI) (Dangl & Jones 2001).

In contrast to PAMPs, effectors are often non-essential dispensable genes that are highly polymorphic between and within species (Raffaele & Kamoun 2012; Lindeberg et al. 2012). *R*-genes co-evolve with effectors and are diverse within and between species, whereas PAMP receptors are generally non-polymorphic within a species.

Defense responses upon perception of a PAMP or effector are similar. Quantitatively ETI usually results in a stronger defense response than PTI. In some cases, ETI results in localized cell death at the site of effector recognition, which is termed the hypersensitive response (HR) (Hinsch & Staskawicz 1996).

In conclusion, resistance to a given pathogen is the sum of recognition of slowly evolving PAMPs and fast evolving effectors triggering PTI and ETI respectively (Jones & Dangl 2006). These distinct defense mechanisms led Schulze-Lefert and Panstruga (2011) to postulate that pathogens, which co-evolve with their hosts, are mostly governed by ETI, whereas those that do not are restricted by PTI.

This introduction will focus on both layers of defense and on mechanisms pathogens employ to avoid host defense with an emphasis on proteinaceous effectors.

## 1.2 PTI signaling

### 1.2.1 Pathogen associated molecular paterns

Although PAMPs are highly conserved molecules they are under a selective pressure to evade recognition. Examples of PAMPs include, flagellin, elongation factor Tu (Ef-Tu) and lipopolysaccharides from gram-negative bacteria (Zipfel et al., 2004, 2006). Eukaryotic examples include β-glucan-binding protein and chitin oligosaccharides (N-acetylchitooligosaccharides) (Kaku et al., 2006). PAMPs are experiencing two evolutionarily opposing forces. To maintain essential cellular function results in strong purifying selection, whereas recognition by the host defense system exerts a strong diversifying selection pressure to avoid recognition by PRRs. Fifty five potential candidates were identified in *Pseudomonas* sp. using bioinformatics, the approach was validated as Ef-Tu was predicted to be a potential PAMP (McCann et al. 2012). Subsequent tests showed that other candidate PAMPs also functioned as such because they triggered ROS accumulation and callose deposition in *Arabidopsis* (McCann et al. 2012). Thus, there is a larger diversity of pathogen derived PAMPs than what is currently known.

### 1.2.2 PAMP recognition receptors

PRRs characterized to date are all located at the plasma membrane. They consist of three structurally distinct parts: an N-terminal extracellular domain, a transmembrane domain, and an intracellular domain. The intracellular domain is subdivided into two groups: those with a short cytoplasmic tail that is without function, and the others that have a kinase domain (Boller & Felix 2009). The extracellular domain is required for PAMP perception and different types of domains have been described, such as leucine-rich repeat (LRR) and LysM domains (Boller & Felix 2009).

The best-studied leucine-rich repeat receptor-like kinase (LRR-RLK) is flagellin-sensing 2 (FLS2) (Zipfel et al. 2004). The extracellular domain of FLS2 is composed of 28 tandem leucine-rich repeats (LRRs), which recognize the bacterial flagellum. A 22 amino acid peptide derived from the most conserved region of the bacterial flagellum, flg22, is sufficient to induce signaling via FLS2 in *Arabidopsis* (Gomez-Gomez et al. 1999). The recognition of flg22 by FLS2 is presumed to be direct, as a point mutation ($FLS2^{G307R}$) in the $10^{th}$ LRR domain severely reduces the binding affinity of flg22 (Gomez-Gomez et al. 2001). However, recognition may depend on a broader region as the LRR repeats 9-15 also contribute to flg22 perception (Dunning et al. 2007). Other examples of LRR-RLKs include: the brassica specific PRR Ef-Tu RECEPTOR that recognizes the bacterial elongation factor Tu (Ef-Tu) (Zipfel et al. 2006), and rice PRR Xa21 (*Xanthomonas resistance* 21), which recognizes the *Xanthomonas oryzae* pv. *oryzae* sulfated peptide Ax21 (activator of Xa21-mediated immunity) (W. Y. Song et al. 1995; S.-W. Lee et al. 2009).

Another type of extracellular domain that recognize PAMPs was initially characterized in bacteria. These LysM domain proteins were found to bind peptidoglycans (PGNs). In plants LysM containing PRRs have been found to recognize both chitin and PGNs (Monaghan & Zipfel 2012). CERK1 (Chitin Elicitor Receptor Kinase 1) is the major chitin binding protein in *Arabidopsis* and is required for signaling (Miya et al. 2007; Iizasa et al. 2010). A recent paper by Liu et al. (2012) describes the crystal structure of the ectodomain of CERK1. The 3 LysM domains form a tightly packed globular structure, but only the second LysM domain is required for chitin binding. Chitin octamers, but not hexamers and pentamers, are ligands of CERK1. Binding of chitin ocatamers leads to homodimerization of CERK1, which activates PTI (T. Liu et al. 2012).

Other LysM containing PRRs bind PGNs but binding and signaling requirements are

different. In contrast to chitin binding, binding of PGNs seem to require tri-complex formation, since knockouts of the PGN binding proteins LYM1 (LysM RLP1) and LYM3 were as insensitive to PGNs as a *lym1/lym3* double knockout. Interestingly, perception of PGNs also requires CERK1 (Willmann et al. 2011). As these PAMPs are recognized by a tri-complex, this suggests that a multitude of potential PAMP recognition complexes could be assembled and allows for recognition of a plethora of different PAMPs.

### 1.2.3 Downstream signaling events and attenuation of signaling

Most known PRRs form a heteromer with a downstream signal amplifier BRASSINOSTEROID INSENSITIVE 1-ASSOCIATED KINASE 1 (BAK1) that is required for downstream signaling but not PAMP binding (Monaghan & Zipfel 2012). A notable exception to this rule is CERK1 (Zipfel 2008). BAK1 is a central regulator of defense signaling to both necrotrophic and biotrophic pathogens (Schwessinger et al. 2011). BAK1 is part of a 5 member family of Somatic Embryo Receptor kinases (SERKs); upon stimulation of PRRs i.e. EFR and FLS2, BAK1 and other SERKs are recruited to the PRRs but with unequal affinities that correlates with the importance in PTI signaling (Roux et al. 2011). A *bak1-5/serk4* double knockout underlines the importance of signal amplification in PAMP perception, as the mutant is almost insensitive to the PAMPs AtPep1, elf18 and flg22 (Schwessinger et al. 2011). This type of signal amplification seems to be evolutionarily conserved as silencing of *BAK1* in *N. benthamiana* results in: loss of ROS signaling upon stimulation with INF1 and CSP22 (cold shock protein 22), and enhanced virulence of isolates of *Phytophthora* (Chaparro-Garcia et al. 2011; Heese et al. 2007).

Mechanistically, the function of BAK1 is to trans-phosphorylate downstream signaling components such as the cytoplasmic kinase BOTRYTIS-INDUCED KINASE 1 (BIK1)

(D. Lu et al. 2010).

Interestingly, attenuation of signaling has been shown to require receptor degradation via ubiquitination (D. Lu et al. 2011; Trujillo et al. 2008). Further downstream events are less well characterized and beyond the scope of this review but they include activation of MAP kinase cascades, accumulation of reactive oxygen species (ROS), changes in ion fluxes, callose deposition, and induction of gene expression (Boller & Felix 2009; Monaghan & Zipfel 2012; Beck et al. 2012; Segonzac & Zipfel 2011).

## 1.3 Mechanisms of effector delivery

### 1.3.1 General strategies to suppress host defense

Physical barriers, PRRs and PTI constitute the first line of defense that provides resistance against a wide range of microbes. Many types of pathogen ranging from bacteria to fungi and oomycetes deploy effectors to suppress PTI. The effector proteins are delivered into the plant cell either by injection via a bacterial type III secretion system (T3SS) or through a haustorium (Jones & Dangl 2006; Dodds & Rathjen 2010). Other mechanisms are used including secretion of toxins and hormone mimics such as ToxA and coronatine (Robert-Seilaniantz, Grant, et al. 2011a; Manning & Ciuffetti 2005).

### 1.3.2 The type III secretion system

The most comprehensively studied effector delivery mechanism is the T3SS, which likely evolved from the bacterial flagellum and is used by a range of gram-negative bacteria (Moreas et al., 2008). The T3SS forms a pilus that is inserted into the plant cell. The T3SS is a key player in effector delivery since lack of a functional T3SS results in the loss of virulence (He et al., 2004). In *Pseudomonas syringae* pv. tomato DC3000 (*Pst* DC3000)

the T3SS is encoded by the hypersensitive response and pathogenicity (*hrp*) genes. These fall into 3 classes (Jin et al., 2003). i) The *hrc* genes (*hrp* genes conserved) encodes the pilus and ii) a class of secreted proteins that codes for the extracellular part of the T3SS. The last class (iii) consists of regulatory genes (eg *HrpRS*), which are induced in minimal media. These have been found to regulate the expression of the T3SS and T3S effector proteins (Innes et al., 1993). Two reports found that approximately 20-30 effectors are secreted into the plant cell during *Pst* infection (Chang et al., 2005, Guttman et al., 2002).

### 1.3.3 Effector secretion from haustoria forming pathogens

Oomycete pathogens such as *Hyaloperonospora arabidopsidis* (*Hpa*) and *Albugo candida* (*Ac*) and fungal rusts such as *Uromyces fabae* and *Puccinia graminis* create specialized plant-pathogen interaction surfaces termed haustoria (Kemen & Jones 2012). These structures are also found in powdery mildews i.e. *Erysiphe orontii* and *Blumeria graminis* f.sp. *hordei*, where they are the most prominent pathogen-plant contact surface. This has led to the hypothesis that haustoria are important sites for nutrient uptake from the host (Voegele & Mendgen 2011).

The oomycete haustorium consists of multiple morphologically different layers (Figure 1.1). The extrahaustorial membrane (EHM) is the outermost layer and is contiguous with the plant plasma membrane and presumably of plant origin. While the subcellular compartment from which the EHM originates is still in question, it has been speculated to be at least in part of tonoplastic origin (Caillaud, Piquerez, Fabro, et al. 2012b). In a study by Lu et al. (2012) the EHM of *P. infestans* and *Hpa* was found not to contain plant membrane proteins such as aquaporin, a calcium transporter, and plant PRRs. The exclusion of these components suggests that only selected proteins are included in the EHM. The EHM is separated from the haustorial cell wall by an electron-dense space

referred to as the extrahaustorial matrix (EHMx) (Soylu 2004). The haustorial cell wall and the extrahaustorial membrane are presumed to be in contact at the haustrial neck. A neckband of callose surrounds the necks of downy and powdery mildews. Interestingly, this neck band is absent from *Albugo* sp. haustoria the first few days and only appears in old infections often after sporulation (Soylu et al. 2003). During an incompatible reaction between *Albugo* and *Arabidopsis* the haustorium is completely encased in callose (Soylu et al. 2003). Beneath the haustorial cell wall is a haustorial plasma membrane. In *Uromyces fabae* it has been shown to harbor transporters for uptake of sugars and amino acids (Hahn et al. 1997; Voegele et al. 2001). For example, cytological and biochemical characterization showed that *Uromyces fabae* secretes an Invertase (INV1p), which metabolises sucrose into fructose and glucose in the EHMx (Voegele et al. 2006). The sugars can then be taken up into the haustorium by a haustorial membrane resident sugar transporter HXT1p dependent on a proton gradient (Voegele et al. 2001). It is therefore likely that haustoria function to absorb nutrients from the host into the pathogen (Voegele & Mendgen 2011).

In addition to its function as nutrient uptake site, the haustorium is also likely the main site where effectors are secreted into the plant cell (Bozkurt et al. 2012). In a pioneering study Kemen et al (2005) immuno-localized the rust transferred protein 1 (RTP1p) from *Uromyces fabae* and found that RTP1p was secreted from the haustorium into the plant cell. Sequence analysis of RTP1p revealed the presence of an N-terminal signal peptide and a two potential glycosylation sites, but no other defining features that would suggest a role in secretion and uptake into the host were identified.

While no obvious linear "effector delivery motif" has been defined for rusts, cloning the avirulence genes *ATR1* (*Arabidopsis thaliana* recognized 1) and *ATR13* from the oomycete *Hpa,* in addition to *Avr3a* (Avirulence 3a) and *Avr1b* from *P. infestans*, revealed that all

possessed an N-terminal signal peptide (SP) and a positionally constrained consensus sequence of sequence arginine, any amino acid, leucine, arginine (RXLR) motif and typically followed by a stretch of acidic (D/E) amino acids (dEER) (R. Allen et al. 2004; W. Shan et al. 2004; Rehmany et al. 2005; Armstrong et al. 2005; Win et al. 2007).

Based on these findings Whisson and colleagues (2007) speculated that the RXLR motif was required for translocation into the plant cell. To test this hypothesis they used Avr3a that is recognized by R3a from *Solanum demissum* (Armstrong et al. 2005). Specifically, they transformed *P. infestans* isolate 88069 (*avr3a*) with either wild type Avr3a or a mutant version Avr3a (AAAA), in which the RXLR or dEER motif is mutated to alanines. Infection of *Solanum demissum* plants revealed an RXLR and dEER motif dependent recognition by R3a suggesting that it was required for delivery to the plant cell. This was independently verified by expressing the N-terminal part of Avr3a fused to *gusA* in *P. infestans* growing on potato and showing that uptake was dependent on the RXLR and dEER motif. Interestingly, the notion that the haustorium is a site of focal secretion of proteins into the plant cell was substantiated by the observation that fluorescently tagged Avr3a expressed in *P. infestans* accumulates around haustoria, presumably in the EHMx (Whisson et al. 2007). These results were corroborated by experiments with the *P. sojae* effector Avr1b. Transgenic *P. sojae* expressing either wild type Avr1b or a Avr1b with the RXLR motif substituted with alanines was shown to be avirulent on soy plant seedlings harboring Rps1b, dependent on the RXLR motif (Dou et al. 2008). Therefore, the haustorium seems to be a site of RXLR effector delivery in oomycetes.

### 1.3.4 Effector delivery in *P. falciparum*

Oomycetes are taxonomically closer to green algae and *Plasmodium falciparum* than fungi and interestingly the RXLR domain has a comparable position and shares similarity to the

PEXEL/VTS (Plasmodium export element/vacuolar targeting signal) motif found in effectors from *Plasmodium falciparum*. Like the RXLR motif, the PEXEL/VTS motif has been shown to be necessary for secretion of malarial effector proteins into host erythrocytes (Hiller et al. 2004).

These two motifs have been shown to be functionally equivalent, as GFP tagged *P. infestans* effectors expressed in *Plasmodium falciparum* are transferred into the host erythrocyte (Bhattacharjee et al. 2006). Reciprocal experiments with PEXEL/VTS - Avr3a chimeric proteins expressed in *P. infestans* isolate 88069 show that the normally virulent strain becomes avirulent on R3a-expressing potatoes. A similar effector delivery system may exist in *Hpa*, as the RXLR domains from ATR1-NdWsB and ATR13 also induce resistance in an equivalent experiment (Grouffaud et al. 2008).

Whether the existence of a common mechanism for protein translocation is a consequence of convergent evolution or common ancestry is a matter of debate. Recent progress in *Plasmodium falciparum* protein translocation sheds more light on the effector delivery mechanism. Interestingly, the PEXEL/VTS motif containing (PTEX) proteins are transported to and concentrated in specialized secretory structures called Maurer's clefts (Bhattacharjee et al. 2008). Therefore, the PEXEL/VTS motif seems to function as a sorting signal within the parasite.

While, the entire secretory route has not yet been elucidated, some PTEX proteins are directed to the ER in a SP (signal peptide) dependent manner; in the ER they are processed by a PEXEL/VTS protease in lieu of a signal peptidase (Boddey et al. 2010). The identity of the aspartyl protease was recently discovered; Plasmepsin V recognizes the PEXEL motif and cleaves PTEX proteins after the conserved leucine (RXL↓xE/D/Q). The C-terminal part of the PTEX protein is subsequently N-terminally acetylated and secreted into the parasitic vacuole (Chang *et al*., 2008). Targeting to the parasitic vacuole is

presumably determined within the parasite, as N-terminal acetylation is insufficient for protein export (Boddey *et al*., 2009).

A putative export mechanism has recently been discovered within the parasitic vacuole. de Koning-Ward and colleagues (2009) identified a protein complex inserted into the parasitophorous vacuole membrane, which is essential for blood-stage growth of *P. falciparum*. The PTEX complex interacts specifically with known PTEX proteins. It is ATP dependent and is a core of two chaperones and a putative membrane channel with homology to the pore forming toxic haemolysin E from *E. coli*. Since homologs of these proteins exist in plant pathogenic oomycetes this mechanism is an interesting candidate for an effector delivery system.

### 1.3.5 Uptake of effectors into plant cells

While Avr3a and Avr1b require the RXLR motif for uptake into the plant, the mechanism of uptake is poorly understood and highly controversial. Dou et al. (2008) reported that C-terminally GFP tagged Avr1b from *P. sojae* purified from *E. coli* could enter soybean cells in an RXLR-dEER motif dependent manner in the absence of the pathogen.

Transient expression in *N. benthaminana* of cerulean fluorescent protein (CFP) versions of AvrM and Avr567 from the flax rust *Melampsora lini* provides additional evidence for uptake in absence of the pathogen (Rafiqi et al. 2010). Based on these observation is seems likely that effector translocation in plants, in contrast to animal systems, could rely on the ability of the effector to bind a plant surface receptor to enter cells (Kale & Tyler 2011).

The mechanism and mode of effector entry is still highly contested and an area of active research (Ellis & Dodds 2011). Recently, Kale et al (2010) provided evidence that several RXLR effectors, among these, Avr3a and Avr1b could bind to phosphatidyl inositol phosphates (PIPs) and the binding was mediated by the RXLR-motif. Evidence for binding

to PIPs came predominantly from binding assays with an *E. coli* produced effector, which was used as a probe against membrane immobilized PIPs. Thus, the specific biophysical binding properties of this binding were not tested. Further support for the role of PIPs in effector uptake, came from competition assays, where exogenous addition of PIPs to plants inhibited uptake of effectors into the plant cell. These data led Kale et al. to propose that oomycete effectors are taken up into the plant cell by binding to PIPs in the plasma membrane, which requires the RXLR motif, resulting in endocytosis and uptake of the effector (Kale et al. 2010).

While effectors do seem to bind PIPs the relevance of this for effector uptake is not clear. AvrM was shown to bind PIPs *in vitro*, but remarkably deletion of the region required for uptake into plant cells was not required for this binding (Gan et al. 2010). Similarly, Yaeno et al. (2012) report that PIP binding of Avr3a and Avr1b is not dependent on the RXLR motif, but rather a positive patch of amino acids in the C-terminus (Yaeno et al. 2011). Mutations in this patch on Avr3a or Avr1b diminishes PIP binding, but does not alter Avr3a avirulence on R3a-expressing *N. benthamiana* plants. Interestingly, the Avr3a mutant is unable to stabilize the host U-box E3 ubiquitin ligase CMPG1, whose stability is inversely correlated with PCD. This led Yaeno et al. to suggest that the role of PIP binding is to stabilize CMPG1 and not facilitate host cell entry.

However, the biological significance of PIP binding by Avr3a has been called into question, as a recent paper describes the biochemical and thermodynamic characteristics of the interaction between PIPs and Avr3a (Wawra, Agacan, et al. 2012a). Surprisingly, they find that only denatured Avr3a protein binds to PIPs.

## 1.4 Identification of effectors in oomycetes

### 1.4.1 Introduction to oomycete genomics

As discussed above the cloning of avirulence genes defined some oomycete proteins as potential effectors. The majority of the avirulence genes cloned to date are from the *Phytophthora* genus, a comprehensive list of cloned oomycete avirulence genes has recently been reviewed (Vleeshouwers et al. 2011; Hein et al. 2009; Chisholm et al. 2006)**.** The cloning of these avirulence genes laid the foundations for the field of oomycete "effectoromics".

Recent advances in DNA sequencing technology has resulted in the genomes of multiple plant pathogenic organisms being sequenced, which significantly advances the understanding of plant-pathogen effector biology by cataloging potential effectors (Spanu 2012; Kemen & Jones 2012; Raffaele & Kamoun 2012). In the following part of this introduction the focus will primarily be on oomycete genomes with special emphasis on *Phytophthora infestans, Hpa, Phytium ultimum*, and *Albugo laibachii* (Haas et al. 2009; Levesque et al. 2010; Baxter et al. 2010; Kemen et al. 2011; Links et al. 2011).

Within the oomycetes there does not seem to be a correlation between total genome size and choice between biotrophic or necrotrophic lifestyle. However, obligate biotrophy has been associated with a lack of certain genes that encode for different classes of metabolic pathways, such as sulphate assimilation and other molybdopterin cofactor-requiring biosynthetic pathways. Further a loss of some classes of plant cell wall hydrolases has been reported. Given these events have occurred several times independently of each other within the oomycetes, and also fungi, it is most likely that the gene losses are due to convergent adaptation to plant parasitism (Raffaele & Kamoun 2012).

### 1.4.2 Oomycete effectors are modular proteins

#### *1.4.2.1 RXLR effectors*

All oomycete avirulence genes cloned to date have been cloned from either *Hpa* or *Phytophthora* species. In *Hpa* several *Arabidopsis thaliana* recognized (ATR) genes have been cloned: ATR1, ATR5 and ATR13 (Rehmany et al. 2005; Bailey et al. 2011; R. Allen et al. 2004). However, the majority of avirulence genes identified have come from different *Phytophthora* species. The best-characterized *Phytophthora* effectors are: *Avr1b* from *P. sojae,* in addition to *Avr3a* and *Avrblb2* from *P. infestans* (W. Shan et al. 2004; Armstrong et al. 2005; Oh et al. 2009; Bozkurt et al. 2011).

Inspection of the amino acid sequences of these avirulence genes reveal some defining features: a signal peptide directing the protein into the ER, and a positionally constrained RXLR motif, which is found within the first 80 amino acids of the protein and typically followed by a stretch of acidic (D/E) amino acids (Win et al. 2007). Using these features as criteria RXLR effectors have been predicted in oomycetes.

Within the *Peronosporales* the *Phytophthora* species have more than 400 RXLR effectors encoded in the genome (370 *P. ramorum*, 385 *P. sojae,* and 563 *P. infestans*), whereas 134 RXLR containing secreted proteins are predicted in *Hpa* (Baxter et al. 2010). Thus, it has been speculated that the reduced number of RXLRs reflects host range; *Hpa* has only been reported on *Arabidopsis*, whereas *Phytophtora* sp. have an extended host range (Slusarenko & Schlaich 2003).

The 240 Mb *P. infestans* genome is the largest oomycete genome sequenced to date. The expansion in genome size compared to *P. ramorum* and *P. sojae* is mainly due to a proliferation of repeats, which make up 74% of the genome. There are gene dense and gene sparse regions in the genome (Haas et al. 2009). Interestingly, effector genes predominantly reside in the gene sparse regions, which are rich in transposable elements

and repeated sequences. Another defining trait of gene sparse regions is that they are likely to represent syntenic breakpoints, which are rapidly evolving parts of the genome. Corroborating this, less than 25% of the RXLR effectors pass tests for orthology between *P. sojae*, *P. ramorum* and *P. infestans* (Tyler *et al*., 2006, Tyler, 2009). A comparative study of the RXLR effectors of *Hpa*, *P. ramorum* and *P. sojae* revealed that the C-terminal region of the effectors had a higher probability of being under positive selection evident as a higher ω (the rate ratio of non-synonymous/synonymous substitution in codons) values compared to N-terminal region of the gene (Win et al. 2007).

A complementary study based on recursive BLAST similarity searching and hidden Markov modeling predicted the RXLR effector complement of *P. ramorum* and *P. sojae* to be approximately 700 members large and dominated by a single superfamily, which include all experimentally verified effectors (Jiang et al. 2008). While the C-terminal region has been considered to be variable, evidence was found of modules/motifs repeated up to eight times termed W, Y and L. The recent elucidation of the 3D structures of the C-termini of the RXLR effectors PexRD2 (putative extracellular protein RD2) and Avr3a11 from *P. infestans* and *P. capsici* respectively revealed that W, Y and L motifs are structural elements essential for formation of a unique WY-fold (Boutemy et al. 2011). Interestingly, the WY-fold creates a stable structural scaffold that is the core of RXLR effectors. This led Boutemy and coworkers to speculate that sequence-unrelated RXLR effectors containing W, Y, and L motifs could have a core WY-fold. The structures of ATR13 (*Hpa*) and Avr3a4 (*P. capsici*) corroborates this notion (Yaeno et al. 2011; Chou et al. 2011). Inclusion of the N-terminal RXLR region resulted in non-diffracting crystals suggesting that this region is unstructured. This was confirmed by the NMR structure of ATR13 which shows that the RXLR region is intrinsically unstructured (Leonelli et al. 2011).

It is important to note that not all RXLR effectors are predicted to have WY-domain folds (Jiang et al. 2008; Boutemy et al. 2011). For example, no WY-fold is predicted for ATR13 and the NMR structure shows that the structure instead consists of a helical domain and a disordered loop (Leonelli et al. 2011).

Secreted proteins with an RXLR-like motif are also found in other species. A survey of approximately 14% of the genome of *Pseudoperonospora cubensis* showed no evidence of RXLR + dEER motif containing secreted proteins. However 29 secreted proteins had high homology to RXLR effectors from *Hpa* and *Phytophthora* sp. Surprisingly, these carry a QXLR motif due to a substitution of an arginine to a glutamine and a dEER stretch of acid amino acids (Tian et al. 2011). As this family is under diversifying selection and is up regulated during plant infection, QXLRs could be bona fide *Pseudoperonopora cubensis* effectors.

While there is considerable evidence suggesting that RXLR type effectors are predominantly expanded in the *Peronosporaceae*, they are not statistically overrepresented in the secretome compared to the proteome of the *Albuginales* and *Phytiales* sequenced to date (Kemen et al. 2011; Links et al. 2011; Levesque et al. 2010). It is therefore unlikely that RXLR effectors play a central role in virulence in these species. However, an elicitin with homology to an RXLR effector was identified in the fish pathogen *Saprolegnia parasitica* (van West et al. 2010). It will therefore be interesting to determine if RXLR effectors are expanded in this organism, as this could suggest an independent adoption of RXLRs as a major class of virulence proteins.

#### *1.4.2.2 Crinkler effectors*

Crinklers comprise another prominent family of effectors, which initially was identified in a bioinformatic screening of a *P. infestans* EST collection for secreted proteins. Two

proteins CRN1 and CRN2 were found to cause leaf-crinkling and tissue necrosis (crn or crinkler) accompanied by induction of defense related genes when expressed *in planta* (Torto et al. 2003). The crn protein family, including CRN1 and CRN2, includes a diverse set of proteins that range from 450-850aa in size (Torto et al. 2003). Approximately 196 *crn* genes and 255 pseudogenes are encoded in the *Phytophthora infestans* genome (Haas et al. 2009). A genome wide study of crn gene evolution revealed that the large diversity of crn genes is due to extensive recombination occurring between different N- and C-terminal modules (Haas et al. 2009). Therefore recombination may be the major driving force behind crns. A similar type of effector evolution has been reported for phyto-pathogenic bacteria, where effectors evolve through a shuffling process termed terminal reassortment (Stavrinides et al. 2006).

Sequence analyses have shown that crns have a modular structure that is analogous to RXLRs. The crns have a hypervariable C-terminal module and a more conserved N-terminal region (Torto et al. 2003; Haas et al. 2009). A defining feature of the N-terminal region of crns in *Phytophthora* is the presence of a secretion peptide and a ~50 amino acid LXLFLAK (thus: leucine, any amino acid, leucine, phenylalanine, leucine, alanine and lysine) domain within the first 60 amino acids followed by an adjacent DWL (Asparagine, Tryptophan and Leucine) motif (Torto et al. 2003; Win et al. 2007; Haas et al. 2009). In contrast to RXLRs the end of the N-terminal domain is flanked by a highly conserved HVLVxxP motif, which marks the beginning of a variable C-terminal domain. As observed for RXLR type effectors the N-terminal domain LXLFLAK is required for translocation into the plant cell (Schornack et al. 2010).

To date, 37 conserved C-terminal regions have been defined in addition to 8 unique C-termini. Transient expression of representative C-terminal domains in *N. benthamiana* inside the plant cell reveals that four domains (DC, DBF, D2 and DXW-DXX-DXS)

trigger cell death. Interestingly, three of these domains, specifically DC, D2 and DBF, have a striking homology to protein kinases. Kinase activity has been demonstrated for CRN8 that belongs to the D2 family; interestingly, kinase activity is required for cell death induction in addition to nuclear localization (van Damme et al. 2012). The nucleus seems to be the major compartment targeted by crns, as 5 out of 5 crn-domain types tested for subcellular localization in *N. benthamiana* transient assays were found to be nuclear localized (Schornack et al. 2010).

Homology based BLAST searches reveal evidence of crn families in the obligate biotrophic species *Albugo* and *Hpa* but also in the necrotrophs *Pythium ultimum* and *Aphanomyces euteiches* although the crns are not completely conserved (Links et al. 2011; Kemen et al. 2011; Win et al. 2007; Cheung et al. 2008; Levesque et al. 2010; Schornack et al. 2010). This suggests, that contrary to their name, crns may also be required to suppress host defense in biotrophic pathogens.

Furthermore, as crns are absent from the diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum* crns could be an old effector class that evolved in oomycetes (Levesque et al. 2010).

### 1.4.3 Extracellular effectors

Apart from effectors translocated inside the plant cell, another class of fungal and oomycete effectors exist, which functions in the plant apoplast. Apoplastic effectors were initially identified in the tomato leaf mould pathogen *Cladosporum fulvum*, which exclusively colonizes the apoplast. The functions of apoplastic effectors are diverse and include: binding soluble chitin oligomers or chitinous pathogen cell wall, to inhibition of apoplastic host proteases (de Jonge & Thomma 2009; Hein et al. 2009). *C. fulvum* secretes two effectors Ecp6 (extracellular protein 6 ) and Avr4 into the apoplast, where they function to protect the fungus from detection and the fungal cell wall from degradation by

the host (de Jonge et al. 2010). Avr4 contains an invertebrate chitin-binding domain, which has affinity for chitin found in the cell wall of the fungus. Thus, Avr4 contributes to virulence by shielding the fungal hyphae from hydrolysis by chitinases secreted from the plant into the apoplast (van den Burg et al. 2006).

Ecp6 is able to bind chitin but is unable to protect fungal hyphae against degradation by host enzymes. Instead the three LysM domains of Ecp6 are thought to compete with host PRRs for soluble chitin (de Jonge et al. 2010). Thus, the role of Ecp6 is to sequester chitin oligosaccharides that are released from the fungal hyphae in order to evade recognition by PRRs and induction of PTI. Recently it was reported that the apoplastic effector Secreted LysM Protein1 from *Magnaportha oryzae* is required for suppression of CEBiP (chitin oligosaccharide elicitor- binding protein) induced PTI (Mentlak et al. 2012). Thus, this mechanism is conserved between various fungi.

To inhibit host proteases *C. fulvum* secretes protease inhibitors into the apoplast such as Avr2 that inhibits the host cysteine protease RCR3 (required for Cladosporium fulvum resistance 3) (Jing Song et al. 2009).

Oomycete genomics have revealed similar apoplastic proteins are encoded in the genomes, but most apoplastic oomycete effectors biochemically characterized to date are from *Phytophthora* sp. (Baxter et al. 2010; Tyler et al. 2006). Similar to *C. fulvum Phytophthora infestans* secretes protease inhibitors; two cystatin-like proteins EPIC1 (extracellular protease inhibitor with cystatin- like domain 1) and EPIC2B are secreted into the apoplast. Initially they were found to inhibit the function of the apoplastic proteases PIP1 (Phytophthora inhibited protease 1) and RCR3 from tomato (Tian et al. 2007; Jing Song et al. 2009). Subsequent work showed that EPIC1 and EPIC2B also inhibit the papain-like cysteine protease C14 from potato (Kaschani et al. 2010).

Interestingly, the binding affinity of EPIC1 and EPIC2B to C14 is much higher than to RCR3 and PIP1 suggesting that EPIC1 and EPIC2B have evolved to target C14 (Kaschani et al. 2010). Corroborating this notion tomato RCR3 is under purifying selection consistent with the *Phytophthora* sp. being atypical pathogens (Jing Song et al. 2009).

In contrast, *C14* from potato is under diversifying selection in wild potato (*Solanum demissum*, *Solanum verrucosum*, and *Solanum stoliniferum*) consistent with a participation in an evolutionary arms race. Structural homology modeling suggests that the residues under positive selection are located around the substrate-binding groove, which also is the EPIC-C14 interaction site (Kaschani et al. 2010; Kaschani & van der Hoorn 2011). These data suggest that the susceptibility target C14 that contributes to immunity is under diversifying selection to avoid inhibition by EPICs, but only in the natural host, potato.

Not all apoplastic effectors directly target the host defense machinery; Nep1-like proteins (NLPs) cause necrosis and are induced at the onset of the necrotrophic phase of *Phytophthora* sp. (Gijzen & Nürnberger 2006; Qutob et al. 2006). In contrast to NLPs from *Phytophthora* sp. NLPs from *Hpa* are mainly expressed at early stages of infection and do not cause necrosis (Cabral et al. 2012; Baxter et al. 2010).

Apart from these more well defined classes of apoplastic effectors other classes exists such as elicitins and other small cysteine rich proteins are also secreted into the apoplast but the function of these are unknown (Hein et al. 2009).

Pathogens also secrete cell wall degrading enzymes such as hydrolases. The number of genes belonging to this class encoded in the genome correlates with lifestyle, thus in necrotrophs many cell wall degrading enzymes are encoded in the genome, whereas only a few are found in biotrophic oomycetes (Baxter et al. 2010; Kemen et al. 2011; Levesque et al. 2010).

## 1.5 Biochemical functions of effectors inside the plant cell

Baxter et al. (2010) predicted 134 high confidence RXLRs that included ATR1 and ATR13 but did not assign or test for function. Fabro et al (2010) screened 64 *Hpa* RXLR effectors individually on 12 *Arabidopsis* accessions by using the EDV system to deliver the effectors via the T3SS of *Pst* DC3000 into the plant cell (Fabro et al. 2011; K. H. Sohn et al. 2007b). The relative contribution of the effector to virulence was measured as increased or decreased growth of *Pst* DC3000 compared to strains expressing a non-functional version of AvrRps4 (K. H. Sohn et al. 2009). The major finding was that 70% of the effectors caused enhanced virulence on at least one *Arabidopsis* ecotype mostly by suppressing PTI. This suggests that *Hpa* presumably utilizes multiple effectors with weak accession-specific effects to create a suitable environment for host colonization. Consistent with *Hpa* being a non-adapted pathogen of *Brassica rapa,* assays revealed that most effectors caused no or a reduction in virulence of *Pst* DC3000. Thus, effectors have been evolutionarily shaped to specifically suppress distinct host processes and to avoid recognition limited by pathogen host range.

A subsequent study tested the subcellular localization of the *Hpa* RXLR effectors and found that most plant compartments are targeted by at least one effector. Interestingly, the majority of effectors localize in the nucleus: 33% are nuclear, and another 33% have a nuclear/cytoplasmic distribution. The remaining effectors predominantly target the plant membrane network including the ER (Caillaud, Piquerez & Jones 2012a). Thus in general, *Hpa* RXLRs mainly target the nucleus and the membrane network (Lindeberg et al. 2012).

Interestingly, a protein-protein interaction screen revealed that some host proteins are targeted by both bacterial and oomycete effectors. Specifically, a yeast two-hybrid interaction screen of approximately 8000 full length *Arabidopsis* proteins against bacterial T3S effectors from *Pseudomonas syringae* and *Hpa* RXLR effectors concluded that

effector proteins converge onto a set of highly interconnected host proteins, termed hubs (Mukhtar et al. 2011). These hubs may play important roles in virulence, as T-DNA knockout of 15 out of 17 hub genes resulted in enhanced disease phenotypes upon infection with bacterial and oomycete pathogens. In general the work suggests that bacterial and eukaryotic pathogens target a limited set of host proteins to suppress host immunity.

The aforementioned screens have greatly expanded our knowledge on effectors. However the precise mechanisms by which effectors suppress immunity are poorly understood. The functional characterization of bacterial effectors is most advanced. However, a *Phytophthora infestans* effector Avr3a has been studied in detail (Armstrong et al. 2005; Bos et al. 2010; Whisson et al. 2007; Bos et al. 2009; Yaeno et al. 2011). Knockdown of Avr3a results in almost complete loss of virulence on both potato and *N. benthamiana* leaves suggesting it to be a major virulence factor of *Phytophthora* sp. (Bos et al. 2010). Mechanistically, Avr3a is a secreted RXLR type effector that is translocated into the host cytoplasm where it interacts with the host E3 ligase CMPG1 (Whisson et al. 2007; Bos et al. 2010). The host protein CMPG1 is required for cell death execution upon recognition of the PAMP-like protein INF1, which is one of the most abundant *Phytophthora* proteins (Kamoun et al. 1997). Bos et al. (2010) found Avr3a stabilizes CMPG1, which results in a suppression of plant cell death upon INF1 stimulus. Plants that carry the resistance gene *R3a* undergo HR upon infection with a *Phytophthora infestans* carrying a recognized allele of Avr3a (Armstrong et al. 2005).

Thus Avr3a can trigger two responses: R3a-dependent HR, and suppression of INF1 dependent PCD. Interestingly, these dual roles of Avr3a can be uncoupled at a structural level. Two major alleles of Avr3a exist, which only differ in two amino acids but have significantly different functions. The $Avr3a^{KI}$ (K80/I103) allele triggers R3a-dependent HR

and is a strong suppressor of INF1 cell-death. In contrast, the Avr3a$^{EM}$ (E80/M103) allele avoids recognition by R3a but is a weaker suppressor of INF1 cell-death (Bos et al. 2006). Further a C-terminal tyrosine-147 is absolutely required for suppression of INF1 cell-death by both alleles; deletion or mutation of this residue results in loss of inhibition via the loss of interaction with CMPG1 (Bos et al. 2009; Bos et al. 2010). Yeano et al (2012) provide evidence that loss of CMPG1 interaction is correlated with loss of binding to immobilized PI3P spots on nitrocellulose membranes. Thus, the current model for Avr3a function is that inside the host cell Avr3a binds to phospholipids, which guides it into a compartment, where it can interact with and stabilize CMPG1 resulting in suppression of INF1 dependent cell death.

Most effectors cause small increases in virulence on a given plant accession, but some effectors are recognized by the plant immune system such as Avr3aKI by R3a. Systematic screens for avirulence have identified 8 effectors as potential avirulence genes by screening 270 candidate *Phytophthora infestans* effectors on 17 potato accessions (Rietman 2011). Given 4590 interactions were screened, avirulence is a rare phenomenon. Corroborating this a test of 18 RXLR effectors from *Hpa* Emoy2 on 83 *Arabidopsis* ecotypes from the Nordborg collection identified a single interaction that resulted in HR out of 1494 binary interactions screened (Goritschnig et al. 2012). Specifically, ATR39-1 is recognized by RPP39 in Wei-0.

As approximately half of the Nordborg lines are resistant to *Hpa* Emoy2, recognition of single effectors resulting in a HR is a rare event although resistance is widespread (Nemri et al. 2010; Krasileva et al. 2011).

## 1.6 Plant surveillance of host processes

### 1.6.1 Mechanisms of R-protein function

Plants employ resistance genes (*R*-genes) to detect the presence of pathogen-secreted effectors. Plant R-genes have been found against bacteria, fungi, nematodes, oomycetes and viruses, but they all belong to one of only five different classes of *R*-genes (Dangl and Jones, 2001). Most *R*-genes encode proteins of the nucleotide-binding-site and leucine rich repeat (NB-LRR) class. This class of proteins is highly polymorphic within plant populations (Guo et al. 2011). Sequence alignments show that the most variable region between different NB-LRR proteins is the C-terminal LRR domain, which may have protein-protein, protein-carbohydrate, and peptide-ligand interaction properties. The C-terminal LRR domain is the major, but not exclusive determinant for substrate specificity (Maekawa, Cheng, et al. 2011a).

The NB domain belongs to the Signal Transduction ATPases with Numerous Domains (STAND) subclade of AAA-ATPase superfamily; NB domains are able to bind and hydrolyze ATP (Bonardi et al. 2012; Tameling et al. 2002). The function of the NB domain is to act as a molecular on/off switch upon recognition of a pathogen effector. The NB domain is inactive in an ADP-bound state. Upon pathogen recognition ADP is exchanged for ATP resulting in a conformational change into an active state leading to downstream signaling. Hydrolysis of ATP results in return to the inactive state (Bonardi et al. 2012).

This molecular switch is central to plant immunity (Yuelin Zhang et al. 2003; Shirano 2002). A D555V mutation in the flax rust resistance gene M result in auto-activation due to enhanced ATP binding (Williams et al. 2011). This auto-activation results in constitutive activation of defenses and ultimately in complete tissue death.

Based on the N-terminal region the NB-LRRs can be further separated into two groups (Dangl & Jones 2001). In *Arabidopsis* the largest group shows homology to the N-terminus of *Drosophila* Toll and mammalian interleukin (IL-1) receptor (TIR-NB-LRR), the second group has a coiled-coil (CC-NB-LRR). These domains likely engage downstream signaling components (Dangl and Jones 2001, detailed later). For example, expression of the TIR domain of the TIR-NB-LRR RPS4 (*P. syringae* resistance 4) result in ectopic cell death via an enhanced disease susceptibility 1 (EDS1), Suppressor of G-two allele of skp1 (SGT1), heat shock protein 90 (HSP90) dependent pathway (Swiderski et al. 2009).

Interestingly, dimerization of CC- or TIR-domains is required for downstream signaling. The recent crystal structures of the N-terminal regions of the CC-NB-LRR MLA-10 and of the TIR-NB-LRR L6 revealed these regions to form homo-dimers (Bernoux et al. 2011; Maekawa, Cheng, et al. 2011a). Point mutations in residues critical for dimer formation underline the importance of homo-dimerization for downstream signaling, as both L6 and MLA10 dimerization-mutants are unable to trigger PCD.

Other mechanisms for activation exist. For example the ADR1 (ACTIVATED DISEASE RESISTANCE 1) class of NB-LRRs are required for downstream activation of *lsd1* (lesions simulating disease 1) induced PCD. Interestingly, this function is independent of ATP binding as a mutant in the ATP binding pocket is still functional. Consequently, this class of NB-LRRs has been hypothesized to function as scaffolds ensuring correct docking of downstream signaling components (Bonardi et al. 2012; Bonardi et al. 2011).

Numerous studies of NB-LRR localization within the plant cell have revealed that they can be either membrane associated, nuclear or cytoplasmic. Some NB-LRRs are localized at membranes, for example the CC-NB-LRR RPM1 (resistance to Pseudomonas syringae pv. maculicola 1) is plasma membrane-localized before and during activation (Gao et al.

2011). For other NB-LRRs such as RPS4 a nuclear localization has been shown to be required to trigger EDS1 dependent signaling (Wirthmueller et al. 2007).

The nucleus has been further implicated in NB-LRR activation, as a screen for suppressors of suppressor of NPR1 constitutive 1 (SNC1) identified the nuclear pore complex components AtNup88/MOS7 (modifier of SNC1) and AtNup96/MOS3 as suppressors (García & Parker 2009). Nuclear import presumably requires the importin α3/MOS6 as this was found as a suppressor of *snc1* (Palma et al. 2005).

While nuclear-cytoplasmic partitioning is essential for activation of defense the exact mechanisms are poorly understood and further work is required to determine these events. Given some resistance genes have C-terminal WRKY domains this implicates them in transcriptional regulation (Narusaka et al. 2009).

### 1.6.2 The guard hypothesis

Upon infection effectors are secreted into the plant cell, where they execute functions to suppress host immunity. In some cases they are recognized or perturb a process monitored by a NB-LRR resulting in its activation. The simplest model describing the genetic relationship between *avr* and *R*-gene is through a direct interaction between an avr effector and a host R-protein (Jones & Dangl 2006; Dodds & Rathjen 2010). However data supporting this model are limited, and only four reports describe a direct interaction between the avr- and R-protein (Jones & Dangl 2006; Krasileva et al. 2010). For example, Dodds and coworkers (Dodds et al. 2006) have shown a direct interaction between avrL567 from the flax rust *Melampsora lini* and the corresponding flax R-protein L6. Likewise, the *Hpa* RXLR effector ATR1-Emoy2 is recognized by a direct interaction with the R-proteins RPP1-WsB and RPP1-NdA (Rehmany et al. 2005; Krasileva et al. 2010). This direct interaction has led to an evolutionary arms-race between recognition and

evasion of detection, which has resulted in the diversification ATR1 and RPP1 alleles (Rehmany et al. 2005; Botella et al. 1998).

In general the literature supports a model, where Avr-R protein recognition is indirect (Jones & Dangl 2006; Dodds & Rathjen 2010; Dangl & Jones 2001). In such a model, R-proteins detect modifications of plant proteins by effectors: this indirect recognition would in principle allow one R-protein to "guard" against multiple pathogen effectors targeting the same plant protein. The principle has been formulated as the "guard hypothesis", which resembles the mammalian immune systems "modified self" theory. This strategy nullifies the evolutionary advantage faster evolving pathogens have, as the host uses the effector's virulence mechanism to trigger the recognition. Since its proposal, ample support for the model has been obtained and it has been the subject of comprehensive reviews (Dodds & Rathjen 2010; Dangl & Jones 2001; Chisholm et al. 2006). The plant pathogen interactions revolving around the *Arabidopsis* protein RIN4 (RPM1 interacting protein 4) provide a prominent example illustrating the guard hypothesis.

The RIN4 is a small (211aa), plasma membrane localized protein that negatively regulates PAMP induced defenses (J. Liu et al. 2009). *Pseudomonas syringae* exploit this by secreting the two effectors AvrB and AvrRpm1 that modulate RIN4 by phosphorylation (Kim et al., 2005, Mackey et al., 2002). However, two CC-NB-LRR proteins, resistance to *Pseudomonas syringae* pv. *syringae* 2 (RPS2) and RPM1 that co-localize and interact directly with RIN4 at the plasma membrane *in planta*, act as guards.

The host resistance protein RPM1 recognizes the phosphorylation of RIN4 and induces effector triggered immunity (ETI). A third effector protein, AvrRpt2, also targets RIN4 and can abolish RPM1 recognition of hyper-phosphorylated RIN4 (Mackey et al., 2003). AvrRpt2 is a cysteine protease that contributes to virulence by cleaving the C-terminus of membrane bound RIN4 and 19 other host proteins with the consensus site VPxFGxW

(Chisholm et al. 2005). However, the disappearance of RIN4 is detected by the resistance protein RPS2 leading to ETI.

Since the *rin4* mutant is lethal, but the *rin4/rpm1/rps2* mutant is viable, it can be said that RIN4 negatively regulates RPM1- and RPS2- induced HR (Belkhadir et al., 2004; Mackey et al., 2002). Although it is important in regulating *RPM1* and *RPS2*, *RIN4* is expendable for virulence functions of *AvrRpm1* and *AvrRpt2* in disease-susceptible *rin4/rpm1/rps2* mutant plants, which suggests the existence of other host targets (Belkhadir et al., 2004). In line with this notion, AvrRpt2 has been shown to cleave other proteins (Chisholm et al., 2005; Shang et al., 2006).

### 1.6.3 Decoy model

The decoy model is the most recent elaboration on the guard hypothesis (Van der Hoorn and Kamoun, 2008). The rationale behind this is, since plant R-proteins are highly polymorphic in the natural plant population "guardees" are under opposing selection forces depending on the presence or absence of the associated R-protein, and thus exist in an unstable evolutionary state. In the presence of the cognate R-gene, natural selection would favor plants evolving a stronger interaction between effector and the guardee, whereas the opposite is true in absence of the R-gene.

A decoy with no other function than specializing in perception of effectors by R-proteins would relax this evolutionary disequilibrium. While no population genetic studies have analyzed the diversity of the proposed decoys yet, examples based on individual ecotypes exist (Stukenbrock and McDonald, 2009). In tomato, the effectors AvrPto and AvrPtoB targeting PRRs illustrate the model.

*Pseudomonas syringae* harboring AvrPto or AvrPtoB are avirulent on tomato plants containing the genes *Pto* and *Prf* (*Pseudomonas* resistance and Fenthion sensitivity) (Kim

et al., 2002; Xiang et al., 2008). Pto has been found to interact with Prf in planta (Loh et al., 1998). To trigger Prf- dependent ETI AvrPto interacts directly with Pto (Mucyn et al., 2006, Chang et al., 2000).

The AvrPtoB protein has ubiquitin ligase activity (Janjusevic et al., 2006, Abramovitch et al., 2006) and has been found to target Pto and also the related kinase Fen (fenthion) for degradation (Abramovitch et al., 2003, Rosebrock et al., 2007).

Further, a screen for PAMP suppressors based on the inhibition of *FRK1* transcription upon flg22 treatment in *Arabidopsis* identified AvrPto and AvrPtoB as potent suppressors. Since MPK3 (MAP kinase 3) and MPK6 (MAP kinase 6) signaling were suppressed, though not abolished, this indicated that AvrPto and AvrPtoB target early steps in flg22 PAMP signaling (He et al., 2006). Further studies found that AvrPto interacts directly with EFR and FLS2 to inhibit their autophosphorylation and thus activation PTI (Xiang et al., 2008). Similarly, AvrPtoB was found to target the PPRs EFR, FLS2 and CERK1 for degradation (Göhre et al. 2008; Gimenez-Ibanez et al. 2009). However, the physiological relevance of these PRR-Avr interactions has been put in question, because under limited expression of AvrPto FLS2 and BAK1 in protoplasts AvrPto interacts only with BAK1 (Shan et al. 2008). Supporting this notion, stable overexpression lines of AvrPto and AvrPtoB in *Arabidopsis* phenotypically resemble *bak1* and weak *bri1* (brassinosteroid-insensitive-1) mutants (Shan et al. 2008).

From an evolutionary perspective this host-pathogen interplay can be interpreted as follows. AvrPtoB suppresses MAMP- induced basal defense by inhibition of PRRs and BAK1. To counter this, tomato plants evolved a decoy, the Fen kinase, conferring resistance to *AvrPtoB* mediated by the *Prf* R gene. As a consequence, a new AvrPtoB allele evolved with the ability to induce ubiquitination and degradation of Fen, once again

allowing infection of host plants. Finally, the host plants developed a Fen related kinase, Pto, which again allowed resistance towards bacteria expressing *AvrPtoB* by activating Prf mediated defense responses.

### 1.6.4 The zig-zag-zig model

The two previous examples also illustrate another important aspect of plant-pathogen interactions, which has been formulated into the zig-zag-zig model (Jones and Dangl, 2006). This model explains how small differences can explain why plants are susceptible or resistant to different pathogens. Pathogens reveal PAMPs to their host, which are recognized by PRRs resulting in PTI. In some cases, the pathogen inhibits PRR function by effector secretion leading to effector-triggered susceptibility (ETS). These effectors will sometimes be recognized by plant R-proteins, which will induce ETI. Since pathogens evolve with their hosts, they may secrete other effectors countering the ETI effect of the R-protein resulting in ETS. The model could be summed up in the equation,

basal resistance = $\Sigma$(PTI) - $\Sigma$(ETS) + $\Sigma$(ETI).

The model is generally applicable. For example in *P. infestans* the apoplastic protein INF1 induces a BAK1 dependent PTI (Hein *et al*, 2009), which is suppressed by the active allele of the RXLR effector Avr3a. In a "final zig", potato plants carrying the R-gene R3a detect Avr3a resulting in ETI (Bos *et al*., 2006). In other less studied plant/pathogen systems only the cognate R-genes are known. These pathogens provide an interesting opportunity to further validate the model(s) and gain new insights into virulence and defense.

### 1.6.5 Events downstream of R-protein recognition

Events following ETI by activation of NB-LRRs are linked to a dramatic release of $Ca^{2+}$ into the cytosol and accumulation of reactive oxygen species followed by transcriptional reprogramming normally resulting in localized programmed cell death (PCD) (Lam et al. 2001; Maekawa, Cheng, et al. 2011a). The entire pathway resulting in PCD is currently not completely understood (Kemen et al. 2011; Aarts et al. 1998; Glazebrook 2005; Spoel & X. Dong 2012; Maekawa, Kufer, et al. 2011b).

To induce resistance, both RPS2 and RPM1 have been shown to depend on *NON RACE-SPECIFIC DISEASE RESISTANCE* (*NDR1*) that generally is required for CC-NB-LRR R-protein induced resistance (Century et al. 1997). Interestingly, NDR1 was recently shown to play a role at the plasma membrane-cell wall junction in maintaining membrane integrity in response to biotic and abiotic stimuli (Knepper et al. 2011).

TIR-NB-LRR signaling requires EDS1 (Aarts et al. 1998). EDS1 is a nucleo-cytoplasmic protein with homology to eukaryotic lipases. It interacts with phytoalexin deficient 4 (PAD4) and senescence-associated protein 101 (SAG101) and controls the amplification and production of SA-related signals during TIR-NB-LRR induced resistance (Feys et al. 2005; Wiermer et al. 2005). Interestingly, the loss of TIR-NB-LRR mediated defense phenotype of a *pad4*/*sag101* double mutant is more severe than in *eds1* or either mutant alone, suggesting that EDS1 function could act as a scaffold for correct PAD4-SAG101 function (Wiermer et al. 2005). Genetically, all three proteins function upstream of SA signaling, but their transcription is also induced by exogenous application of SA, suggesting the existence of an SA amplification loop (Falk et al. 1999; Jirage et al. 1999).

Nevertheless, other pathways do exists as some CC-NB-LRRs are able to activate defense independently of *EDS1* and *NDR1* (Aarts et al. 1998; McDowell et al. 2000; Bittner-Eddy et al. 2000).

## 1.7 Origin of the oomycetes

Oomycetes are able to reproduce both sexually and asexually. They are fungus-like heterotrophs that are either saprophytes or parasites. Oomycetes together with brown algae, diatoms and *Plasmodium falciparum* belong to the clade of stramenopiles. Fungi belong to the clade of opisthokonts. While morphologically both oomycetes and fungi are filamentous eukaryotic microbes, defining differences exists (Kemen & Jones 2012). The cell wall of oomycete microbes consists predominantly of cellulose-based polymers generally low in chitin (Beakes & Sekimoto 2009). The oomycete clade contains pathogens with a wide host range from humans and fish to plants (Thines & Kamoun 2010). Some of the most devastating plant diseases, such as sudden oak death and potato late blight, are caused by oomycete pathogens. Oomycete plant pathogens adopt diverse pathogenic lifestyles from completely biotrophic to necrotrophic. Necrotrophic pathogens such as *Pythium ultimum* kill their host and extract nutrients from the dead tissue comparable to saprophytes (Cheung et al. 2008). In contrast, obligate biotrophs are dependent on their host to complete their life cycle and thus suppress host responses and in some cases prolong the life of the host tissues (Cooper et al. 2008).

Given the intimacy of this interaction biotrophs are highly adapted to their host and this relationship may shape both host plant and oomycete evolution (Spanu 2012). Interestingly, biotrophic plant parasitism has independently evolved at least three times within the oomycota; once in the *Saprolegniales* and once in *Peronosporales* and once in *Albuginales* (Thines & Kamoun 2010). The biotrophic pathogens are: *Aphanomyces euteiches, Albugo* sp., and *Hpa*.

Between these two extremes are pathogens that have an initial biotrophic phase followed by a necrotrophic phase, which are called hemi-biotrophs. In the hemi-biotroph *P.*

*infestans* two distinct sets of genes govern each phase (Haas et al. 2009). Interestingly, biotrophic infection of *Brassica* plants by *Albugo* species (sp.) is almost symptomless. *Albugo laibachii* parasitizes *Arabidopsis* and is a remarkably effective pathogen as susceptibility is widespread.

### 1.7.1 Albugo sp.

The *Albugo* genus is widely distributed and infects a range of dicotyledonous host plants. The most economically important hosts are crucifers such as cabbage, turnip and broccoli on which *Albugo* causes diseases known as white rust or white blister (Holub et al. 1995). *Albugo* species are thought to have broad host ranges (Y.-J. Choi et al. 2006; Voglmayr & Riethmüller 2006; Holub et al. 1995).

The *Brassicaceae* infecting clade is the best studied and has recently been found to consist of at least four distinct species based on *ITS* and *cox2* phylogeny (Thines et al. 2009; Y. Choi et al. 2008). Within this clade one branch is thought to have a broad host range and comprises the generalist species *A. candida*. Other, more defined branches specialize on a small host range, for example *A. laibachii*, whose only defined host is *A. thaliana* (Thines et al. 2009). Either the very broad host range of the different *Albugo* species might be more restricted than previously thought, or alternatively, yet undefined specialist species (like *A. laibachii*) exist within the *A. candida* generalist group.

### 1.7.2 The life cycle of *Albugo* sp.

Appearance of white blisters on host tissue is a characteristic feature of *A. candida* infection. The blisters are most likely formed by the outward pressure of sub epidermal unbranched zoosporangia growth extending from the mother cell. Ultimately these blisters

rupture and release zoosporangia into the surroundings. The mature zoosporangium is dehydrated and water is essential for germination. Each germinating zoosporangium releases 4-6 biflagellate, cell wall-less, zoospores through the operculum within 1 hour of rehydration. On *Arabidopsis* these motile zoospores are attracted to guard cells, where they shed their flagellum and produce a cell wall. The zoospores germinate by producing a germtube that extends through the stomatal opening into the stomatal chamber, where they produce an appressorium and penetrate a mesophyll cell growing the first haustorium. In the case of a susceptible interaction this leads to hyphal elongation with haustorial formation occurring at regular intervals in mesophyll cells. Colonization of cotyledons occurs within 48 hours.

*A. candida* is capable of replicating both sexually and asexually which provides means for an epidemic. Blisters that are evidence of asexual reproduction become evident 5 days post infection and are released after 7-9 days (Holub et al. 1995).

Sexual reproduction is characterized by production of spherical male and female gametangia at the hyphal tip. The male gametangium grows towards and penetrates the female oogonium fertilizing it. The fertilized oogonium matures into a golden brown oospore within 7 days. The oospores are released into the soil when the host tissue is degraded. This could provide an important overwintering mechanism in cold or dry climates.

### 1.7.3 Arabidopsis resistance to *A. candida* and *A. laibachii*

In general *Arabidopsis* ecotypes are susceptible to *A. laibachii*. A field survey of natural *Arabidopsis* populations over two consecutive seasons in southwest England found more than 25% were infected with *A. laibachii* (Holub et al. 1994). Corroborating this, an

experiment with ~300 ecotypes planted in Norwich and Cologne found only 15 ecotypes with resistance to *A. laibachii* isolate Nc14 (*AlNc14*) at both locations (Kemen et al. 2011).

Interestingly, resistance has been linked to R-gene mediated defense. The *A. laibachii* strain Acem1 (*Acem1*) has been found to trigger resistance, characterized as no blister formation, in Ksk-1 and Ksk-2 (Borhan *et al*., 2004). A cross between resistant Ksk-1 and the susceptible Wei-1 ecotype identified two resistance to *Albugo candida* (RAC) loci, named *RAC1* and *RAC3.* Interestingly, *RAC3* conferred semi-dominant resistance, whereas *RAC1* was inherited dominantly. A third weakly dominant resistance locus, *RAC2* that is independent of *RAC1* and *RAC3* was identified in a cross between Ksk-2 and Wei-1.

Map based cloning identified RAC1 as a TIR-NB-LRR protein. Proving that resistance to *Acem1* was conferred by *RAC1* alone, transgenic expression of *RAC1* conferred resistance in the susceptible ecotype Col-0. This resistance was dependent on *EDS1* but remarkably independent of *PAD4* and the salicylate hydroxylase *NahG* (Borhan et al. 2004).

The *Arabidopsis* ecotype Col-0 is resistant to a broad range of *A. candida* races. *White rust resistance 4* (*WRR4*) was identified by mapping for resistance to Acem2 on a $F_9$ Col-gl x Nd-1 inbred line. Introduction of WRR4 into susceptible Ws-3 plants introduced full resistance to *A. koreana ex. Capsella bursa-pastoris* (Race 4) and *A. candida* races isolated from *B. rapa* (Race 7), *B. juncea* (Race 2) and *B. oleracea* (Race 9) (Borhan et al. 2008). Likewise, Col-*wrr4* knockout plants were found to be susceptible to race 2 and 7. Thus, WRR4 confers resistance to a broad range of *A. candida* isolated from discrete hosts and like RAC1 it is a TIR-NB-LRR protein, implying R-proteins to be major players in *Arabidopsis* resistance to *Albugo* sp.

Since TIR-NB-LRR proteins reside inside plant cells this suggests that *Albugo* sp. like other oomycetes deliver effector proteins into host cells. It was the primary aim of this

PhD study to describe and characterize the effector complement of *Albugo laibachii.* To achieve this some R-genes mediating *Albugo* sp. resistance have been characterized.

## 1.8 Aims of this research

A crucial part of the oomycete infection strategy is to secrete proteins into plant cells where they interfere with host processes and suppress defenses. The biochemical function(s) of most effectors are poorly understood. Effectors that were identified, as avirulence genes, have been biochemically characterized in most detail.

We are primarily interested in how the plant innate immune system is suppressed during a compatible interaction. *A. laibachii* is a remarkably well-adapted pathogen on its host. This study focuses on the *Arabidopsis-Albugo laibachii* pathosystem. Interestingly, *A. laibachii* infection enables secondary infections on the host by pathogens normally resisted by the plant (Cooper et al. 2008). We infer that *A. laibachii* may have some remarkably powerful effectors that target core components of the host defense. It is the goal of this study to define the effector complement of *A. laibachii* and analyze how some effectors modulate host processes to enable host colonization by the pathogen.

In **chapter 3**, we sequence the genome of *AlNc14* using an Illumina sequencing strategy only (a group effort but primarily conducted by Eric Kemen). We define candidate effectors and test these for virulence activities. Our initial screening revealed that CHXC1 could enhance virulence of *Pst* DC3000 ΔAvrPto/ΔAvrPtoB. Intriguingly, CHXC1 has homology to HECT E3 ligases and is conserved in other oomycetes. In **chapter 4** we further describe and functionally characterize CHXC1 and hypothesize that it is a core oomycete effector.

Fast evolving species-specific effectors have been abundantly described in the literature. In chapter 3 we show that the SSP6 gene is encoded at a heterozygous locus and the proteins are candidate effectors. A follow up study in **chapter 5** shows that multiple variants of SSP6 exist and these are found in a region, which is absent from the *A. candida*. One variant, SSP6-2c is able to enhance virulence of *P. infestans* on *N. benthamiana* by suppression of PTI.

Eric Kemen collected Albugo isolates used in this study in the winter of 2007. In **chapter 6** we use these isolates to test if other types of isolate specific resistances are found in addition to those already reported. We find that known *RAC/WRR* genes confer resistance to *AlNc14* and *AlNc2*. Interestingly, we find that an additional *R*-gene is present in Col-5, which is active against *AcNc2*.

As genomics leads to the generation of many hypotheses that require extensive testing in the lab we describe the development of a method for a higher throughput screening system of *Hpa* in **Chapter 7.**

A part of the work presented in this thesis (Chapter 3) has led to a publication on the genome sequence of *A. laibachii* (Kemen et al. 2011), which can be found at the end of the thesis in the appendix.

# 2. General Materials and Methods

## 2.1 Plants

### 2.1.1 Plant growth

To synchronize germination the seeds were stratified for 1 week at 5$^{o}$C at a light intensity of 15-17 µmol/m$^2$/s. Subsequently the seedlings were grown at 75% relative humidity and a light intensity of 180 µmol/m$^2$/s at 22°C with a 10 hrs light and 14 hrs dark cycle. After 2

weeks the seedlings were transplanted into either P24 trays or FP9 pots containing an 8:1 mixture of F2 compost (Scotts Levington) and grit.

*Brassica oleracea* cv. Maris kestrel was grown in the same conditions as *Arabidopsis thaliana* though without stratification.

### 2.1.2 Seed sterilization

Up to 100µl seeds were aliquoted in 1.5 ml Eppendorf tubes, which were incubated in chloride gas for 6-12 hr created by mixing 96 ml bleach with 4 ml concentrated (37% w/v) HCl. This is toxic and thus was contained in a sealed dome put inside a fume hood.

### 2.1.3 *Arabidopsis* ecotypes and mutants used in this study

The following *Arabidopsis* ecotypes were used in this study: Col-0, Col-5, Ler-0, Nd-0, Ksk-1 and Ws-0. The *Arabidopsis* mutants used in this study are shown in Table 2.1. Mutant stocks were obtained from the European Arabidopsis Stock Centre (http://arabidopsis.info) and screened by PCR for presence of the T-DNA insert. The progeny of these lines were scored as homozygous for the T-DNA insertion if PCR reactions revealed the presence of a T-DNA insert band and absence of a genomic product spanning the integration site.

**Table: 2.1: List of *Arabidopsis* mutants used in this study.**

| Ecotype | mutant | Agi number | Reference |
|---|---|---|---|
| Col-0 | *arf9-1* | AT4G23980 | Robert-Seilaniantz *et al.*, 2010 |
| | *ein5-1* | AT1G54490 | Olmedo *et al.*, 2006 |
| Ws-0 | *eds1* | AT3G48090 | Parker *et al.*, 1996 |

### 2.1.4 Arabidopsis transformation

Arabidopsis transformants were created in house using the floral dip method (Clough & Bent 1998). Briefly, flowering plants were dipped in a suspension of *Agrobacterium tumefaciens* harboring a binary plasmid with the construct of interest. Subsequently the

plants were allowed to set seeds. $T_1$ seeds were harvested and selected on selective GM medium. Antibiotic resistant plants were analyzed for expression of the protein of interest either by western-blotting or fluorescence microscopy. Positive plants were transferred to soil and allowed to set seed. $T_2$ seeds were grown on selective GM media and the proportion of antibiotic resistant versus susceptible progeny were counted to identify single T-DNA insertions. $T_2$ lines where a quarter of the plants were susceptible to the antibiotic were considered to have a single insertion. Eight plants per $T_2$ line were grown and allowed to set seed. The $T_3$ progeny was tested for resistance and resistant lines were considered homozygous for the T-DNA insertion and used in subsequent experiments. Protein expression was confirmed in each generation by western-blot analysis or fluorescence microscopy. *Arabidopsis* transformants carrying the *bar* resistance gene were not selected on GM plates. Instead they were selected at the 4-leaf stage by spraying soil grown plants with a 100 mM aqueous solution of BASTA®. The transformants generated in this study are shown in Table 2.2.

**Table 2.2: List of *Arabidopsis* transgenics used in this study.**

| Ecotype | Transgene | Reference |
|---|---|---|
| Col-0 | 35S:GFP-ΔSP-CHXC1 | This thesis |
| | 35S:GFP-ΔSP-CHXC1(C651A) | This thesis |
| | estradiol:StrepII-3xHA-ΔSP-CHXC1 | This thesis |
| | estradiol:StrepII-3xHA-ΔSP-CHXC1(C651A) | This thesis |
| | estradiol:StrepII-3xHA-ΔSP-YFP | Fabro et al. 2011 |
| | 35S:GFP-ΔSP-SSP6-A | This thesis |
| | 35S:GFP-ΔSP-SSP6-2c | This thesis |
| | 35S:AFB1 | Robert-Seilaniantz *et al.*, 2010 |
| | 35S:GFP | This thesis |

### 2.1.5 *Agrobacterium tumefaciens* transient expression

Bacteria were streaked on L plates supplemented with appropriate antibiotics and grown for 2 days at 28°C. Single colonies were inoculated into 10 ml L broth with the appropriate

antibiotics and incubated overnight. Cells were harvested and washed twice in 10 mM $MgCl_2$ and prepared to a final $OD_{600}$ of 0.3 in 10mM $MgCl_2$. Transient expression was induced by pressure infiltrating the bacterial suspension into 4-week-old *N. benthamiana* plants. The 3rd and 4th leaf from the top were used for expression.

### 2.1.6 Reactive oxygen measurements

Up to 4 constructs per leaf were transiently expressed in the 3rd leaf of 4 week old *N. benthamiana* plants. After two days of incubation, 3 leaf discs (cork borer size 1) were sampled from each construct. Leaf discs were obtained from 8 different leaves resulting in a total of 24 leaf discs per construct. The leaf discs were incubated overnight at RT in 200 µl sterile milliQ water in a 96 well elisa plate. Prior to PAMP elicitation the water was carefully removed and 100 µL of assay solution (17 mM luminol [Sigma], 1 mM horseradish peroxidise [Sigma], and 100 nM flg22 [Peptron] or 100 µg / mL chitin [Nacosy]) was added to each well. Luminescence was detected for up to 80 minutes using a Photek camera system (Photek Ltd., St Leonards-on-sea, UK).

## 2.2 Plant pathogens

### 2.2.1 List of bacterial and oomycete strains

#### *2.2.1.1 Oomycetes isolates*

*Albugo laibachii* isolate Nc14 (*AlNc14*) (Originates from Norwich, Norfolk, UK.)

*Albugo laibachii* isolate Em1 (*Alem1*) (Originates from East Malling, UK)

*Albugo candida* isolate Nc2 (*AcNc2*) (Originates from Norwich, Norfolk, UK.)

*Hyaloperonospora arabidopsidis* (*Hpa*):

*Hpa isolate Noco2* (Originates from Norwich, Norfolk, UK.)

*Hpa isolate Waco9* (Originates from Wageningen, The Netherlands.)

*Phytophthora infestans* blue 13

*Phytophthora infestans* 88069

*Phytophthora infestans* NL07434

#### 2.2.1.2 Bacterial strains

**Table 2.3: Bacterial strains used in this thesis.**

Bacterial strains used in this study

| Bacterial strain | harboring plasmids |
|---|---|
| *Pseudomonas syringae* DC3000 | n/a |
| *Pseudomonas syringae* DC3000 luxCDABE | ΔSP-CHXC1:pEDV6 |
| | ΔSP-CHXC1(C651A):pEDV6 |
| | ΔSP-SSP6-A:pEDV6 |
| | AvrRps4(AAAA):pEDV6 |
| *Pseudomonas syringae* DC3000 *ΔAvrPto/ΔAvrPtoB* | n/a |
| | AvrRps4(AAAA):pEDV6 |
| | ΔSP-CHXC1:pEDV6 |
| | ΔSP-CHXC1(C651A):pEDV6 |

### 2.2.2 Bacterial strains and growth conditions

Media

All recipes are for the scale of 1 liter.

L broth 10 g tryptone, 5 g yeast extract, 5 g NaCl, 1 g glucose, pH7.0. For solid medium, 10 g agar was included.

King's B 20 g Peptone, 10 mL 100% glycerol, 1.5 g Heptahydrated Magnesium Sulfate, 1.5g Potassium Hydrogen Phosphate, adjusted to pH 7.0. To make solid medium, 10 g agar was included.

GM 4.3 g MS salts, 0.59 g MES, 0.1 g myo-inositol, 1 ml of 1000x GM vitamin stock, 8 g Bacto agar, pH adjusted to pH 5.7 with KOH. 100 ml of 1000x

GM vitamins contains 0.1 g thiamine, 0.05 g pyridoxine, 0.05 g nicotinic acid.

## 2.3 Antibiotics

The final concentrations of antibiotics used for selection of bacteria were of 100 µg/ml, 25 µg/ml, 50 µg/ml, 50 µg/ml, and 100 µg/ml for carbenicillin, gentamycin, kanamycin, rifampicin and spectinomycin, respectively. All antibiotic solutions were filter-sterilized using a 22 µm micro-filter.

Transgenic *Arabidopsis* lines were selected with either kanamycin, gentamycin or spectinomycin at 50 µg/ml. Screening for resistance to BASTA was performed by applying a 100 mM aqueous solution.

## 2.4 Microbial methods

### *2.4.1 Hpa* Waco9 infection

3-week-old plants were spray-infected (taking care that each plant received a similar inoculum) using 10 ml per 24 plants of an aqueous suspension of 5 x $10^4$ spores/ml. Similarly, 10 day old cotyledons were infected with 10 ml aqueous solution per 8 FP7 pots. To achieve 100% humidity the sprayed plants were covered with a transparent plastic lid and grown at 16°C with a 10 hour light and 14 hour dark cycle. For propagation and sub-culturing fresh spores were collected 1 week after infection.

### *2.4.2 A. laibachii* and *A. candida* infection

Heavily infected Col-Tho (Col-5 transformant containing multiple insertions of RPW8, highly resistant to Powdery mildews) leaves (leaves from ~4-6 plants) were collected and suspended by rigorous vortexing in ~30 ml ice cold sterile water and incubated on ice for 40 minutes. Subsequently the suspension was filtered through mira-cloth. The cleaned

suspension was spray-inoculated onto 4-6 week old *Arabidopsis* plants. The plants were kept in the dark at 4°C and 100% humidity in a plastic bag overnight. The following morning the plants were transferred to a Sanyo cabinet and kept under 10-hr light and 14-hr dark cycles with a 20°C day and 16°C night temperature. The plants were kept in the plastic bag to maintan 100% humidity for 1 day in the Sanyo cabinet, after which the bag was removed and the plants kept at ambient humidity. Infections were scored at 7, 14 and 21 dpi.

For propagation and sub-culturing fresh spores were collected 1 week after infection.

### *2.4.3* Pseudomonas syringae infection

Frozen glycerol stocks of *Pseudomonas syringae* strains were streaked onto King's B media supplemented with appropriate antibiotics. The plates were incubated at 28°C for 2-3 days, where after single colonies were resuspended in 200 µl L media, which was spread onto a new fresh KB plate supplemented with appropriate antibiotics and incubated overnight at 28°C.

Bacteria were recovered and resuspended in 10 mM $MgCl_2$ and diluted to a final $OD_{600}$ of 0.2 corresponding to $2x10^8$ CFU/ml. Six 4-5 week old *Arabidopsis* plants were spray infected with 10 ml bacterial solution supplemented with 0.02% v/v Silwet-77 within 1 hour of preparation of the bacteria. The plants were covered with a transparent lid post infection.

Bacterial populations within the leaves were determined 3 or 4 days post infection (dpi).

### *2.4.4* Quantification of bacterial populations within leaves

Infected leaves were sampled with a cork borer (no. 3 taking) three punches per leaf corresponding to 1 $cm^2$. The leaves were added to a 1.5 ml Eppendorf tube supplemented with 200 µl 10 mM $MgCl_2$, the leaves were ground and a further 800 µl 10 mM $MgCl_2$ was

added. The suspension was vortexed and a 150 µl aliquot was diluted in 150 µl 10 mM $MgCl_2$. This was considered as a $10^{-1}$ dilution. This was then diluted 1:10 in 300 µl 10 mM $MgCl_2$. This procedure was continued until a $10^{-6}$ dilution was reached.

Each of the serial dilutions were plated on L-plates containing appropriate antibiotics and incubated until colonies appeared which then were counted.

### *2.4.5* Trypan blue staining of *Hpa* structures

In a fume hood, the *Hpa* infected leaves were transferred to 15 ml falcon tubes and 4 ml (or enough to cover the leaves) of 1% Trypan blue w/v in Lactophenol (NBS Biological Limited) diluted 1:1 with 100% EtOH. The leaves were boiled for approximately 2 minutes in a water bath. The trypan blue solution was discarded, and leaves were destained for 1 hr in 5 ml chloralhydrate (2,5 g/ml). To reduce unspecific staining the chloralhydrate solution was replaced 3 times during a 36 hr period where each wash took 8-16 hrs. Finally, the leaves were mounted in v/v 60% glycerol on glass slides.

### *2.4.6* Quantification of *Hpa* conidiophores on adult trypan blue stained leaves

Trypan blue stained conidiophores were counted using differential interference contrast and a 10 – 40x magnification (depending on level of infection) on a Leica DMR microscope. The size of the leaf was determined by taking a picture. The leaf area was cropped using Photoshop and the area measured using ImageJ. Finally the number of condiophores was normalized to the area of the leaf ($cm^2$).

### *2.4.7* Fluorescent detection of aerial *Hpa*, powdery mildew *and* P. *infestans*

#### 2.4.7.1 *Staining of Hpa, powdery mildew* and *P. infestans structures*

Three fully expanded adult leaves from the same rosette were sampled from each plant. The leaves were removed with a pair of forceps around the middle of the petiole and submerged dorsal side down in 0.01% w/v uvitex 2B dissolved in distilled water. The

leaves were incubated for at least 1 minute. Subsequently, the leaves were washed for approximately 5 seconds by submersion 3 times in water. Note, washing removes background, but also decreases the intensity of signal. The leaves were dried on paper towel to remove excess water and then mounted on glass slides dorsal side up.

A similar approach was applied to stain *P. infestans* infected *N. benthamiana* leaves and powdery mildew infected *Arabidopsis* leaves.

#### 2.4.7.2 *Imaging Hpa, Powdery mildew* and *P. infestans structures with Uvitex 2B*

Uvitex 2B stained leaves were imaged using a Leica M165 FC Fluorescent Stereo Microscope fitted with an EL6000 laser. Pictures were recorded in 16 bit grey scale (1728x1296 pixels) for 329.1 ms with 100% open iris using 4-6x gain at 7.3x magnification. The pictures were stored as JPEG image files.

The leaf shape and size was determined by exciting chlorophyll at 470nm ± 40nm and recording the autofluorescence using a GFP2 filter (Leica order #10447407). A second picture was recorded of the uvitex 2B stained aerial *Hpa* structures using the UV filter (Leica order# 10447415) (Figure 7.2). Bright field pictures were taken using transillumination. A similar approach was applied to stain *P. infestans* infected *N. benthamiana* leaves and powdery mildew infected *Arabidopsis* leaves.

#### 2.4.7.3 *Quantification of Hpa conidiophores on cotyledons using uvitex 2B*

Five days post infection the number of conidiophores on a cotyledon was determined. Staining of the conidiophores was performed by submerging the cotyledon in an aqueous 0.01% w/v uvitex 2B solution for 5 sec. Excess stain was removed by submersion of the cotyledons in tap water for 5 sec. The washing step was repeated twice. Afterwards the cotyledons were dried on paper towel for 30 sec to remove excess liquid. The cotyledons

were mounted on glass slides bottom side up. The stained conidiophores were counted using a Leica M165 FC Fluorescent Stereo Microscope fitted with an EL6000 laser using the UV filter (BP 350/50, LP 420).

### 2.4.8 Infection of *N. benthamiana* leaves with *Phytophthora* sp.

#### *2.4.8.1 Rye A plates*

Take 60g rye grains and soak for 36 hr in H2O. Remove liquid and add 20 g sucrose. Macerate for 10 s, and incubate for 3 hr at 49°C. Filter the macerate through a muslin and adjust to pH 7. Add 15g agar and H2O to 1L. Autoclave and pour plates.

#### *2.4.8.2 Phytophthora growth and infection*

From a heavily infected plate take a 1 $cm^2$ square Rye A agar and transfer to a fresh Rye A plate. Incubate for 10-14 days at 18°C depending on stain. To produce infectious inoculum, wet plate with 4 ml $H_2O$ and incubate for 2 hr at 4°C. Scrape off P. *infestans* and adjust to 50000 zoospores/ml with $H_2O$. Infect by drop inoculation with approximately 500 zoospores/drop on detached *N. benthamiana* leaves. Incubate leaves in Petri dishes (20 cm x 20 cm) fitted with wet blue roll for 3 to 6 days depending in strain.

## 2.5 Molecular biology methods

### 2.5.1 DNA

#### *2.5.1.1 Rapid DNA isolation for amplicons less than 1000 bp*

DNA was isolated using an aqueous 10% w/v chelex 100 solution (Hwangbo et al. 2010). Approximately 0.1 $cm^2$ leaf material was ground using a pipette tip in 50 µl chelex

solution and this was boiled for 5 min. The chelex resin and leaf material was pelleted and 2 µl supernatant was used for PCR reactions.

#### 2.5.1.2 High molecular weight DNA isolation from Albugo sp. and Arabidopsis

DNA was extracted using a slightly modified version of the proteinase K/SDS method from McKinney et al. (1995). Briefly, tissue was ground in liquid nitrogen and resuspended by inversion in a extraction buffer (50mM Tris pH 8.0, 200 mM NaCl, 0.2 mM EDTA, 0.5% w/v SDS, 100 mg/ml Proteinase K) in a 1:2 ratio. Subsequently an equal volume of phenol:chloroform was added and mixed by inversion. Following a centrifugation step the top aqueous phase was transferred to a new tube and chloroform:isoamyl alcohol (24:1) was added in a 1:1 ratio and mixed by inversion. Following a centrifugation step the top aqueous phase was transferred to a new tube and the DNA precipitated using sodium acetate and isopropanol. The DNA was washed twice in 70% v/v EtOH, pellet dried, and resuspended in TE.

#### 2.5.1.3 Ethanol precipitation of DNA

Two volumes of ice cold EtOH and 0.1 volume of 3M NaOAc were added to 1 volume of DNA and mixed by inversion. DNA was pelleted by centrifugation 15000g at 4°C, the supernatant decanted, and the pellet washed twice with 70% v/v EtOH. Pellet was air dried and resuspended in TE.

#### 2.5.1.4 Plasmid isolation

Plasmid minipreps were prepared using NucleoSpin® Plasmid (Macherey-Nagel) following manufactures instructions and eluting in a 60 µl volume.

#### 2.5.1.5 DNA Digestion

~500 ng DNA was typically digested in a 20 µl volume following manufacturers instruction. Prior to digestion PCR products were cleaned though sepharose.

#### *2.5.1.6 Polymerase chain reaction (PCR)*

PCRs were carried out using 20 to 100 ng DNA preparations as templates. Each 20 µL reaction contained: 1 X Phusion HF buffer, 0.2 mM dNTPs, 0.002 U Phusion polymerase (Finnzymes), and 0.5 µM of each primer. PCR cycles were optimised for different primers and the length of amplicons. Normally an annealing temperature of 58°C was used and extension of 30 sec per kb. PCR was performed in a DNA thermal cycler (Peltier Thermal Cycler 225, MJ Research). Phusion PCRs were used for cloning and allele mining purposes.

PCRs for mapping, colony PCR and mutant genotyping were performed with home made Taq polymerase. Colony PCR was used to identify positive colonies containing recombinant plasmids during cloning. The PCR was performed as above except that the DNA template was substituted with bacterial cells diluted in 50 µl MilliQ water.

## Table 2.4 Primers used in this study.

| Identifier | Sequence | Notes |
|---|---|---|
| **Used to resolve the SSP6 locus in AlNc14** | | |
| TSL316 | taggccctctgttggcatac | |
| TSL317 | taacaaccattgtcccacga | |
| TSL318 | cctgattcgattgtgacgtg | |
| TSL319 | taacaaccattgtcccacga | |
| | | |
| **Mapping (Chapter 7)** | | |
| TSL312CIW8F | TAGTGAAACCTTTCTCAGAT | |
| TSL313CIW8R | TTATGTTTTCTTCAATCAGTT | |
| TSL306T2K10-3-fw | CATTAAAATCAAGAAATACATGT | mapping AKA map4 |
| TSL307T2K10-3-r | gacttttatctcatgttctattaa | mapping AKA map4 |
| TSL308ST101 | ATGTCCAAATTGACCAACCG | mapping AKA map8 |
| TSL309ST102 | CAAAATAAACACCCCCAACT | mapping AKA map8 |
| TSL84RAC1fwd | TGTCGTGATATCATTAAAACGTA | |
| TSL85RAC1rev | GGGAGATTCTCGAGATGCAA | |
| TSL86Cra1fwd | GCATGACAATATTTTGGCGACTTGTC | |
| TSL87Cra1rev | GCACATTAGGAGCGGTGATACCATTGC | |
| TSL88NGA129fwd | CACACTGAAGATGGTCTTGAGG | |
| TSL89NGA129rev | TCAGGAGGAACTAAAGTGAGGG | |
| TSL90ciw9fwd | CAGACGTATCAAATGACAAATG | |
| TSL91ciw9rev | GACTACTGCTCAAACTATTCGG | |
| | | |
| **RACE-PCR primers** | | |
| *General primers* | | |
| TSL20TN | ttttttttttttttttttttV | |
| TSLn5'Primer | TATCAACGCAGAGTGGCCATTAC | |
| 5'PCR primer | AAGCAGTGGTATCAACGCAGAGT | |
| GeneRacer 3' | GCTGTCAACGATACGCTACGTAACG | |
| GeneRacer 3' nested | CGCTACGTAACGGCATGACAGTG | |
| TSL141RCA5p | TGAGCTAAGGACAGCGTGGT | |
| TSL142RCA5pn | GTGAGTTCCGGTTCCCATAG | |
| TSL143RBCS1A5pn | CGTGCTCACGGTACACAAAT | |
| TSL144RBSC1a5p | TCCATCATAGTATCCGGGTGA | |
| *CHXC1* | | |
| TSL72CHXC1'3 | AGATTTGCGTTACCGCTGCTTAAA | |
| TSL73CHXC1'3n | ATCCTCGGAGTTCCCATTTCATTT | |
| TSL74CHXC1'5 | CGGGCTCATTCAAAAACTCAATTC | |
| TSL75CHXC1'5n | CAGTGAAATACCCCATCGAGATCC | |
| TSL229CHXC1p5wak | ACGCTTCTTTCGGAGTACCA | |
| TSL236CHXC1cDNAf | TAGTCACGTATTAGCATACAATTTCAT | |
| TSL237CHXC1cDNAr | CAGCAGGAGAATTATGACCAA | |
| *SSP6* | | |
| TSL261SSP6seqfw | CGGCTTGCTGATTTGGTATT | |
| TSL262SSP6seqfw | TCTACGCAACAGGATCAGGA | |
| TSL263SSP6seqrev | TCAAATGGATCCTGTGTGTG | |
| TSL246SSP6gDNAf | AAAAATCCAGGAGGGTTCCGCTTT | |
| TSL247SSP6gDNAr | CCCTGCAATGAAGAATTGCGGATA | |
| TSL248SSP6cDNAfw | CAAACAAGGTGTTGTCATTACGATTTAAC | |
| TSL249SSP6cDNArev | GAAGGAGAGTAGTATTAATAGAGATC | |
| TSL178ssp6_3p | GAGCCAGCAAAAAGGGAGCATGGA | |
| TSL179ssp6_3pn | GTTTGGCGTTGCAAAATCCGCAAG | |
| TSL180ssp6_5p | CTTGCGGATTTTGCAACGCCAAAC | |
| TSL181ssp6_5pn | TTGATCGCTTGGCGCTGTAATTCG | |
| | | |
| **Cloning of CHXC1 constructs** | | |
| TSL333dSP-CHXC1f | GGCTGAAGCTGAATTCGTATGCAGAAGCAGAGTT | |
| TSL335dSP-CHXC1r | TGGGCCACGTGAATTCTTCCAGTCCAAATCCGGT | |
| TSL336CHXC1_157f | aagttctgtttcagggcccgGCTCATCACACGCAACAAAA | |
| TSL337CHXC1_285f | aagttctgtttcagggcccgCTTGTTTCACTACAGTTTCG | |
| TSL301CHXC1PopinF | aagttctgtttcagggcccgGTATGCAGAAGCAGAGTTT | |
| TSL302CHXC1PopinR | ATGGTCTAGAAAGCTTTATTCCAGTCCAAATCCGGTGA | |
| TSL325CHXC1infus | GAATTCCCGGGGATCATGGTATGCAGAAGCAG | |
| TSL326CHXC1infus | CATGGTCGACGGATCTCATTCCAGTCCAAAT | |
| TSL295CHXC1-mutF | CACGGGCACACACAGCCTTTAACCGAATCGATTTGCCCA | |
| TSL296CHXC1-mutR | TGGGCAAATCGATTCGGTTAAAGGCTGTGTGTGCCCGTG | |
| TSL45CHXC1SPnew | CACCATGGTATGCAGAAGCAGAGTTT | |
| TSL71CHXC1r | TCATTCCAGTCCAAATCCGGTGA | |
| | | |
| **Cloning of SSP6 variant constructs** | | |
| TSL226SSP6sp | CACCATGGAAGAGCCAGCAAAAAGGGAGCATG | |
| TSL227SSP6rev | CCATTTATTATTTGCGAAAGCCATCTT | |
| TSL298SSP6c2rdST | ATGTTTCGGCCATTTTGTAATGTTGTA | |
| | | |
| **Standard primers used in the study** | | |
| TSL-M13uni(-21) | tgtaaaacgacggccagtg | |
| TSL-M13rev(-29) | caggaaacagctatgaccatg | |
| TSL-edv6-FOR | tacacccaaatccctattgg | |
| TSL-Lba1 | tggttcacgtagtgggccatcg | |
| P35Sfw | TGGAGAGGACTCCGGTATTTTT | |
| T35Srev | TTTTGCGGACTCTAGCATGG | |

#### 2.5.1.7 Agarose gel electrophoresis and gel extraction

For DNA electrophoresis of sequence length polymorphism (SLP) markers with amplicons of ~200 bp ± 20bp, 1.35% w/v Agarose gels supplemented with 0.7% w/v Synergel (Diversified Biotech, ES) (equivalent of a standard 3% w/v agarose gel) were used. The gels were made following manufactures instruction and run as a normal Agarose-gel.

Routine DNA visualization was performed by electrophoresis on 1% w/v agarose gels supplemented with 0.5 µg/ml ethidium bromide, in 1 x TAE. Gels were run at 100-150V and DNA visualized under UV light. Gel DNA extractions were performed using the QIAEX II kit (QIAGEN) or NucleoSpin Gel and PCR Clean-up (Macherey-Nagel) following manufactures instructions.

#### 2.5.1.8 Cloning

RACE-PCR products and allele mining PCRs were ligated into a StuI digested dephosphorylated pCR-blunt vector (Invitrogen) using T4 ligase (Invitrogen) following the manufacturers instructions. Bacterial and yeast expression constructs were generated using *Infusion* HD cloning kit (Cat # 639648, Takara Bio Europe/Clontech) technology following manufacturers recommendations with the exception that 25% of the recommended reaction volume was used. Cloning of constructs for use *in planta* were created using the pENTR D-TOPO entry vector according to manufacturers instruction except for performing ¼ size reactions. Clones were verified by DNA sequencing.

To create *in planta* expression constructs entry vector and destination vector were recombined using LR-II clonase mix (Invitrogen, UK) following manufacturer's instructions but using 50% of the recommended reaction volume.

#### 2.5.1.9 DNA sequencing

DNA sequencing reactions were performed in a 10 µl volume containing approximately 150 ng template plasmid DNA, 0.5 µl of a 3.2 µM sequencing primer, 1.5 µl 5x buffer and 1 µl ABI Big Dye Terminator Ready Reaction Mix (Big Dye 3.1 by Perkin Elmer). The following PCR cycle conditions were used: initial step at 96°C for 1 min, denaturation at 96°C for 10 sec, annealing at 50°C for 5 sec and elongation at 60°C for 4 min (25 cycles total). Sanger sequencing was carried out at Genome-Enterprise (TGAC, Norwich). Sequencing data were trimmed, analyzed, and aligned using Sequencher software (http://genecodes.com/).

#### 2.5.1.10 Site directed mutagenesis

Primers were designed using the recommendations from the QuickChange XL Site Directed Mutagenesis Kit (Stratagene). Phusion PCR was run as previously described but with 25 cycles and a 140 seconds extension time. The PCR product was cleaned using a Macherey-Nagel Nucleo-Spin kit following manufacturers recommendations. 10µl of the PCR was digested with DpnI in NEBuffer 4 (New England Biolabs) for 2 hr at 37°C to remove template plasmids. The digestion was cleaned though sepharose resin and transformed into *E. coli* Mach-1 cells (Invitrogen).

### 2.5.2 RNA

#### 2.5.2.1 Isolation of total RNA

Plant tissue was ground in liquid nitrogen. Frozen powder was transferred to a pre-cooled 1.5 mL Eppendorf tube and 1 mL Tri-Reagent (Sigma) was added and incubated at room temperature for 10 min. The tube was centrifuged for 20 min at 12,000 × g and the supernatant was transferred to a new tube and a 1:1 volume of isopropanol was added. The tube was incubated overnight at −20°C and subsequently centrifuged for 15 min at 12,000

× g at 4°C. The supernatant was discarded and pellets were washed with RNAse free 70 % w/v ethanol, air dried, and resuspended in RNase-free water. Yield and protein contamination was determined using the optical density at 260 nm and 280 nm Micro-Volume UV-Vis Spectrophotometer for Nucleic Acid and Protein Quantitation (Nanodrop, Thermo scientific, UK) and RNA quality determined by gel electrophoresis.

#### 2.5.2.2 *3' and 5' RACE PCR*

cDNA with 3'- and 5' adaptors suitable for rapid amplification of cDNA ends (RACE) - PCR was generated using the GeneRacer Kit following manufacturer's instructions (Invitrogen, UK). The RACE-PCR was performed using Invitrogen 3' and 5' primers and nested primers (see Table 2.4) using $1/10^{th}$ of the recommended cDNA. Both primary and nested RACE-PCR was performed using phusion PCR with 25 cycles and 3 min extension time. RACE-PCR products were evaluated by agarose gel electrophoresis. DNA fragments of interest were extracted and sequenced.

### 2.5.3 Bacteria

#### *2.5.3.1 Agrobacterium tumefaciens and E. coli plasmid transformation by electroporation*

Electro competent *Agrobacterium tumefaciens* GV3101 (~$10^6$ CFU) or *E. coli* DH10B (~5 x $10^8$ CFU) cells were transformed in 0.1 cm cuvettes using a BioradMicropulser at standard *Agrobacterium* or *E. coli* setting. Cells were recovered in liquid L medium for 1-2 hr at 28°C (Agrobacterium) or 45 min at 37°C (*E. coli*) before plating on 1% Agar L plates with appropriate antibiotics.

#### *2.5.3.2 Transformation of chemically competent E. coli cells*

*E. coli* Mach-1 cells (derived from an *E. coli* W strain and not standard K12 strain) were transformed following manufacturers instructions (Invitrogen UK).

#### 2.5.3.3 Triparental mating

Overnight cultures of acceptor (i.e. *Pseudomonas syringae* pv. *tomato* DC3000 ΔAvrPto/ΔAvrPtoB or *Pseudomonas syringae* pv. *tomato* DC3000 luxCDABE), donor (i.e. *E. coli* harboring effector candidates in pEDV6) and helper (*E. coli* harboring pRK2013) were grown on L media with appropriate antibiotics at 37°C with agitation. The following day fresh cultures were set up in the morning and grown for 10 hr. The cells were harvested, washed once, resuspended in 10 mM $MgCl_2$ and mixed in a ratio of 1 donor : 8 acceptor : 1 helper. The bacteria were mixed and spotted on an L plate without antibiotics and grown overnight at 28°C. The following day bacteria were re-streaked onto KB plates with selective media and incubated for 2 days at 28°C. Colonies formed were re-streaked onto two plates, which were incubated at either 28°C or 37°C. Colonies that only grew at 28°C were considered to be *Psuedomonas synrigae* transconjugants. Presence of the plasmid was confirmed by PCR.

### 2.5.4 Protein

#### 2.5.4.1 Total protein extraction

Three leaf punches of cork borer size no. 3 were ground in liquid nitrogen with a pestle in a 1.5 ml eppendorf tube. Total protein was extracted by adding 200µl 2x Laemmli buffer (62.5mM TrisHCl, 4.37% v/v glycerol, 1% w/v SDS, 125 mM DTT). The solution was mixed by vortexing. Proteins were denatured prior to SDS-PAGE by boiling for 5 minutes at 95°C.

#### 2.5.4.2 SDS page

Proteins were separated by electrophoresis on acrylamide gels (8% or 15% w/v acrylamide, 0.375 M Tris pH 8.8, 0.1% w/v SDS, 0.1% w/v APS, 0.04% v/v TEMED for separation gels and 5.1% w/v acrylamide, 125mM Tris pH 6.8, 0.1% w/v SDS, 0.1% w/v

APS, 0.1% v/v TEMED for stacking gels) using 1 mm spacers in the Biorad mini protein gel casting system. Loading buffer (final concentration 50 mM Tris-HCl pH 6.8, 100 mM DTT, 2% w/v SDS, 0.1% w/v bromphenol blue, 10% v/v glycerol) was added to the protein samples and the gels were run in a standard SDS page running buffer (24 mM Tris, 193 mM Glycine, 0.1% w/v SDS). Gels were run 125 V.

#### *2.5.4.3 Coomassie staining of protein gels*

SDS-PAGE gels were washed briefly in reverse-osmosis water and then stained by incubating gels for up to 1 hr in 15 ml Instant blue (Expedeon, UK). Excess stain was removed by washing in reverse-osmosis water.

#### *2.5.4.4 Western blot*

Proteins separated by SDS-PAGE were transferred to a PVDF membrane (Biorad); the membranes were blocked with TBS supplemented with 5% w/v milk powder. The membranes were probed with primary antibody (for concentrations, see Table 2.5) overnight at 4°C. The membrane was washed three times in TBS + 0.1% v/v Tween20 and thrice with TBS prior to probing with HRP coupled secondary antibody for 1 hr. Subsequently, the membrane was washed three times with TBS + 0.1% v/v Tween20 and six times with TBS. All antibodies were diluted in TBS supplemented with 5% w/v milk powder and washing steps were more than 8 min. Western-blot membranes were initially developed with ECL plus (GE Healthcare) as substrate for HRP reactions. If no reaction was observed, Femto solution (Thermo Scientific) was used as substrate.

**Table 2.5: List of antibodies used in this study.** * The custom generated antibody was raised in two guinea pigs injected with a combination of two 16 aa long peptides (H2N-KKFSQRQRGAQRRKLC-CONH2 and H2N-ERGMDGRLQWKQTDTC-CONH2) together with an adjuvant. The animals were sacrificed 28 days post immunization and the purified antibody was tested against both peptides by Ariane Kemen. Plant species in parenthesis describe where the given antibody works best.

| Purpose | Raised in | Specificity | Company | Working dil. | Ref Supp |
|---|---|---|---|---|---|
| Primary | rabbit | Anti-GFP (for A. thaliana) | Amsbio | 1/5.000 | n.d. |
| | mouse | Anti-GFP (for N. benthamiana) | Roche | 1/3.000 | 11814460001 |
| | Guinea Pig | Anti-CHXC1 (custom made antibody *) | Eurogentec | 1/1000 | n/a |
| | rat | Anti-HA 3F10 | Roche | 1/2000-1/3000 | 11867423001 (50ug) |
| Peroxidase-Conjugated | | Anti-HA 3F10 Peroxidase High Affinity | Roche | 1/3.000 | 2013819 |
| | | Anti-GFP (for N. benthamiana) | Santacruz | 1/10.000 | sc-9996 |
| Secondary Antibodies | | Anti-Rat IgG HRP raised in Rabbit | Sigma | 1/12.000 | A 5795 |
| | | Anti-Rabbit IgG HRP raised in goat | Sigma | 1/10.000 | A0545 |
| | | Anti-Guinea pig | | 1/1000 | |
| | | Anti-Mouse IgG HRP raised in goat | Sigma | 1/10.000 | A 2304 |

#### *2.5.4.5 CHXC1 protein expression and refolding*

**CHXC1 protein expression**

*E. coli* soluBL21 cells harboring ΔSP-CHXC1::pOPIN-F were grown overnight in 10 ml L broth with appropriate antibiotics at 37°C with agitation. Subsequently, the culture was diluted 1:100 to a final volume of 101 ml and grown at 37°C with agitation until an $OD_{600}$ of 0.8 was reached. Protein expression was induced by adding 100 µl 1 M IPTG to the culture and the incubating for 5 hr at 28°C with agitation. The cells were collected in 50 ml Falcon tubes by centrifugation at 5000 g at 4 °C for 5 min. The supernatant was discarded and the bacterial pellet stored at - 80°C or directly used for protein purification.

**CHXC1 protein purification**

Bacterial pellets were resuspended in ice cold 40 ml PBS supplemented with a P1000 tip of lysozyme and lysed by sonication (1 sec pulse, 3 sec rest, 40% amplitude and 2 min sonication). The lysate was spun at 4°C for 15 min at 9000 rpm and the supernatant discarded. The pellet was resuspended in 40 ml ice cold PBS + 1% v/v triton X100. Inclusion bodies were collected by centrifugation at 4°C for 15 min at 9000 rpm and the supernatant discarded. The pellet was washed three times in 40 ml ice cold PBS + 1% v/v triton.

After three washes the pellet containing 6xHIS-ΔSP-CHXC1 was resolubilized in 20 ml ice cold PBS + 1.5% w/v sarkosyl by incubation for 1 hr at 4°C with constant mixing on a shaking table. Non-solubilized proteins were pelleted by 15 min centrifugation at 9000 rpm at 4°C, and the supernatant was flash diluted 1:20 into ice cold PBS with rigorous stirring. The proteins were allowed to refold overnight at 4°C. The solution was then filtered through a 22 µm filter to remove protein aggregates. The resulting liquid was run over a 5 ml HIS-trap column (GE-healthcare). The column was washed with 50 ml PBS + 25 mM imidazole, and then proteins were eluted in 20 ml PBS + 250 mM imidazole, which was passed twice over the column. The proteins were concentrated to a final volume of 1 ml using an amicon column (Millipore Ltd) with a molecular weight cut off of 5000 Da following manufacturers instructions.

## 2.6 Confocal microscopy

Leaf samples were mounted in water on glass slides and optical sections visualized by confocal microscopy (Leica DM6000B/TCS SP5, Leica Microsystems). Constructs transiently expressed in *N. benthamiana* were analyzed between 1 and 3 days after infiltration. Where indicated plasmolysis in *N. benthamiana* was performed by vacuum infiltrating leaf discs with 1 M NaCl. Perfluorodecalin (Sigma Ltd.) was used instead of water to mount leaves infected with *AlNc14* to enhance the Z-depth, which is required to visualize haustoria and hyphae in infected *Arabidopsis* tissues.

GFP-tagged constructs were excited at 488 nm and emission recorded at 535 ±35 nm. Chlorophyll autofluorescence was recorded at 700±50 nm. Identical settings were used to visualize GFP-tagged constructs in stable *Arabidopsis* lines.

## 2.7 Computational prediction of secreted HECT E3 ligases in proteomes

The proteomes of *Albugo candida Nc2, Albugo laibachii Nc14, Arabidopsis lyrata, Arabidopsis thaliana, Chlamydomonas reinhardtii, Dictyostelium discoideum, Hyaloperonospora arabidopsidis Emoy2, Magnaporthe oryzae, Phytophthora capsici, Physcomitrella patens, Phytophthora infestans, Phytophthora ramorum, Phytophthora sojae, Plasmodium falciparum, Puccinia graminis, Pythium ultimum, Saccharomyces cerevisiae, Saprolegnia parasitica, Schizosaccharomyces pombe, Thalassiosira pseudonana, Toxoplasma gondii and Ustilago maydis* were screened using a HECT E3 ligase pFAM profile (PF00632) and a hidden Markov based search strategy with HMMER3 (Punta et al. 2011; Eddy 2011). Significant hits were tested for presence of a signal peptide using SignalP3.0 with default parameters (Bendtsen et al. 2004).

# 3 Sequencing of *A. laibachii* Nc14 and Em1 reveal candidate effectors

## 3.1 Introduction to the project

*A. laibachii* is remarkably well-adapted on *Arabidopsis* and is thus a good model to dissect *Arabidopsis* defense pathways (Holub 2008). Oomycetes secrete various types of effectors into the host cell to suppress defense (Kamoun 2006; Kale & Tyler 2011). Effectors often have a modular structure with a N-terminal translocation domain and a variable C-terminal domain (Schornack et al. 2009; Raffaele & Kamoun 2012).

The N-terminal translocation domain has a signal peptide targeting the protein for secretion, which typically is followed by a linear amino acid motif that is required for translocation (Schornack et al. 2009; Kamoun 2006). The C-terminal domain harbors the virulence function and is often highly variable within and between species (Win et al. 2007; Jiang et al. 2008). Effectors often occur in large expanded families and their genes show signatures of accelerated evolution evident as elevated levels non-synonymous SNPs and rapid birth and death evolution (Kamoun 2006; Raffaele & Kamoun 2012).

The advent of next generation DNA sequencing technologies has resulted in eukaryotic plant pathogenic species becoming economically feasible to sequence.

Here, we report the genome sequence of *AlNc14* and analysis of the effector complement. We functionally validate a novel class of effectors, the "CHXCs", by showing delivery to the host cell and effector functionality. This chapter will describe the results that are central to this study. Readers are referred to the article "Gene Gain and Loss during Evolution of Obligate Parasitism in the White Rust Pathogen of *Arabidopsis thaliana*" for a thorough discussion of all the results obtained.

## 3.2 Results and Discussion

### 3.2.1 RACE-PCR data aid *de novo* gene prediction and genome assembly

The *AlNc14* genome was sequenced using next-generation sequencing (NGS) techniques. Two libraries with an insert size of either 400 bp or 800 bp were constructed from *AlNc14* genomic DNA. These were sequenced using an Illumina GAII to generate 76 bp paired-end reads. The reads were *de novo* assembled using Velvet (Zerbino & Birney 2008) to generate short read assemblies (of each sequenced library or in combination).

*De novo* Velvet assemblies generated from NGS data correctly assemble non-repetitive regions, which normally encompass the gene rich regions. In contrast, repetitive regions such as simple repeats and recently diverged paralogs do not. They become continuity break points, as Velvet fails to find a single continuous path through the de Bruijn graph. This results in paralogs being split into a number of discrete contigs dependent on the number of diverged base-pair positions within a homologous duplicated region and other parameters set in Velvet. Thus, most nucleotides are represented in the genome assembly, but the final assembly is not necessarily contiguous due to complex branching in the underlying de Bruijn graph structure.

For this reason we generated a more contiguous genome assembly, but theoretically less correct, with the Minimus meta assembler (Sommer et al. 2007) using the Velvet *de novo* genome assemblies as input. This method takes advantage of the variation in continuity in assembly using Velvet with different parameters and different insert length libraries. The quality of the Minimus meta assembly was assessed by realigning both pair-end reads and single-end reads back to the genome. Merged regions with no read coverage were subsequently split into two. The advantage of this type of two-phase assembly is that, while not all regions are assembled, more regions are completely assembled into more contiguous sequence.

Following the giant panda genome, this is to our knowledge the second eukaryotic genome to be entirely assembled by 76 bp paired-end reads at the time of publication (R. Li et al. 2010).

As the genome relies entirely on *de novo* gene prediction, we independently verified candidate effector gene models by RACE-PCR experiments. In initial genome assemblies, the effector candidates were largely located at contig ends, which resulted in incorrect prediction of the 3' region of the gene. The RACE-PCR results revealed that most of these genes extended beyond the assembled contig. Consequently, we used primers designed to the 3'- and 5'-end of the cDNA determined by RACE-PCR for PCR on gDNA to obtain data on intron/exon structure and merge contigs in the genome assembly.

The RACE-PCR data was used as an aid in assembly of problematic areas in the final version of the *AlNc14* genome. In addition, the *de novo* gene calling for the genome was calibrated using the RACE-PCR results. These efforts resulted in the assembly of a ~37 Mb genome containing 13032 gene models, of which 672 encode for proteins that are predicted to be secreted without a trans-membrane domain (Eric Kemen personal comm.). Within the secretome there are both apoplastic and putative host cell translocated effectors (Kemen et al. 2011).

### 3.2.2 The *A. laibachii* effectorome

Interestingly, RXLR effectors are less likely to be important for *A. laibachii* virulence, as they are not over represented in the secretome compared to the overall proteome (Kemen et al. 2011). Similar conclusions can be drawn about the CRNs. Homology BLAST searches revealed three CRNs that have a signal peptide. An additional eight CRN proteins are identified that lack signal peptide.

Remarkably, a novel type of effector is found in the secretome of *A. laibachii* -the CHXC family. Analyses suggest that CHXC are over represented in the secretome compared to

the proteome and they have a modular protein structure with a conserved N-terminus and divergent C-terminus. We aligned all CHXC protein sequences with a signal peptide using ClustalW. The phylogenetic relationships were inferred by a Maximum Parsimony analysis using Phylip v3.69. The phylogenetic reconstruction indicated that most CHXC proteins exist in distinct clusters, but some also appear to be singletons (Figure 3.1). This is consistent with reports from other oomycetes, such as *Hpa* and *P. infestans*, where most effectors are highly redundant, and thus cluster, but others occur as singletons (Haas et al. 2009; Baxter et al. 2010).

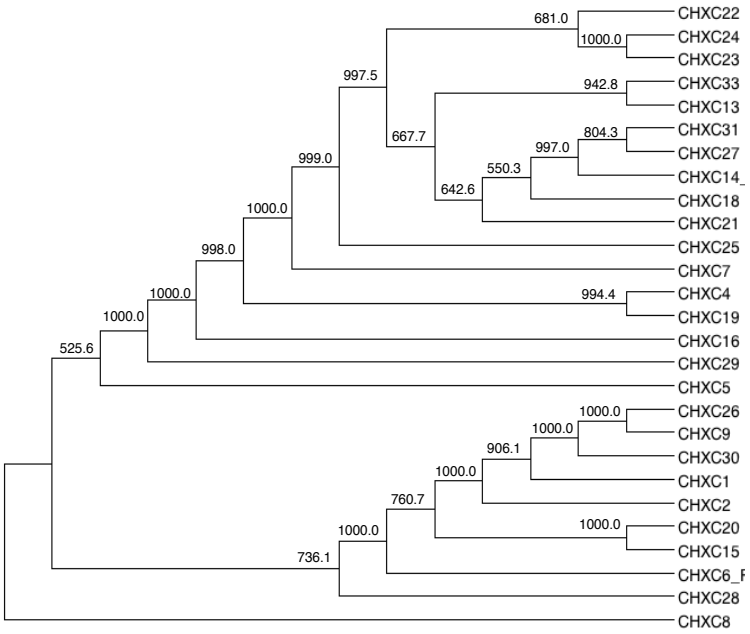



**Figure 3.1: Phylogenetic relationships of *A. laibachii* CHXCs.** A phylogenetic consensus tree representing all *AlNc14* CHXCs. The alignment was created using ClustalW with default parameters. The cladogram was computed using Maximum parsimony with a 1000 bootstraps. Numbers above branches indicate of the number of times that the branch separated the proteins into two sets, out of a 1000 trees.

Since the above-mentioned effector prediction approaches are biased towards candidate effector protein families containing easily identifiable linear motifs, a scoring index was applied to prioritize all secreted proteins prior to selecting effector candidates for experimentation. The ranking system is as follows. Given that most characterized

oomycete effectors are small, secreted proteins, this was weighted heavily, as well as the presence of known effector uptake motifs. Less importance was assigned to being on a short contig, which could be evidence of paralogs, or a location in repetitive regions of the genome. Some effector genes are under strong diversifying selection pressure; therefore if genes had non-synonymous SNPs (single nucleotide polymorphisms) or INDELs (insertions and deletions) in proteins either between or within, *AlNc14* or *Alem1* (Holub et al. 1995) this contributed to the score. To facilitate cloning and to limit artefacts, full cDNA coverage as well as high expression was also a criterion. We hypothesized that effector proteins would be expressed at early time points before sporulation, thus expression before day 4 also was assigned points. 719 genes encoding SSPs (for short secreted protein) were ranked.

### 3.2.3 Candidate *A. laibachii* effectors contribute to virulence

Effector screens in *Hpa* have revealed the majority of oomycete effectors can suppress host defenses in at least some *Arabidopsis* accessions (Fabro et al. 2011). We therefore tested if the predicted *A. laibachii* effector classes could suppress host defenses. To this end we cloned and tested effectors for potential virulence functions using the effector-detector-vector (EDV) system (K. Sohn et al. 2007a). The EDV system uses the bacterial T3SS to deliver effectors into the plant cytoplasm via fusions of candidate effector to the N-terminus of the *P. syringae* effector protein, AvrRps4. Upon translocation into the host cell AvrRps4 is processed *in planta* to cleave off the N-terminal 133 amino acids of the protein. Substitution of the KRVY motif (positions 134-7) to quadruple-alanine results in AvrRps4 becoming non-functional (K. H. Sohn et al. 2009).

We used *Pst* DC3000 luxCDABE delivering AvrRps4(AAAA) as a baseline for virulence; we found that CHXC5, CHXC7, RXLR1, RXLR2 and SSP6-A but not CRN2 conferred enhanced virulence of *Pst* DC3000 luxCDABE on *Arabidopsis* Col-0 compared to

AvrRps4(AAAA). On Nd-0 these effectors also enhanced the virulence of *Pst* DC3000luxCDABE compared to AvrRps4(AAAA) (Figure 3.2 and Kemen et al. 2011). In addition, CHXC1, CHXC5 and RXLR1 enhanced virulence of *Pst* DC3000 ΔAvrPto/ΔAvrPtoB compared to AvrRps4(AAAA) on Col-0 (Figure 3.2).

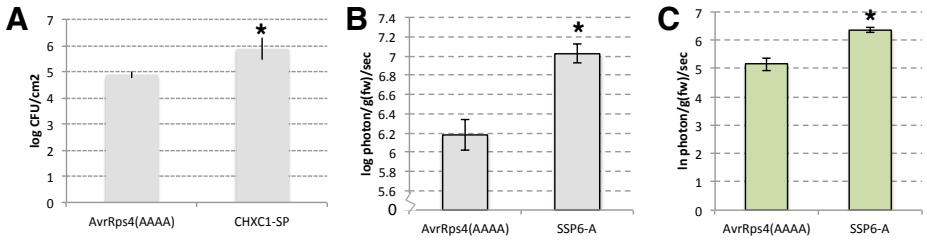



**Figure 3.2 CHXC1 and SSP6-A confers enhanced virulence.** A) 4-5 week old Col-0 plants were infected with 5x10$^8$ CFU of *Pst* DC3000 ΔAvrPto/ ΔAvrPtoB harboring AvrRps4(AAAA) or ΔSP-CHXC1 in pEDV6 were spray infected onto and growth assessed 3 dpi. The experiment was repeated thrice with similar results B+C) 4-5 week old Col-0 B) or Nd-0 C) plants were infected with 5x10$^8$ CFU of *Pst* DC3000 luxCDABE harboring either AvrRps4(AAAA) or ΔSP-SSP6-A. Bacterial growth was measured 2 dpi as an increase in luciferase photon esmission per gram fresh weight per second (photon/g[fw]/ sec. The histograms represent the log median of photon emission of three independent experiments. *: p<0.05, student two-tailed t-test.

## 3.3 Conclusion

The sequencing of the *AlNc14* genome revealed novel classes of effectors that suppress host defenses. Thus our data suggests that the effector classes employed in the suppression of host defenses by *A. laibachii* differ to those from *Phytophthora* sp. and *Hpa.*

We chose to study the candidate effectors CHXC1 and SSP6 in further detail for the following reasons: CHXC1 has homology to HECT (homologous to the E6-AP carboxyl terminus) E3 ligases, which are enzymes that function as ubiquitin ligases; many plant pathogenic oomycetes encode a secreted CHXC1 homolog. In contrast, SSP6 did not have homology to known proteins, but it was encoded at a heterozygous locus suggesting the gene could be fast evolving. Thus, CHXC1 and SSP6 represent two different types of

effectors, both implicated in suppression of defense. They are the focus of the studies in chapters 4 and 5 respectively.

# 4. CHXC1 has homology to a secreted HECT E3 ligase, and confers enhanced virulence in *Arabidopsis*

## 4.1 Introduction

Plant and animal pathogens produce effectors to interfere with the host immune system and to create a suitable environment for colonization. Plant pathogens secrete effectors into the apoplast and some effectors are translocated via the haustorium inside the host cell (Dodds & Rathjen 2010). The functions of most effectors that enter the host cell are unknown.

Oomycete effectors that are translocated into the host cell such as CHXCs, CRNs, and RXLRs often occur in large expanded families. These are characterized by having a conserved N-terminus that is required for delivery and a divergent C-terminus harboring the effector function that undergoes a strong diversifying selection (Win et al. 2007). Most effectors do not have similarity to known proteins. For this reason, their biochemical functions are largely unknown (Schornack et al. 2009). Knowledge about effectors and their host targets are required to elucidate these underlying mechanisms. Focusing on the events occurring after delivery of the effector inside the plant cell, by expressing the effector *in planta*, has been a highly successful strategy to characterize effector functions (Bozkurt et al. 2012).

*The P. syringae* effector AvrPtoB is an E3 ubiquitin ligase that hijacks the host ubiquitination machinery to target PRRs for degradation (L. Shan et al. 2008; Göhre et al. 2008; Janjusevic 2006). Similarly, VirF encodes an F-box protein that is required for *Agrobacterium tumefaciens* virulence; it recruits the plant E3 adaptor protein Skp1 to the T-complex protein VirE2 and VIP1 resulting in their degradation (Tzfira et al. 2004). Recently, a novel E3 ubiquitin ligase NopM (nodulation outer protein M),

which is secreted via the T3SS from the symbiotic bacterium Rhizobium sp. Strain NGR234, was shown to suppress PAMP induced ROS accumulation and be required for nodulation on *Lablab purpureus* (Xin et al. 2012). Thus, for symbiotic and many pathogenic bacteria living on plants the enzymatic activities of their effectors and how these alter plant immunity has been elucidated. By comparison, the enzymatic functions of filamentous eukaryotic pathogen effectors have been hard to predict due to their lack of sequence similarity to known proteins.

Exceptions to this rule are the fungal effector AvrPita from *Magnaporathe oryzae* that shows similarity to zinc-dependent metalloproteases; mutations in key catalytic residues resulted in abolished virulence (Orbach et al. 2000). Likewise, Avr3b from *P. sojae* carries a C-terminal Nudix hydroxylase domain that was shown to be functional but not required for Rps3b (resistance to *Phytophthora sojae* 3b) dependent resistance (S. Dong et al. 2011). Recently, CRN8 from *P. infestans* was shown to encode a kinase whose activity and nuclear localization is required for virulence (van Damme et al. 2012).

In Chapter 3 we reported that CHXC1 confers enhanced virulence of *Pst* DC3000 Δ*AvrPto*/Δ*AvrPtoB* when delivered inside the host cell via EDV (K. H. Sohn et al. 2007b). CHXC1 has similarity to HECT E3 ligases and is a candidate effector.

There are several major types of E3 ligases for ubiquitin, including the RING (really interesting new gene) and the HECT classes, that all require activated ubiquitin to function (Kerscher et al. 2006). Ubiquitin is activated by E1 activating enzymes that hydrolyzes ATP to create an ubiquitin-E1-cysteine-thioester high entropy bond. E2 conjugating enzymes have a higher affinity for ubiquitin-E1 complexes than E1s alone (Kerscher et al. 2006). Thus, E1-ubiquitin docks with E2 conjugating enzymes leading

to transfer of the ubiquitin onto the E2 conjugating enzyme and the expulsion of the E1. Similarly, E3 conjugating enzymes have a higher affinity for the activated E2-ubiquitin complex than E2 alone. These affinities create an energetically favorable cascade, which begins with a highly labile E1-ATP anhydrous bond and ends with stable covalent attachment of a 76 amino acid ubiquitin moiety to (usually) a lysine on the target substrate.

The Ubiquitin E3 ligase class RING has a finger motif that is defined by $Zn^{2+}$-chelating amino acids, which form two cross-brace-arranged free loops. The RING E3s are not enzymes in the strictest interpretation, as they function like scaffolds that bring the target protein and an activated E2 conjugating enzymes in close proximity. This proximity allows a nucleophilic lysine on the target substrate to attack the highly charged ubiquitin-thioester-E2 bond, which results in transfer of the ubiquitin moiety from the E2 onto the target substrate (Deshaies & Joazeiro 2009).

In contrast, HECT domain containing E3 ligases function as true enzymes. Crystal structures of HECT E3 ligases suggests they have a bilobal structure, where the N-terminal lobe is required for E2 docking and the C-terminal lobe determines ubiquitin chain specificity (H. C. Kim & Huibregtse 2009). HECT E3 ligases have a distinct ubiquitination mechanism: the ubiquitin is initially transferred from the E2 onto an active site cysteine in the C-terminus of the HECT domain. This is presumed to require a dramatic conformational change in lobe structure, as mutations in the hinge-region between the two lobes significantly reduce catalytic activity (H. C. Kim et al. 2011; Verdecia et al. 2003). Subsequently, the ubiquitin is transferred from the HECT domain onto a lysine on the target substrate by a nucleophilic attack on the ubiquitin-cysteine-thiol-ester bond. Single ubiquitin moieties are added in a stepwise

manner to generate polyubiquitin chains (H. C. Kim et al. 2011). Polyubiquitinated proteins are recognized and degraded by the proteasome (Vierstra 2009).

While no plant-pathogen effectors have been described with HECT E3 ligase activity, the virulence proteins VirF (*Agrobacterium* tumefaciens) and P0 (*Polerovirus*) have been demonstrated to function as F-box proteins likewise AvrPtoB from *P. syringae* functions as an E3 ligase, sharing homology with plant RING E3 ligases (Y. T. Cheng & X. Li 2012). Several host-translocated HECT E3 ligases that interfere with the host immune system have been described in bacterial pathogens of mammals (Anderson & Frank 2012). For example, NleL (Non-Lee-Encoded effector Ligase) from enteropathogenic *E. coli* O157:H7 is a HECT E3 ligase that is secreted as a T3SS effector required for virulence by reducing pedestal formation (Piscatelli et al. 2011). *Salmonella enterica* is a Gram-negative bacterium that secretes effectors to survive inside the host cell and cause inflammation; one such effector is the HECT E3 ligase SopA (Ying Zhang et al. 2006).

Both SopA and NleL have been co-crystalized with the E2 conjugating enzyme UbcH7 (Lin et al. 2012). Interestingly, unlike eukaryotic HECT E3 ligases, bacterial HECT E3 ligases interact with E2 conjugating enzymes at the opposite end of the N-terminal lobe. However, the same residues on UbcH7 define the interaction surface for both eukaryotic and bacterial HECT domain interactions. The implication of this finding is currently unknown, but it suggests that bacterial HECT E3 ligases employ different forms of molecular mimicry to hijack the host ubiquitin proteasome pathway.

In plants, over 6% of the proteins in the proteome are associated with protein turnover and degradation (Vierstra 2009). For this reason it is not surprising that ubiquitination plays a central role in plant signaling and development. For example,

organ size is controlled by the ubiquitin receptor DA1 and the RING E3 ligase Big brother (Y. Li et al. 2008). Interestingly, many plant hormones require the ubiquitin proteasome system for signaling (Santner & Estelle 2010). A prominent case is TIR1 (Transport Inhibitor Response 1) the auxin receptor, which encodes an F-box protein that upon binding of auxin targets AUX/IAA proteins for degradation (Gray et al. 2001). Hormonal induced protein degradation and signaling represents favorable susceptibility targets for pathogens. The jasmonic acid isoleucine (JA-Ile) receptor COI1 (coronatine insensitive 1) is required for degradation of JAZ (jasmonate ZIM domain) proteins that are transcriptional repressors, which inhibit the transcription factor AtMYC2 (Robert-Seilaniantz, Grant, et al. 2011a). Stimulation with JA-Ile leads to a COI1 dependent degradation of JAZ proteins and induction of jasmonate responses by AtMYC2. This induction results in production of NAC transcription factors that repress *ICS1* (Isochorismate Synthase) and induce *BSMT1* (SA methyl transferase 1), which are components of SA biosynthesis and metabolism respectively (Zheng et al. 2012). Jasmonate responses are induced upon recognition of necrotrophic pathogens and result in suppression of biotrophic defense (Robert-Seilaniantz, Grant, et al. 2011a). The hemi-biotrophic bacterium *P. syringae* secretes coronatine that mimics JA-Ile to suppress biotrophic defenses, which creates a favorable environment for the bacterium (Mittal & Davis 1995).

While these examples describe pathogens using the host ubiquitination machinery to their own advantage, protein degradation is also required for successful plant defense. For example, the tobacco E3 ligase CMPG1 is required for successful Avr9/Cf-9, Pto/AvrPto and INF1 resistance responses (González-Lamothe et al. 2006). Further, ubiquitination has also been implicated in R-protein signaling, as MOS5/UBA1 and the U-box proteins MAC3a/MAC3b are required for *snc1*

autoactivation of defense and disease resistance (Goritschnig et al. 2007; Monaghan et al. 2009).

Given the central role of the proteasome in plant signaling, it is unsurprising that pathogens have exploited this to their advantage. Examples of eukaryotic effectors from plant pathogens being delivered to the plant cell to hijack this system have currently not been reported. However, Lévesque et al (2010) found that many HECT E3 ligases were present in the *P. ultimum* proteome and 16 were secreted suggesting that the host ubiquitin proteasome system could be targeted.

Here, we show that HECT E3 ligases are an expanded group of proteins in oomycetes compared to diatoms, fungi, and plants. Interestingly, a subset of these is secreted including CHXC1 from *Albugo laibachii*. We characterized CHXC1 and found that CHXC1 localizes to the plant cell nucleus and enhances the virulence of *Hpa* Noco2 in *Arabidopsis.* However, expression of CHXC1 in *N. benthamiana* did not enhance the virulence of two different *Phytophthora* isolates. Further, we show that the virulence of CHXC1 depend on the predicted catalytic residue cys651 by using *Pst* DC3000 ΔAvrPto/ΔAvrPtoB in EDV experiments, and *Hpa* Noco2 infections of estradiol inducible *Arabidopsis* lines. Finally we find that the stability of CHXC1 was slightly increased by treatment with the proteasome inhibitor MG132 dependent on cys651.

## 4.2 Results

### 4.2.1 CHXC1 confers enhanced susceptibility in *Arabidopsis*

We expressed CHXC1 (Genbank: CCA15224.1) directly in plants to determine if the enhanced virulence of *Pst* DC3000 ΔAvrPto/ΔAvrPtoB harboring ΔSP-CHXC1::pEDV6 on Col-0 was due to an effect of CHXC1 perturbing defense inside the plant host cell.

We transformed an N-terminally GFP-tagged ΔSP-CHXC1 (without signal peptide) under a 35S promoter into Col-0. The $T_1$ transformants were small in size. When we selected the $T_2$ progeny for resistance to kanamycin all lines segregated for resistance. However, the plants were not reduced in size as previously observed. We screened for GFP fluorescence and found that 14 of 42 lines had detectable GFP-CHXC1 fluorescence. Interestingly, the GFP signal segregated within the population (n>50 plants analyzed per line): only 4 lines expressed the construct in all plants.

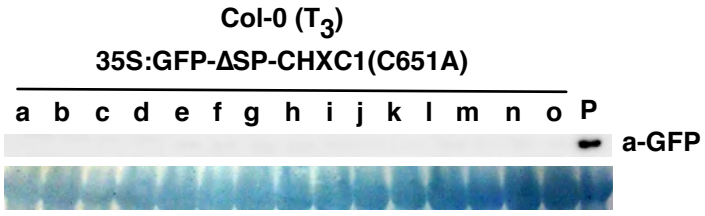



**Figure 4.1: 35S:GFP-ΔSP-CHXC1 protein is not observed in Col-0 $T_3$ transformants.** Top panel: membranes were probed with anti-GFP and developed with HRP. lanes a-o are 15 individual $T_2$ lines transformed with 35S:GFP-ΔSP-CHXC1 construct. P is a GFP positive control. Bottom panel: Amido-black staining show equal loading of protein.

We generated $T_3$ homozygous lines, but were unable to obtain lines that expressed amounts of GFP-ΔSP-CHXC1 that were detectable by fluorescence microscopy. More than 80 $T_3$ lines were tested and approximately a third were homozygous. A subset of the $T_3$ lines was tested for presence of the protein in total plant extracts by western blotting with α–GFP. However, we were unable to detect a specific signal corresponding to GFP-ΔSP-CHXC1 with the antibody (Figure 4.1).

As we were unable to obtain homozygous $T_3$ lines expressing GFP-ΔSP-CHXC1; we analyzed segregating $T_2$ lines. Initially, we tested 5 $T_2$ lines that were strongly GFP fluorescent for presence of the GFP-ΔSP-CHXC1 fusion protein by Western blotting. The α–GFP antibody reacted with a band at ~110 KDa that corresponds to the GFP-

ΔSP-CHXC1 fusion protein. Smaller bands also reacted with the α–GFP antibody, which we assume was partially degraded protein (Figure 4.2). As some GFP-ΔSP-CHXC1 fusion protein was full length in some lines, we used these lines for pathology experiments. To ensure all lines carried the transgene we selected segregating $T_2$ lines for kanamycin resistance and transferred these to soil.

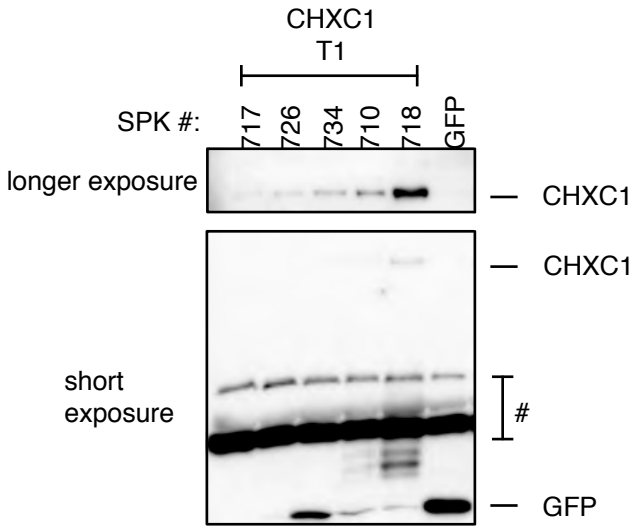



**Figure 4.2: Full length CHXC1 is degraded *in planta*.** Col-0 T1 plants expressing ΔSP-CHXC1 or EV in pK7WGF2 were selected on GM Kanamycin plates and subsequently transferred to soil. 4 weeks old tissue was sampled and separated on 8% v/v SDS-PAGE and the membrane probed with Hrp coupled anti-GFP. # denote unspecific cross reacting bands. SPK denote the transgene line tested.

Four-week-old plants were infected with *Hpa* Noco2; the infection was scored 6 dpi using the uvitex 2B method (Chapter 7). We observed enhanced virulence of *Hpa* Noco2 on Col-0 plants expressing GFP-ΔSP-CHXC1 compared to GFP (Figure 4.3). This was evident as a four-fold increase in area of the leaf, which was covered with conidiophores on Col-0 35S:GFP-ΔSP-CHXC1 compared to Col-0 35S:GFP.

The $T_2$ lines were selected on kanamycin, which could interfere with plant growth or *Hpa* infection. Thus, we performed an experiment without pre-selection of transgene carrying plants on kanamycin plates. Unselected soil-grown segregating $T_2$ seedlings

were infected with *Hpa* Noco2; the number of conidiophores per cotyledon was determined 5 dpi (Figure 4.3). Consistent with results on true leaves, cotyledons expressing GFP-ΔSP-CHXC1 were more susceptible to *Hpa* Noco2 than Col-0 plants.

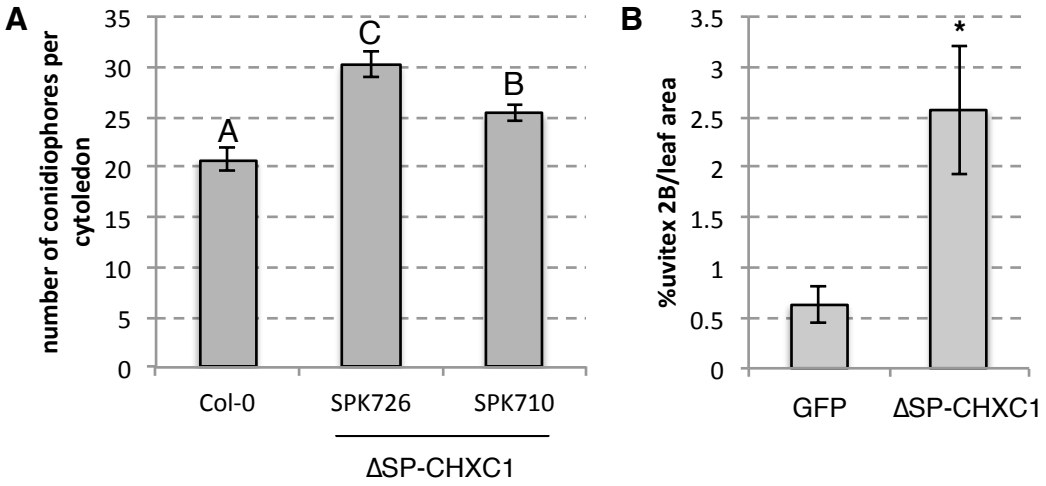



**Figure 4.3: *In planta* expressed CHXC1 enhanced virulence of *Hpa* Noco2.** Col-0 $T_2$ plants expressing ΔSP-CHXC1 in pK7WGF2. A) 10 day old cytoledons were infected with *Hpa* Noco2 and scored 5 dpi for each line more than 30 cytoledons were scored. B) Segregating T2 plants expressing EV or ΔSP-CHXC1 in pK7WFG2 were selected on MS+Kan plates for resistance to Kan and GFP signal transferred to soil. The plants were infected with *Hpa* Noco2 when 3 weeks old and scored 6 dpi. 18 leaves per line were analysed. Letters denote $p<0.05$ by Tukey post hoc A) or * 2-tailed t-test B). Error bars denote standard error. The experiment has been repeated once.

### 4.2.2 CHXC1 localizes to the nucleus

In summary, these data suggest that CHXC1 when expressed *in planta* conferred enhanced virulence of *Hpa* Noco2. Given CHXC1 also enhanced *Pst* DC3000 ΔAvrPto/ΔAvrPtoB virulence when delivered via T3SS, this suggested that CHXC1 inside the plant cell confers enhanced disease susceptibility of *Arabidopsis* Col-0 plants.

When effectors are secreted into the plant cell they target various compartments (Caillaud, Piquerez, Fabro, et al. 2012b); knowledge about the localization gives indications regarding a potential function.

We used confocal microscopy to determine the subcellular localization of *in planta*-expressed GFP-ΔSP-CHXC1. This protein localized to the nucleus and was excluded from the nucleolus (Figure 4.4). Using nuclear localization prediction tools we were unable to identify a nuclear localization signal (NLS) (Nguyen Ba et al. 2009). However, a stretch of positively amino acids $K_{164}$KFSQRQRGAQRRK$L_{178}$ may function as a mono-partite NLS.

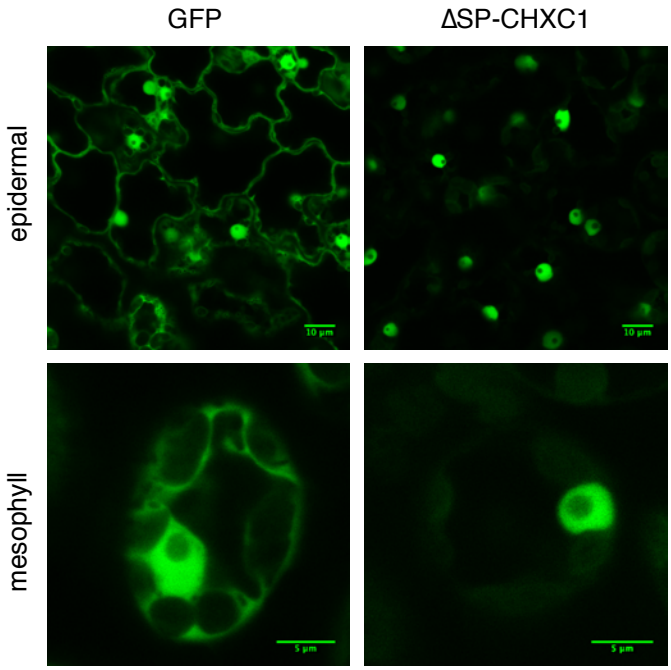



**Figure 4.4: CHXC1 localises to the nucleus.** Stable Col-0 transformants expressing either ΔSP-CHXC1 or GFP in pK7WFG2 were analysed by *in vivo* confocal microscopy 14 days post germination. CHXC1 localises to the nucleus (right), whereas GFP is nuclear and cytoplasmic (left) in epidermal cells (top panels) and mesophyll cells (bottom panels). More than 50 cells were analysed. The experiment was repeated three times with a similar result.

### 4.2.3 CHXC1 encodes a HECT E3 ligase

We queried the full-length sequence of CHXC1 against the PFAM database and found a C-terminal HECT domain. The catalytic site of these enzymes is highly conserved; this enabled us to identify the catalytic cysteine residue by creating an alignment of

CHXC1 with sequences of successfully crystalized HECT domains (Figure 4.5). The alignment, PFAM and PROsite predictions suggest that cys651 facilitates formation of the ubiquitin-thioester-intermediate (de Castro et al. 2006; Punta et al. 2011).

We generated a C651A mutation in the putative catalytic residue to determine, if the catalytic activity of CHXC1 was required for virulence. The mutant ΔSP-CHXC1(C651A) was introduced into a pEDV6 plasmid for EDV experiments.

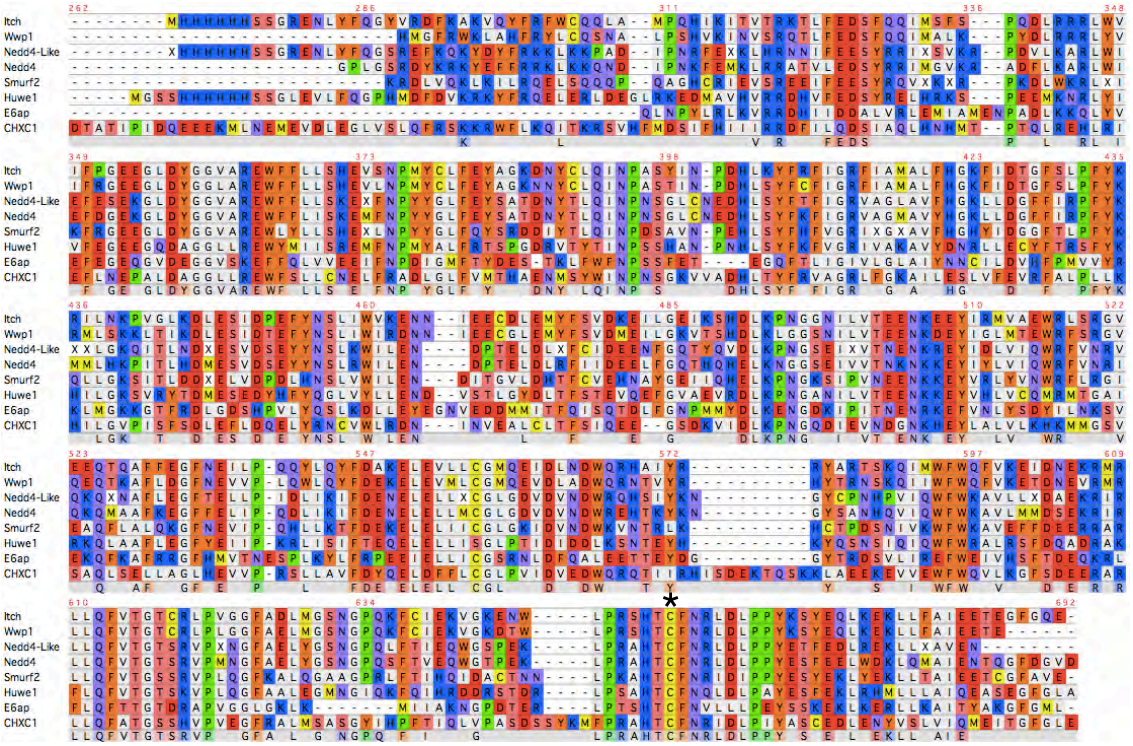



**Figure 4.5: The predicted catalytic residue cys651 is conserved in the CHXC1 HECT domain.** Alignment of CHXC1 from residue 262 to 684 (C-terminus) together with crystallised HECT domains. Sequences the HECT domains of Huwe1 (GI:270346506), Smurf2 (GI: 75765918), Nedd4-like (GI:146387319), Itch (GI:350610814), Nedd4 (GI:326634047), Wwp1 (GI: 37926893), E6AP (GI:6573516), CHXC1(GI:325180814) were downloaded from NCBI and aligned using ClustalW with default parameters. * denote the putative catalytic cystein residue in CHXC1 and other HECT domains.

We spray-infected Col-0 plants with either *Pst* DC3000 ΔAvrPto/ΔAvrPtoB harboring ΔSP-CHXC1(C651A)::pEDV6, ΔSP-CHXC1::pEDV6, or AvrRps4(AAAA)::pEDV6 and measured the growth at 4 dpi. Colony counts of *Pst* DC3000 ΔAvrPto/ΔAvrPtoB harboring ΔSP-CHXC1(C651A) were not significantly different from AvrRps4(AAAA), whereas counts of those carrying ΔSP-CHXC1 were significantly higher (Figure 4.6).

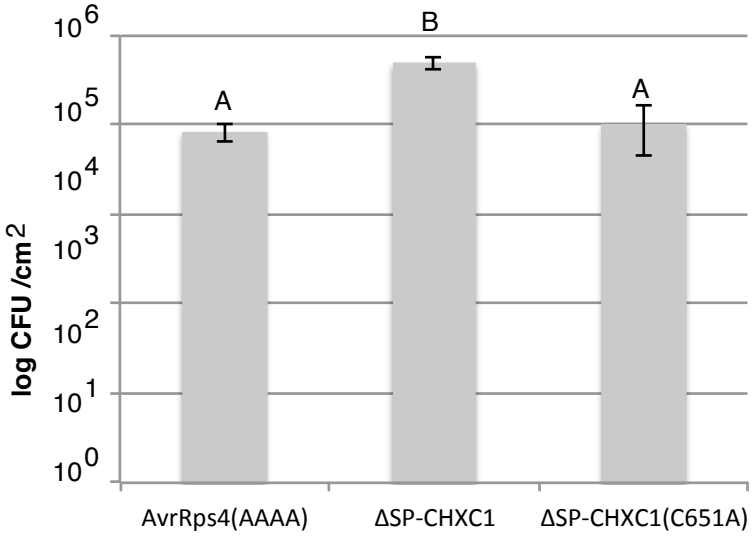



**Figure 4.6: Enhanced virulence of *Pst* DC3000 ΔAvrPto/ ΔAvrPtoB carrying CHXC1 depend on Cys651 in *Arabidopsis*.** Colony counts of 4 week old Col-0 plants spray inoculated with *Pst* DC3000 Δ*AvrPto*/Δ*AvrPtoB* harbouring either AvrRps4(AAAA), ΔSP-CHXC1, or ΔSP-CHXC1(C651A) in pEDV6. The plants were incubated for 4 days before sampling. The experiment was performed blind and with 8 replicate samples. The experiment was performed three times with similar results. Error bars denote standard error. Different letters (A,B) denote significant differences ($P<0.05$) assessed with a Tukey post hoc test.

ΔSP-CHXC1 confers enhanced virulence to *Pseudomonas* in EDV assays compared to ΔSP-CHXC1(C651A). We therefore generated estradiol inducible lines of ΔSP-CHXC1(C651) and ΔSP-CHXC1 in Col-0 and tested phenotypes after infection with *Hpa* Noco2. Interestingly, we found that expression of ΔSP-CHXC1 in Col-0 enhanced the virulence of *Hpa* Noco2, whereas ΔSP-CHXC1(C651A) was indistinguishable from YFP controls (Figure 4.7). We interpret these data as showing that CHXC1 cys651 is required for attenuating host resistance.

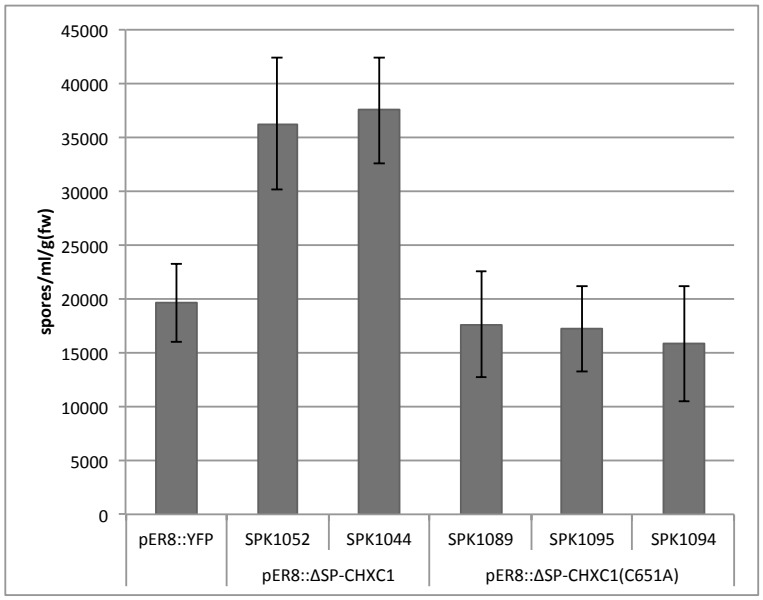



**Figure 4.7: Transient expression of ΔSP-CHXC1 but not ΔSP-CHXC1(C651A) enhance virulence of *Hpa* Noco2 in Col-0.** Three week old Col-0 $T_3$ plants carrying either estradiol inducible (pER8 vector) YFP, ΔSP-CHXC1, or ΔSP-CHXC1(C651A) were spray inoculated with 5 x $10^4$ spores/ml *Hpa* Noco2. The day prior to infection and 3 dpi the plants were sprayed with 2 μM β-estradiol. The infection was scored 6 dpi. Error bars denote standard error, and SPK number identifies the specific independent transgenic line used.

### 4.2.4 Expression of CHXC1 in *N. benthamiana* does not alter *P. infestans* virulence

In some cases, effectors target pathways or proteins that are conserved in other species. For example, *Agrobacterium tumefaciens* induced transient expression of the *Rhizobial* effector NopM in *N. benthamiana* results in flg22 dependent ROS suppression requiring catalytic activity (Xin et al. 2012). Therefore, we speculated that CHXC1 might be able to modulate the susceptibility target in *N. benthamiana* even though it is a non-host for *A. laibachii*.

To determine if CHXC1 localized to the nucleus in *N. benthamiana* as in *Arabidopsis*, we expressed GFP-ΔSP-CHXC1 or GFP in *N. benthamiana* using *Agrobacterium* transient assays. Consistent with our observations in *Arabidopsis* CHXC1 was distinctly nuclear localized in contrast to free GFP (Figure 4.8).

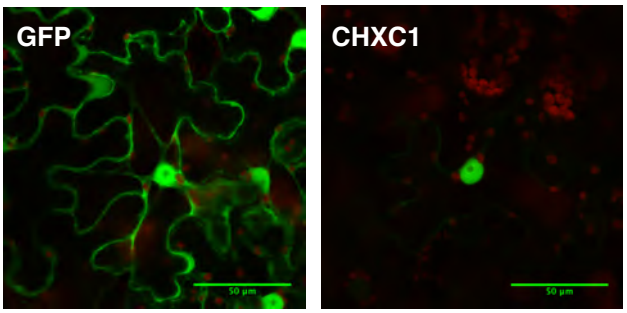



**Figure 4.8: CHXC1 localises to the nuclues in *N. benthamiana*.** Transient expression with Agrobacterium GV3101 pMP90 in N. benthamiana of N-terminally GFP (pK7WGF2) tagged ΔSP-CHXC1 (right) or (left) free GFP (Empty vector) were analyzed 2dpi. More than 20 cells were analysed. The experiment was repeated three times. Scalebar 50 $\mu$m.

This led us to perform an experiment, where we expressed *in planta* either ΔSP-CHXC1 or ΔSP-CHXC1(C651A) and then superinfected with *P. infestans* isolates. Our prediction was that ΔSP-CHXC1 should modify the susceptibility target, and thus result in enhanced virulence of *P. infestans* relative to ΔSP-CHXC1(C651A).

We expressed ΔSP-CHXC1 and ΔSP-CHXC1(C651A) in either side of the leaf to minimize variation between infections. The leaves were superinfected with the *P. infestans* blue 13 or *P. infestans* NL07434 strains, which have a long and short biotrophic phase respectively (Tolga Bozkurt pers. Comm.). We scored the infection by determining lesion size at 6 dpi and did not observe a difference in virulence dependent on cys651 (Figure 4.9).

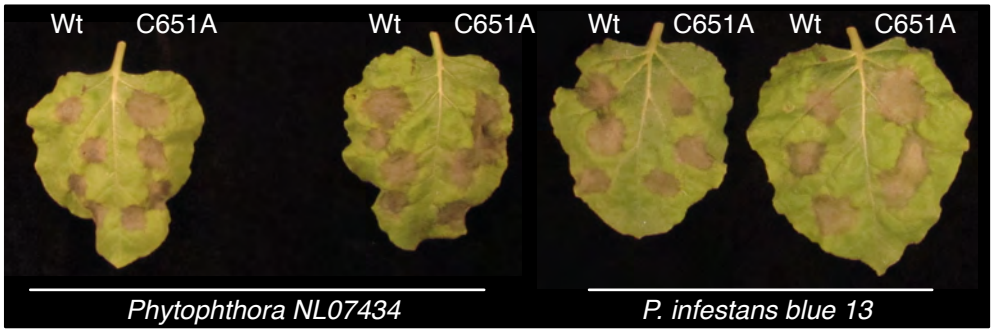



**Figure 4.9: Expression of CHXC1 C651A in *N. benthaminana* does not the alter virulence of *Phytophthora* strains.** Representative leaves of transiently expresssed ΔSP-CHXC1 (wt) or ΔSP-CHXC1(C651A) (C651A) in pK7WGF2 in *N. benthamiana*. Expression was induced 2 days proior to superinfection with *Phytophthora* NL07434 or *P. infestans* blue 13. Disease progression was scored at 6 dpi. The experiment was repeated twice with a similar result.

### 4.2.5 Expression and purification of CHXC1

PFAM predictions suggested that CHXC1 is a HECT E3 ligase and Cys651 is required for CHXC1 function. This indicated that CHXC1 is a functional HECT E3 ligase, which ubiquitinates host proteins inside the plant cell. To determine if CHXC1 has *in vitro* ubiquitination activity, we devised a strategy to obtain CHXC1 protein.

We attempted soluble expression in two expression systems: *Pichia pastoris* and *E. coli.* We initially tested various CHXC1 truncations in *E. coli* for soluble protein expression. We cloned ΔSP-CHXC1 into pOPIN-F and transformed this into 3 *E. coli* expression strains: *E. coli* BL21, *E. coli* soluBL21, and *E. coli* soluBL21 harboring pRARE. Expression was induced at 15°C, 28°C and 37°C and high quantities of completely insoluble CHXC1 were obtained (data not shown). The construct was N-terminally 6xHis tagged, which could interfere with protein solubility. Fusions to the maltose binding protein (MBP) has been extensively used to solubilize proteins (Walls & Loughran 2011). For this reason, we created an MBP-ΔSP-CHXC1 fusion construct. However, this construct expressed markedly worse and was still insoluble (data not shown).

The HECT domain of CHXC1 is highly conserved. This allowed us to generate putative structures of CHXC1 using the Swiss-model automated homology server (Arnold et al. 2006). Two hypothetical structures of CHXC1, based on the crystal structure of the human NEDD4-like protein (PDB: 2oniA, unpublished) and the yeast HECT E3 ligase (PDB: 3olmA) (H. C. Kim et al. 2011) were predicted.

Since these proteins were successfully expressed and crystalized, we created two constructs of CHXC1 based homology to the crystal structures and predicted secondary structure. These constructs, Δ157-CHXC1 and Δ285-CHXC1, were cloned into pOPIN-M and pOPIN-F and expressed in *E. coli* BL21. While we did observe expression, the fusion proteins were insoluble and levels were lower than 6x-HIS-ΔSP-CHXC1 in pOPIN-F (data not shown).

As CHXC1 is secreted from *A. laibachii* it could receive secondary modifications during the passage from the ER to Golgi and subsequent secretion, which may be important for function and solubility. The eukaryote *Pichia pastoris* has been established as a system for high-level protein expression of secreted proteins. We took advantage of this system and cloned CHXC1 from its signal peptide cleavage site to translation termination into pPICZalphaA. This created a α–factor-ΔSP-CHXC1 fusion construct, which was transformed into *Pichia pastoris.* In this system the strong *AOX1* promoter drives transgene-expression. We attempted expression and while AOX1 was strongly induced, we did not detect CHXC1 protein in the culture media, or the *Pichia pastoris* cell pellets (data not shown).

### 4.2.6 Refolding of CHXC1 from inclusion bodies

Further attempts to create a soluble construct were abandoned; instead, we choose to refold insoluble 6xHIS-ΔSP-CHXC1 expressed in *E. coli* SoluBL21. Approximately 90% of eukaryotic proteins that have been expressed in *E. coli* are insoluble, but in some cases they can be refolded (Burgess 2009; Gräslund et al. 2008). Insoluble proteins are generally aggregates of misfolded proteins. Misfolded proteins are thought to be produced when a protein is produced at too high a rate, and the folding machinery is unable to correctly fold proteins. Refolding relies on the principle that given enough time and proper conditions a linearized protein will fold correctly. To refold a protein, it should initially be partially or fully denatured and solubilized. Subsequently, the denaturant is removed and the protein is allowed to slowly refold (Burgess 2009).

Sarkosyl is a zwitterionic detergent, which has been successfully used to denature proteins prior to refolding (Tao et al. 2010). We devised a strategy to isolate inclusion bodies of CHXC1 from *E. coli* and then wash them in 1% v/v triton X-100 to remove membranes and cytoplasmic proteins. The 6xHis-ΔSP-CHXC1 was resolubilized with 1.5% v/v sarkosyl in PBS (Figure 4.10). The resolubilised material was flash diluted 21 fold into PBS and incubated overnight. We removed aggregated proteins by passing the liquid through a 22 µM filter and loaded the eluate onto a HIS-trap column. The column was washed with 5 column-volumes of PBS containing 25 mM imidazole to reduce unspecific binding. CHXC1 was eluted in PBS + 250 mM imidazole, and afterwards concentrated 4x in an amicon column.

We analyzed key fractions by western blotting, with α–histidine and α–CHXC1 antibodies (raised in mouse and guinea pig respectively), and found that our procedure resulted in production of soluble 6xHis-ΔSP-CHXC1 (Figure 4.10).

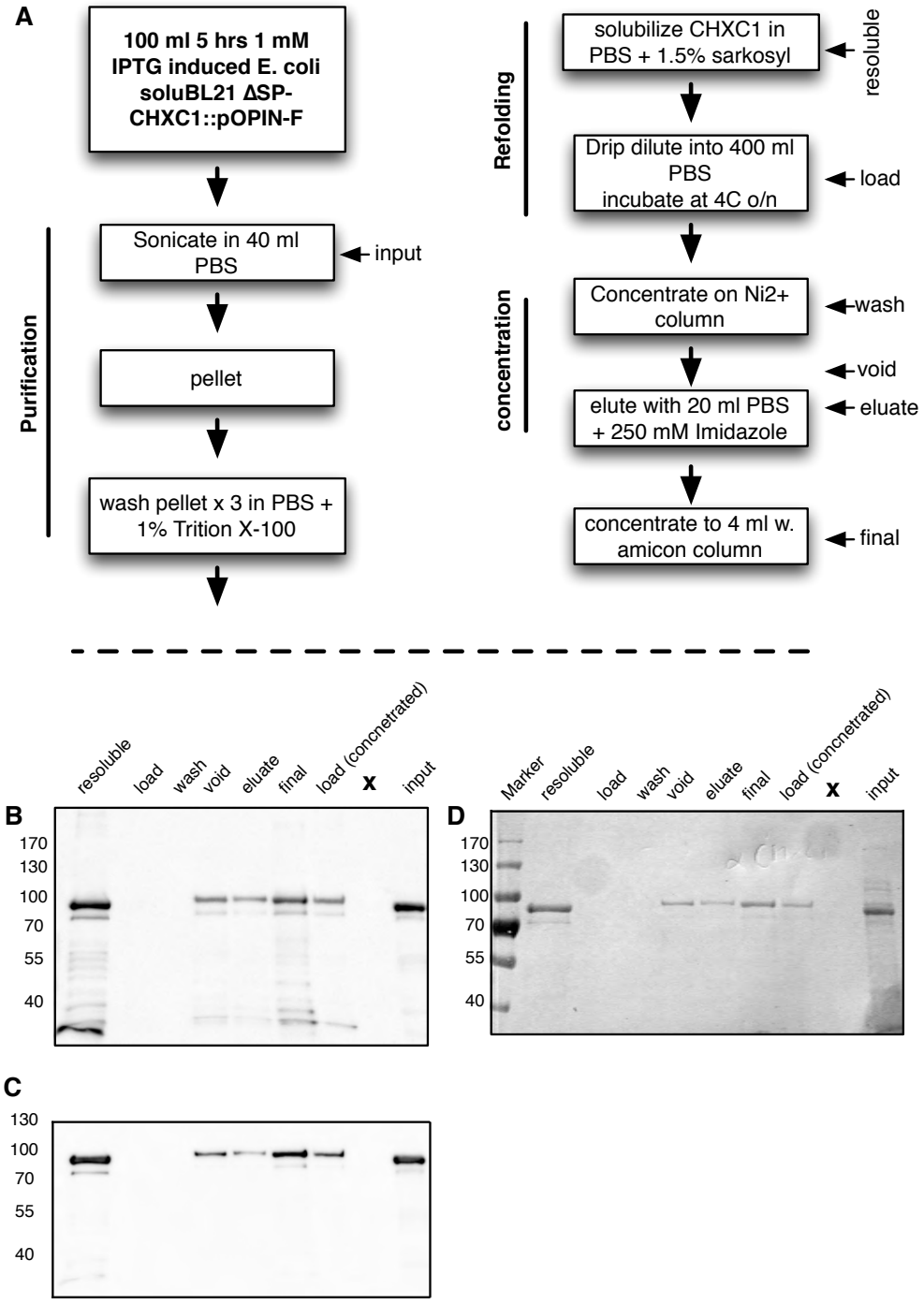



**Figure 4.10: CHXC1 can be refolded in sarkosyl and purified via Ni2+ columns.** A) A schematic representation of the procedure followed to obtain purified soluble CHXC1. Vertical arrows shows at what point samples were obtained in the purification process. The samples were analysed by western blotting and probed with either a B) anti-His or a C) anti-CHXC1 antibody. To determine amounts of protein in fractions the membrane was amindo black stained for proteins D).

Amidoblack staining of the membrane revealed that the purification was not completely pure. The contaminants were likely endogenous *E. coli* proteins and degradation products that bound to the HIS-trap column and co-eluted with 6xHis-ΔSP-CHXC1.

We performed a mass spectrometry analysis to independently determine if the purified 6xHIS-ΔSP-CHXC1 was full length. Mass spectrometry confirmed the identity of 6xHis-ΔSP-CHXC1 as we obtained peptide-spectra covering both ends of the protein (data not shown). Thus, we had established a method to purify and obtain soluble 6xHis-ΔSP-CHXC1 from inclusion bodies and confirmed this by three independent methods.

To determine if CHXC1 is a HECT E3 ligase *in vitro* we incubated the purified protein in an ATP containing buffer with recombinant E1 protein, recombinant ubiquitin, and one of three different recombinant E2s. We tested: UbcH7, UbcH5a, and UbcH5b. To date we have been able to show that the reagents work using the RING E3 ligase EOD1/BB (Big brother, a kind gift from Jack Dumenil, unpublished). However, we were unable to detect any E3 ligase activity of the 6xHis-ΔSP-CHXC1 protein.

The *in vitro* experiments failed, which could be due to use of E2 conjugating enzymes that do not function with CHXC1. For example, Kraft and coworkers (Kraft 2005) tested 10 different E2 ubiquitin conjugating enzymes against an array of 64 RING E3 ligases for autoubiquitination *in vitro*; they found that not all combinations of E2 and E3 enzymes results in ubiquitination. Thus it is conceivable that we have not found the correct E2 conjugating enzyme.

We hypothesized that the correct E2 conjugating enzyme would interact with CHXC1. For this reason, we set out to find CHXC1 interacting proteins. We cloned ΔSP-CHXC1 and ΔSP-CHXC1(C651A) into pLexA and transformed them into the yeast strain

EGY488[p8op-lacZ], but when tested for autoactivation, both constructs auto-activated (Lennart Wirthmueller pers. comm.).

The pLexA system is based on the LEU2 and LacZ reporters, whereas the GAL4 based system uses the HIS3 reporter (Van Criekinge & Beyaert 1999). The HIS3 reporter encodes imidazole glycerol phosphate dehydratase, which is competitively inhibited by 3-amino-1,2,4-triazole (3-AT) (Hilton et al. 1965). Titration with 3-AT can limit histidine biosynthesis and growth of yeast cells, and is therefore often used to reduce growth arising from weak auto-activation by GAL4-AD or GAL4-BD fusion proteins.

To take advantage of this system, we cloned ΔSP-CHXC1 and ΔSP-CHXC1(C651A) into pDB and pAD and transformed this into yeast. We did not observe autoactivation in this system (J. Steinbrenner pers. comm.). Currently, we are awaiting the results of the yeast interactor screening, which could identify a correct E2 interaction partner and potential target substrate(s) of CHXC1.

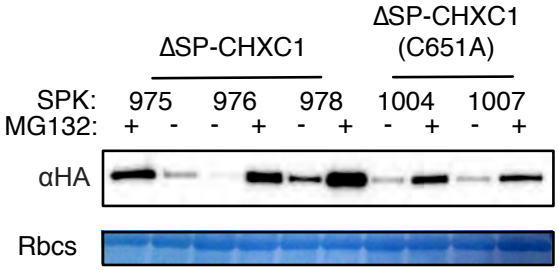
### 4.2.7 Stability of CHXC1 upon MG132 treatment is partially dependent on cys651 *in vivo*

While an *in vitro* experiment could demonstrate enzymatic activity dependent on a defined set of components, it does not prove that CHXC1 is a functional HECT E3 ligase *in vivo.* We transformed 3xHA-ΔSP-CHXC1 and 3xHA-ΔSP-CHXC1(C651A) under an estradiol inducible promoter into Col-0 to test for activity *in vivo*.

Both constructs yielded approximately identical numbers of $T_1$ transformants, but only 3 out of 15 ΔSP-CHXC1 transformants produced protein when expression was induced with estradiol compared to 2 out of 5 ΔSP-CHXC1(C651A). In addition, the ΔSP-CHXC1 signal was much reduced compared to ΔSP-CHXC1(C651A).

HECT E3 ligases can autoubiquitinate when overexpressed *in vivo*, or tested *in vitro* and poly-ubiquitinated proteins are degraded via the 26S proteasomal pathway (Voges et al. 1999). This led us to speculate whether the apparent lack of ΔSP-CHXC1 expressing transformants was caused by proteasomal degradation of autoubiquitinated CHXC1, where catalytic autoubiquitination activity required cys651.

To test this hypothesis we blocked proteasomal degradation with MG132, which is a Z-Leu-Leu-Leu-al oligopeptide and a potent proteasome inhibitor (D. H. Lee & Goldberg 1998). Specifically, we took 3 ΔSP-CHXC1 and 2 ΔSP-CHXC1(C651A) transformants and induced these with estradiol supplemented with either DMSO or 10μM MG132. 24 hours post induction leaf discs were sampled and protein levels assessed by western blotting (Figure 4.11). We determined the intensities of the protein bands and compared the ratio of MG132 treated leaf discs to DMSO controls. Interestingly, we found a 3-6 fold increase in ΔSP-CHXC1 protein, whereas ΔSP-CHXC1(C651A) protein levels increased approximately 2 fold.




| | SPK | Treatment | ImageQuant | MG132/DMSO |
|---|---|---|---|---|
| ΔSP-CHXC1 | 975 | MG132 | 30996727 | 2.95 |
| | | DMSO | 10490675 | |
| | 976 | DMSO | 6070374 | 6.05 |
| | | MG132 | 36720108 | |
| | 978 | DMSO | 17382457 | 4.11 |
| | | MG132 | 71398035 | |
| ΔSP-CHXC1 (C651A) | 1004 | DMSO | 10284952 | 2.20 |
| | | MG132 | 22635192 | |
| | 1007 | DMSO | 9558387 | 1.94 |
| | | MG132 | 18562840 | |

**Figure 4.11: The stability of CHXC1 partially depends on cys 651 *in planta*.** The plants were selected on MS agar plates supplemented with 25μg/ml gentamycin and pricked out into soil after 2 weeks. The plants were grown in soil two weeks prior to infiltration of leaves with 20μM estradiol supplemented with either 10μM MG132 or DMSO. Infiltrated tissue was collected 20 hours post treatment using a cork borer size 3, taking one leaf disc from different 4 plants. Proteins were extracted and analysed by western blotting with aHA-HRP. Relative protein levels were detected using Imagequant software. SPK denote transgene line.

### 4.2.8 Secreted HECT E3 ligases are only found in oomycetes

Pathogenic bacteria have long been known to secrete effectors to modulate their host's ubiquitination machinery to their advantage (Abramovitch et al. 2006; Lin et al. 2011). *Pst* DC3000 secretes AvrPtoB, which encodes a novel E3 ligase that ubiquitinates CERK1 and FLS2 marking them for degradation by the host proteasome (Gimenez-Ibanez et al. 2009; Göhre et al. 2008). Most eukaryotic pathogens have a functional ubiquitin proteasome system, but effectors exploiting the host ubiquitin proteasome system have not been identified. CHXC1 conferred enhanced disease susceptibility when expressed inside the host; we reasoned that other eukaryotic pathogens could employ a similar mechanism.

To define if CHXC1 type secreted HECT E3 ligases are present in other eukaryotes we downloaded 4 plant, 5 fungal, and 13 protist proteomes; 9 of the protist proteomes belonged to the oomycota (Table 4.1). 16 of the organisms had a parasitic lifestyle, whereas 8 had an auxotrophic lifestyle. The non-parasitic auxotrophs belong to 3 different phyla; the presence of secreted HECT E3 ligases in these organisms would argue for this being a general mechanism. To determine if secretion of HECT E3 ligases is a general strategy employed by pathogens we included 5 non-oomycete pathogens in the analysis.

We employed a HMM based search strategy to define the distribution of HECT E3 ligases in the proteomes using a PFAM HECT domain profile (PF00632).

The 4 fungal and the protists not belonging to the oomyctes (*P. falciparum, T. gondii, D. discoideum,* and *Th. Pseudonana*) had less than 10 HECT E3 ligases encoded in their proteome. Similarly, between 10-18 HECT E3 ligases were encoded in plant proteomes (Table 4.1). This was significantly less than 16-58 HECT E3 ligases

encoded in the oomycete proteomes (W = 33.5, p-value = 0.02012, two-tailed Wilcoxon rank-sum test).

**Table 4.1: Secreted putative HECT E3 ligases are found in oomycetes.** HECT E3 ligases from the proteomes of the organisms in various taxonomic groups were tested for presence of secretion signal.

| | HECT E3 ligases | |
|---|---|---|
| **Organism** | **All** | **Secreted** |
| **Plants** | | |
| *Arabidopsis lyrata* | 11 | |
| *Arabidopsis thaliana* | 12 | |
| *Physcomitrella patens* | 18 | |
| *Chlamydomonas reinhardtii* | 10 | |
| ***Fungi*** | | |
| *Magnaporthe oryzae* | 6 | |
| *Ustilago maydis* | 5 | |
| *Puccinia graminis* | 9 | |
| *Schizosaccharomyces pombe* | 7 | |
| *Saccharomyces cerevisiae* | 5 | |
| **Protists** | | |
| *Plasmodium falciparum* | 4 | |
| *Toxoplasma gondii* | 9 | |
| *Dictyostelium discoideum* | 6 | |
| *Thalassiosira pseudonana* | 7 | |
| **Protists - Oomyctes** | | |
| *Saprolegnia parasitica* | 46 | 9 * |
| *Pythium ultimum* | 58 | 15 * |
| *Albugo candida Nc2* | 16# | 1 |
| *Albugo laibachii Nc14* | 16 | 1 |
| *Phytophthora capsici* | 30 | 3 * |
| *Phytophthora infestans* | 36 | 7 * |
| *Phytophthora ramorum* | 29 | 9 * |
| *Phytophthora sojae* | 30 | 8 * |
| *Hyaloperonospora arabidopsidis* | 18 | 1 * |

# there are 27 gene models, but if you only consider single loci and not putative splice variants, then you get 16. * Some have an RXLR motif within the first 80 amino acids after signal peptide cleave site.

We tested for whether other HECT E3 ligases carry a secretion signal peptide. Interestingly, secreted HECT E3 ligases are exclusively found in oomycete species. Further, necrotrophic oomycetes carry between 3 and 15 secreted HECT E3 ligases, whereas only a single one is encoded in biotrophic oomycetes (Table 4.1).

In *Albugo laibachii* this was CHXC1, interestingly the *Hpa* ortholog was identified as RXLRNEE7. Sequence analysis revealed that all oomycete species except *Albugo* sp. and *P. ultimum* had some secreted HECT E3 ligases, which had an RXLR motif within the first 80 amino acids. This suggests that HECT E3 ligases could be important virulence proteins in other plant pathogenic oomycetes.

## 4.3 Discussion

We describe an *Albugo laibachii* effector CHXC1 that belongs to an evolutionarily conserved class of putative secreted HECT E3 ligases. The nuclear localized CHXC1 depend on the predicted catalytic residue Cys651 for virulence. As homologs of CHXC1 are found in other oomyctes, we propose that secretion of proteins with homology to HECT E3 ligases represent an oomycete specific method for suppressing host defense.

### 4.3.1 Is CHXC1 a functional HECT E3 ligase?

Despite substantial efforts we were unable to conclusively show that CHXC1 is a functional HECT E3 ligase. We have demonstrated that the stability of ΔSP-CHXC1 was more increased than ΔSP-CHXC1(C651A) upon treatment with MG132 *in vivo*. While, this could be indicative of ubiquitin dependent proteasomal degradation dependent on cys651, the effect was not dramatic. Thus, it is possible that the observed effect is due to a general inhibition of protein turn over via the proteasome by MG132. The absence of higher molecular weight ΔSP-CHXC1 molecules arising from autoubiquitination on CHXC1 lysine residues supports this notion.

However, CHXC1 has high homology to HECT E3 ligases and virulence assays with *Pst* DC3000 Δ*AvrPto/ΔAvrPtoB* and *Hpa* Noco2 demonstrated a requirement for cys651.

While cys651 could be required for E3 ligase activity, it could also be required for interaction with host targets or proper folding of CHXC1. To test if CHXC1 is a functional HECT E3 ligase, we will pull down ΔSP-CHXC1 protein, from plants treated with the proteasome inhibitor MG132, and probe the immuno-precipitate with α-ubiquitin. This will determine if ΔSP-CHXC1 is ubiquitinated *in vivo*. If so, the identity of the ubiquitinated residues will be determined by mass spectrometry on the immuno-precipitated ΔSP-CHXC1 and ΔSP-CHXC1(C651A). The presence of ubiquitinated lysines dependent on cys651 will determine if CHXC1 is a functional HECT E3 ligase *in vivo*.

We conducted *in vitro* experiments with *E. coli* refolded ΔSP-CHXC1 to determine the minimum set of components required for CHXC1 mediated ubiquitination. To date, we have been unable to define suitable *in vitro* conditions perhaps because CHXC1 is not a functional HECT E3 ligase.

While the tested *in vitro* conditions do not allow CHXC1 autoubiquitination, it is possible that CHXC1 is specifically activated in the plant cell by certain conditions or proteins. To determine if an *in planta* modification is required, we will perform an experiment where CHXC1 expressed *in planta* is purified, and used as substrate for an *in vitro* ubiquitination assay. In addition, complementation of a yeast mutant with the CHXC1 HECT domain would demonstrate E3 ligase activity. Specifically, *Saccharomyces cerevisiae* requires the HECT E3 ligase *Rps5* for internalization of Hxt1 (Hexose transporter 1) upon stimulus with methylglyoxal. We will attempt to complement the mutant *YPH250, plc1Δ::HIS3 rsp5$^{wimp}$* with a Rps5$^{\Delta HECT}$-CHXC1$^{HECT}$ chimera, where the C-terminal HECT domain of Rps5 has been substituted with the CHXC1 HECT domain (Yoshida et al. 2012). Overall these assays will lead to a deeper understanding of CHXC1 function and potentially demonstrate enzymatic activity.

### 4.3.2 Are HECT E3 ligases core effectors deployed by pathogenic oomycetes?

Map based cloning of fungal and oomycete avirulence effectors has shown that these belong to a highly diverse class of short secreted proteins often without any sequence similarity to proteins of known function (Bozkurt et al. 2012). Prominent examples are: ATR1, ATR13, Avr3a, and more recently ATR5 (Rehmany et al. 2005; Armstrong et al. 2005; R. Allen et al. 2004; Bailey et al. 2011). These characteristics have been the corner stone in *de novo* bioinformatic predictions of effectors. These approaches have identified a plethora of potential species-specific effectors (Stassen & Van den Ackerveken 2011; Oliva et al. 2010). However, conserved effectors present in multiple species over a larger evolutionary time have not been studied.

We examined the distribution of HECT E3 ligases in fungi, protists, and plants. Intriguingly, our comparative proteomic analysis revealed that secreted HECT E3 ligases were unique to pathogenic oomycetes; this coincided with a dramatic expansion of the HECT E3 ligases encoded in the proteome.

10 or less HECT E3 ligases were encoded in *Th. pseudana* (diatom), *Pl. falciparum* (*Apicomplexa*), and *Ch. reinhardtii* (*green alga*) proteomes. We speculate that the expansion of HECT E3 ligases family(s) in oomycetes predates the split into sister species but occurred after the divergence of oomycetes from the stramenopiles (Beakes & Sekimoto 2009). The strategy seems to be unique to pathogenic oomycetes, as we did not find evidence of secreted HECT E3 ligases in plant pathogenic fungi.

Remarkably, there was a correlation between choice of lifestyle and number of secreted HECT E3 ligases encoded in the proteomes. A single secreted HECT E3 ligase

was present in *Hpa*, *A. laibachii* and *A. candida*, whereas an average of ~9 (range 3-15) HECT E3 ligases were found in *P. ultimum, P. capsici, P. infestans, P. sojae, P. ramorum* and *S. parasitica*. We interpret this finding as showing that obligate biotrophic oomycetes secrete a single HECT E3 ligase, whereas multiple are secreted by hemibiotrophic and necrotrophic oomycetes. Interestingly, all pathogenic oomycetes carried a homolog of CHXC1. As most pathogens have an initial biotrophic phase (Kelley et al. 2010), we speculate that CHXC1 or its ortholog is required to suppress defense and establish biotrophy. Presumably, other secreted HECT E3 ligases are required at the onset or during the necrotrophic phase in necrotrophic oomycetes.

While experimental characterization of multiple secreted HECT E3 ligases from other necrotophic oomycetes would shed light on the matter, we attempted to block or prolong the biotrophic phase using CHXC1. We did this by expressing ΔSP-CHXC1 or ΔSP-CHXC1(C651A) in *N. benthamiana* to enhance or decrease virulence of *P. infestans*. Contrary to expectation we did not find an alteration in virulence dependent on cys651. Several reasons could explain this finding: First, the level expression of ΔSP-CHXC1 or ΔSP-CHXC1(C651A) could be insufficient to alter the virulence effect of the cognate *P. infestans* HECT E3 ligase ortholog of CHXC1, as this may have enzymatic function. To address this, we will use the pEAQ-HT expression system to induce massive protein expression and infect with *P. infestans* (Sainsbury et al. 2009).

Further, effectors are highly specific and evolve together with their cognate host proteins. For example, *P. infestans* is a pathogen of *Solanum* sp., which secretes EPIC1 and EPIC2B to inhibit the host protease C14. Interestingly, the interaction is much stronger with C14 than with the tomato proteases PIP1 and RCR3 (Kaschani et al.

2010). It is therefore conceivable that target specificity to a *Brassicaceae* protein by CHXC1 results in lack of efficient suppression of defenses in *Solanaceous* plants.

However, it is also possible that CHXC1 fails to recruit the ubiquitin proteasome system to modify the susceptibility target. This could be tested by expressing CHXC1 or CHXC1(C651A) with and without MG132. If the changes in CHXC1 protein levels are insignificantly different from CHXC1(C651A) upon MG132 treatment this would suggest that is unable to recruit the ubiquitin proteasome system.

To test if CHXC1 could suppress defense in other *Brassicaceae* we will transiently express CHXC1 or CHXC1(C651A) in *Brassica oleracea* (Kate Bailey unpublished) and infect with Brassica downy mildew (Constantinescu & Fatehi 2002). These experiments could reveal if the effector targets are highly conserved across host and pathogen species.

### 4.3.3 Conclusion and outlook

In conclusion, we found that the effector CHXC1 from *Albugo laibachii* is nuclear localized; virulence assays demonstrated that CHXC1 enhanced virulence of *Pst* DC3000 *ΔAvrPto/ΔAvrPtoB* and *Hpa* Noco2 dependent on cys651. Further, in pathogenic oomycetes the HECT E3 ligase family is an expanded super family where some members are secreted; we propose that secretion of HECT E3 ligases to modulate host defenses represents a strategy general to all oomycetes. To our knowledge CHXC1 is the first example of a protein with homology to a eukaryotic HECT E3 ligase that is employed as an effector.

# 5 Variants of SSP6 are candidate effectors unique to *A. laibachii* that enhance virulence by suppression of PTI

## 5.1 Introduction

Oomycetes are a group of economically important plant pathogens because they infect a wide range of agricultural crops, such as *Phytophthora infestans* (tomato and potato), *P. sojae* (Soy), and *Albugo candida* (various brassica species). *Albugo laibachii* is phylogenetically closely related to *A. candida*; both species are obligate biotrophs and able to infect *Arabidopsis*. During the parasitic stage of *Albugo* sp. life cycle, the pathogen secretes effectors into the host cell to suppress defense and create an environment favorable to the pathogen (Dodds & Rathjen 2010). Studies of oomycete and bacterial effectors have shown they often target the PTI machinery and modulate it (Block & Alfano 2011; Stassen & Van den Ackerveken 2011). Consequently most effectors are presumed to enhance the virulence of a given pathogen, although the first effector proteins were identified based on avirulence functions resulting in host resistance (Armstrong et al. 2005; R. Allen et al. 2004; Rehmany et al. 2005; W. Shan et al. 2004; Bailey et al. 2011). Resistance to *A. laibachii* in *Arabidopsis* is rare with approximately 90% of the accessions tested showing susceptibility to *A. laibachii* Nc14 (*AlNc14*) (Kemen et al. 2011). As *A. laibachii* is remarkably well adapted, the *AlNc14*/*Arabidopsis* patho-system is a good system to study how pathogen effectors shape their hosts to aid infection.

Several oomycete species including *A. laibachii* have recently been sequenced and the availability of these genomes has stimulated bioinformatic efforts to predict their cognate effector repertoires (Raffaele & Kamoun 2012; Spanu 2012; Kemen & Jones 2012).

Several key findings have arisen from this community effort. Firstly the effector complements of oomycete genomes are large, ranging from 134 RXLR effectors in *Hpa* to 700 in *P. infestans.*

While, RXLR effectors are significantly over represented in the secretomes of *Hpa* and *Phytophthora* sp., this class of effectors is not overrepresented in the secretomes of other oomycetes such as *A. laibachii*, *Phytium ultimum, Aphanomyces euteiches* or *Saprolegnia parasitica*. Thus, contrary to evidence derived from classical cloning of oomycete effectors, which have exclusively been RXLR or RxLR-like effectors, there is a large diversity of effector classes (Bozkurt et al. 2012; Stassen & Van den Ackerveken 2011). Interestingly, crinkler-like proteins have been predicted in all oomycetes to date although generally in lower numbers than what has been reported for *P. infestans*.

In the *A. laibachii* genome, the presence of some CRNs and a large family of novel effectors, the CHXCs, were identified (Kemen et al. 2011). As no genetic evidence has been provided to underline the potential *A. laibachii* effector groups, a ranking system was devised to prioritize the secreted proteins without trans-membrane domains for potential virulence functions. Scores were assigned to short secreted proteins (SSPs) for features that might be expected to be present in an effector protein, as exemplified in the literature: we was hypothesized that effectors could have paralogs or be in repetitive regions as described in *P. infestans* (Haas et al. 2009); thus genes being on a short repetitive contig in the initial illumina assembly (≤3000 bp) or located at the end of a contig gave 1 pt. Since almost all avirulence proteins described to date are short proteins, SSPs ≤400 amino acids were assigned 2 pts. Full cDNA coverage gave 1 pt, while expression in *AlNc14* earlier than 10 days post infection gave 1 pt, and expression before day 4 gave a further 1 pt. SSPs with heterozygous positions within the gene were given 1 pt. Lastly if a known uptake motif

was identified (RXLR, RXLQ, CHXC, and CRN) this yielded 2 pts. The maximum score possible for an SSP was 9 pts.

In this ranking system SSP6 scored 6 points due to: being a protein smaller than 400 amino acids (2 pts), being heterozygous (1 pt), fully covered by cDNA tags (1 pt), expressed before day 10 (1 pt), and on a short contig in the initial illumina assemblies (1 pt). This put it in the top 10 SSP of the prioritized list.

Together with other candidate *A. laibachii* effector classes, this category of effectors were screened on *Arabidopsis* ecotypes to test if they would enhance virulence (Kemen et al. 2011, and Chapter 3). One interesting candidate (SSP6) enhanced virulence of *Pst* DC3000 when delivered to the plant cell via the EDV system on the *Arabidopsis* ecotypes Col-0 and Nd-0. In order to gain a better understanding of the interactions between *AlNc14* and its host *Arabidopsis*, this effector was studied in more detail.

Here, we report that 7 variants of SSP6 are encoded in the *AlNc14* genome and they show signatures of being under a positive selection pressure. Two variants of SSP6 (SSP6-2c and SSP6-A) are predominantly expressed and both localize to the plasma membrane. Interestingly, SSP6-2c, but not SSP6-A, is able to suppress flg22 dependent reactive oxygen species (ROS) bursts and enhance the virulence of *P. infestans* blue 13 in *N. benthamiana.* Localization of SSP6-2c and SSP6-A in *Arabidopsis* plants shows that SSP6-2c localizes around the haustorium of *AlNc14* upon infection. As we were unable to find homologs of SSP6 outside *A. laibachii*, we propose that variants of SSP6 are fast evolving species-specific effectors.

## 5.2 Results

### 5.2.1 Variants of SSP6 are encoded in a genomic region absent from *Ac2VRR*

We anticipated that SSP6 (GenBank: CCA18026.1) could be in a repeat rich region due to our SSP selection criteria. Since the *AlNc14* genome assembly is solely based on NGS, which perform poorly in such regions, we were interested to resolve this locus in more detail. To determine if the locus was heterozygous and properly resolved, we mapped all the Illumina reads to the genome assembly using MAQ (Eric Kemen pers. comm.). This was done two ways: first reads were aligned retaining the read-pair information, and subsequently the reads were aligned as single-reads. The average read depth of paired-end read data suggested the read depth was around 140 deep close to the global average of 120 deep. In contrast the read depths determined by aligning single-end reads that aligned to the SSP6 locus was approximately 400, which is close to double the average read depth (Figure 5.1). This suggested that two closely related loci had been collapsed in the assembly and one locus was correctly assembled. Alignments of single- and paired-end reads onto the genome gave slightly different SNP predictions. As single-end read alignment allowed fewer mismatches than alignment of paired-end reads, which allowed for a more ambiguous alignment of the reads but with high confidence due to the knowledge that the read pairs are physically connected, the paired-end read alignment could predict closely spaced SNPs. Consistent with this paired-end read SNP predictions predicted 13 SNPs, whereas single-end read SNP predictions predicted 5 (Figure 5.1).

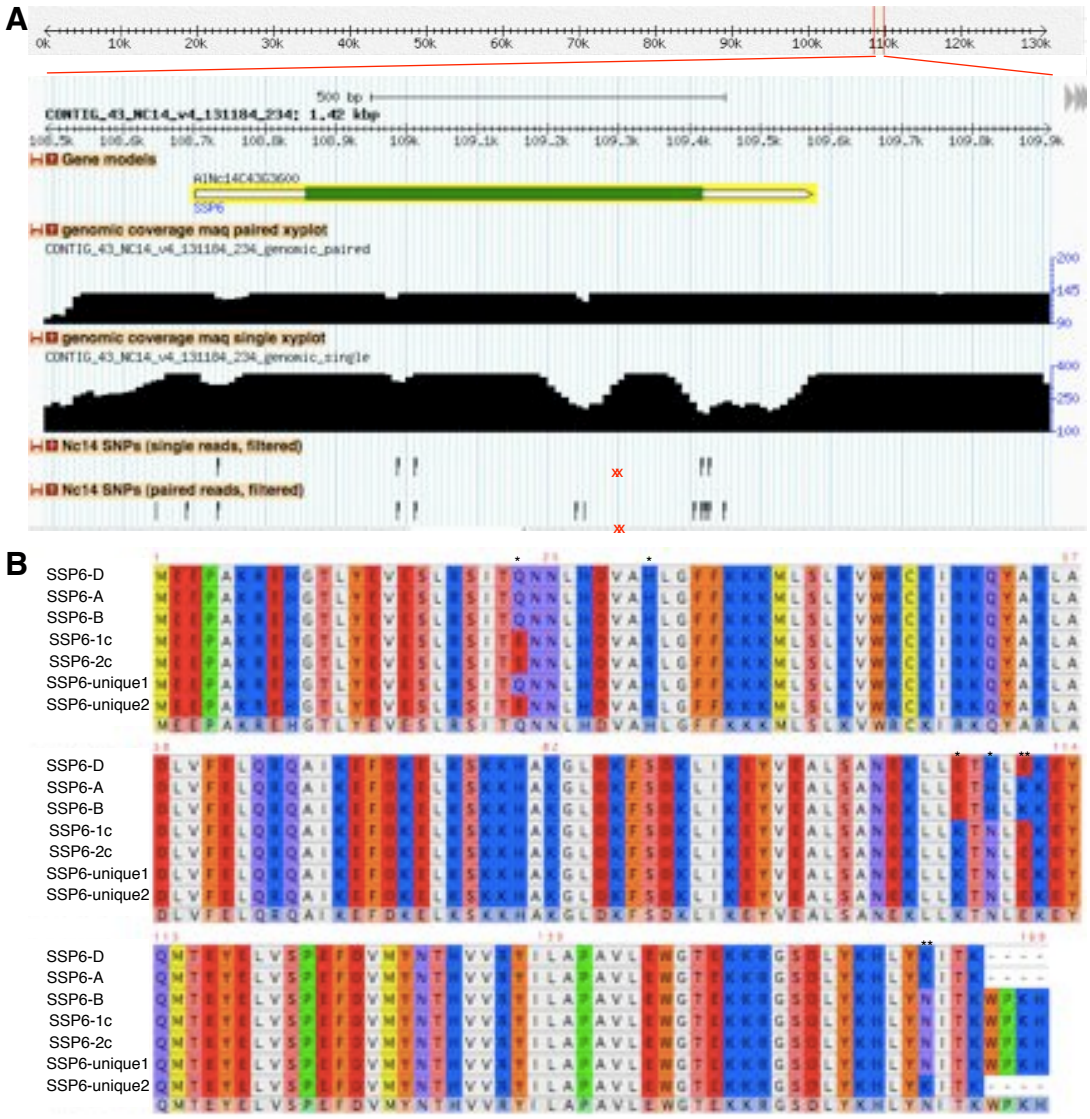



**Figure 5.1: The genomic region containing SSP6 is collapsed in the *AlNc14* genome assembly.** A) A GBrowser visualisation of the genomic region surrounding SSP6. Tracks show that the single-end read depth is close to the double of the global average (~240x), whereas paired-end read depth coverage is close to average (~160x). SNP prediction based on single- and paired-end read alignment using MAQ are shown. Red crosses indicate positions of unique SNPs determined by Sanger sequencing of ESTs. B) Protein alignment of the 7 variants of SSP6 determined by allele-mining. Variant B and D have only been obtained through cDNA sequencing, whereas A, 1c, and 2c have been found both by gDNA and cDNA sequencing. Likewise have unique 1 and 2 only been found once by gDNA sequencing. For clarity reasons, proteins were aligned from SP cleavage site to end. PAML test suggests that the copies are under positive selection (p = 0.005, Chisq). Asterix indicate residues under positive selection: * p<0.05, and ** p<0.01.

Effectors are often found in fast evolving repetitive regions and with paralogs linked in *cis*. For example, 3 paralogs of ATR5 are found within a 35 kb region (Bailey et al. 2011). These fast evolving regions are often not present in related species and can be evident as gaps in syntenic regions (Haas et al. 2009). We speculated that SSP6 could be located in a similar region, where regions proximal to the *SSP6* locus would be syntenic with other oomycetes i.e. *Albugo candida*.

We queried the recently published genome of the related *Albugo* species *Albugo candida* 2VRR (*Ac2VRR*) to test this hypothesis (Links et al. 2011). Specifically, we searched the *Ac2VRR* genome for regions with homology to a 15 kb region either side of the unresolved SSP6 locus. Interestingly, synteny with *Ac2VRR* begins 544 bp upstream of SSP6, but the region downstream of SSP6 was absent from the *Ac2VRR* genome assembly. This suggested that the SSP6 containing region was specific to *AlNc14*.

We performed a reciprocal BLAST with the full length *Ac2VRR* contig to determine if other *AlNc14* contigs would align to this. Interestingly, we found that two contigs were homologous to the *Ac2VRR* contig supporting a 979 bp gap in *Ac2VRR* (Figure 5.2). The high level of conservation in the regions flanking the synteny break points suggested that these two contigs could be linked on the same chromosome. To reconstruct the SSP6 locus we attempted to determine if these contigs were physically connected using PCR. To this end we generated high quality DNA from *AlNc14* and performed PCR reactions with two different primer sets to span the putative gap between the contigs. Interestingly we obtained amplification products and sequencing of these revealed that a 1072 bp sequence was split into multiple contigs in the assembly. Using this 1072 bp sequence we were able to merge the two contigs into a single contig.

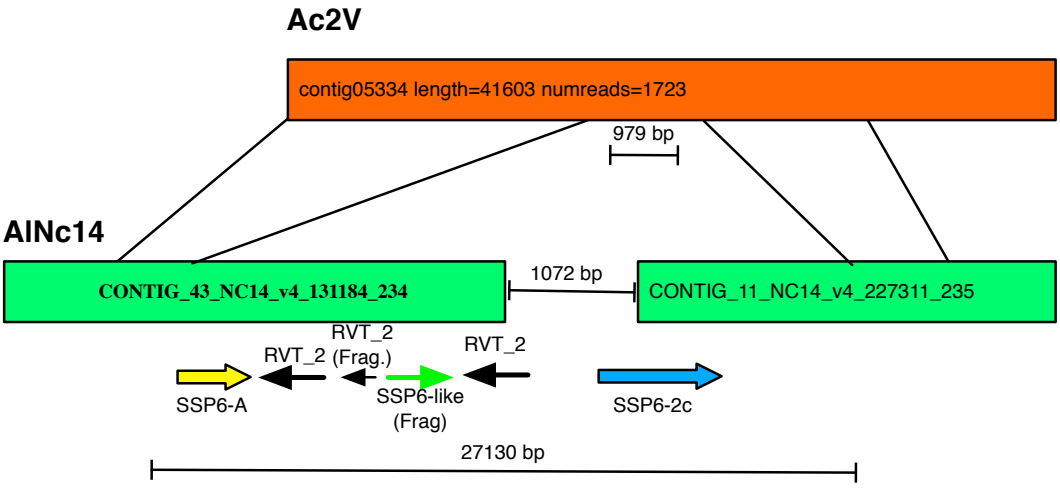



**Figure 5.2: The SSP6 locus is absent from Ac2V in a non-syntenic region.** The Contig_43 hosting SSP6 was queried against the Ac2V genome using BLAST. Reciprocal BLASTing yielded two hits: Contig_43 and Contig_11. The gap (1072bp) was closed by PCR analysis using primers specific to the ends of Contig_43 and Contig_11. In total 27130 bp were inserted in AlNc14, whereas 979 bp were specific to Ac2V in the non-syntenic region. Annotation of the region revealed the presence of transposons containing a RVT2 domain. Block arrows denote genes expresed in AlNc14, whereas arrows describe genes without cDNA evidence. (Frag) denotes pseudogenized genes. The figure is not to scale.

In total, 27130 bp were present in the *AlNc14* genome but absent from the *Ac2VRR* genome. By using blast homology searches and *de novo* gene prediction in this region we found that two SSP6 alleles were encoded at two different loci in this region; further, we identified a pseudogenised copy of SSP6. The two alleles we identified were the SSP6-A and the SSP6-2c alleles. Intriguingly three other ORFs were found in this region, which had homology to RVT transposases. One appeared to be pseudogenized due to the presence of multiple premature translational termination codons (Figure 5.2).

### 5.2.2 Variants of SSP6 show signatures of diversifying selection

We performed PCR allele mining on genomic DNA and cDNA to independently determine the various alleles/variants of SSP6 that existed in the *AlNc14* genome (Figure 5.1). Surprisingly, we found 7 variants of SSP6 in *AlNc14* out of 48 clones sequenced. We tried

using homology-based searches to determine the location of these alleles in the genome assembly but these attempts were unsuccessful.

Many oomycete effectors are highly polymorphic and under a strong positive selection pressure (Win et al. 2007; R. L. Allen et al. 2008). We performed a test for positive selection to determine if some residues in the variants of SSP6 were under a selection pressure using Phylogenetic Analysis by Maximum Likelihood (PAML) (Yang 1997). The analysis suggested that the variants of SSP6 were undergoing positive selection and specifically 6 residues were targeted, which were located in 3 discrete regions ($p<0.005$, $\chi^2$ test). The changes in these regions represent large changes in local physiochemical environment e.g. Q23E is a change from a positively charged side chain to a negative (Figure 5.1). Similarly, the two other regions have dramatic changes: the first regions have a $E_{107}TNLK_{111}$ to $K_{107}THLE_{111}$ change, whereas the second present in the C-terminus of the SSP6 variants is a change from $N_{162}ITKWPKH_{169}$ to $K_{162}ITK_{165}$.

As SSP6 is under positive selection and absent from *Ac2VRR* we reasoned that it could be a fast evolving effector, which are typically species specific. However, the alleles could have been gained though horizontal gene transfer, as virulence and host range of pathogens can be dramatically altered by a spontaneous uptake of foreign DNA elements by horizontal gene transfer. Richards et al (2011) analyzed plant pathogenic oomycetes and found that significant number of the secretome arose from genes of a fungal origin. Among the different classes of transferred genes were genes with host degrading capabilities, genes that are required to combat the plant, and effector genes required to suppress host defense. Thus, to determine if homologs of SSP6 are found in other species including fungi we performed homology based blast searches against the NCBI non-redundant protein database, but found no significant hits (E-value$<10^{-3}$). This suggests that copies of SSP6

likely are unique to *Albugo laibachii* and could represent a fast evolving set of genes consistent with their absence from *Ac2VRR*.

We analyzed the ESTs to determine, which alleles are predominantly expressed, and found that SSP6-A, SSP6-B, SSP6-1c, SSP6-2c, and SSP6-D were expressed, but SSP6-A and SSP6-2c had a markedly higher EST tag count compared to the others. As SSP6-A and SSP6-2c also represent the most allelic diversity, and for these reasons we chose to study them in more detail.

### 5.2.3 SSP6-A and SSP6-2c are plasma membrane localized

Subcellular localization *in planta* can reveal potential modes of function of an effector. To localize SSP6-A and SSP6-2c *in planta*, we fused GFP at the N-terminus of either ΔSP-SSP6-A or ΔSP-SSP6-2c under control of the 35S promoter. The constructs were transiently expressed in *N. benthamiana* and localized using confocal microscopy. Interestingly, both SSP6-A and SSP6-2c localized to the plasma membrane. Plasmolysis with 1M NaCl resulted in formation of Hechtian strands, thus confirming a strong association of the proteins to the plasma membrane (Figure 5.3). We queried the sequence for potential membrane localization signals or transmembrane domains but found none.

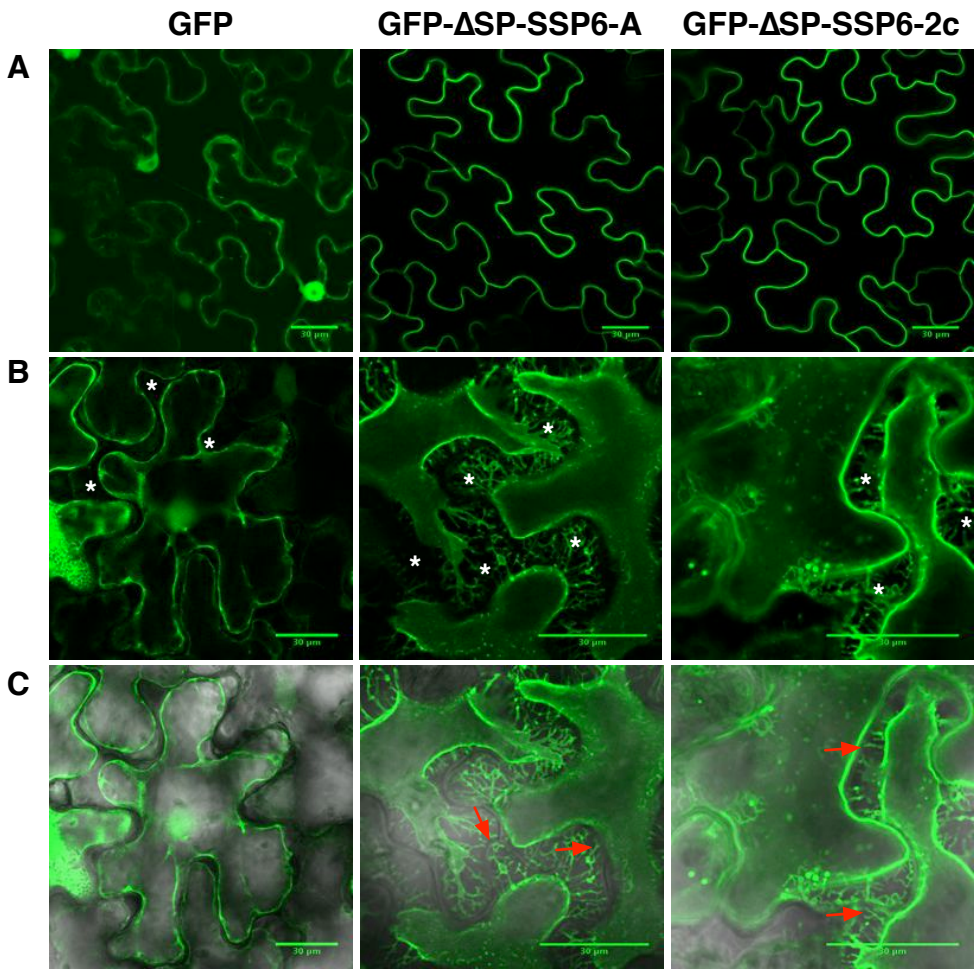



**Figure 5.3: SSP6-A and SSP6-2c localise to the plasma-membrane.** A) Transient expression in *N. benthamiana* of N-terminally GFP-tagged ΔSP-SSP6-A, ΔSP-SSP6-2c or EV (GFP). A) untreated control. B) Leaf discs were plasmolysed in 1M NaCl; white asterisks indicate area between cell wall and tonoplast. C) Plasmolysed GFP picture overlaid with DIC to visualise hechtian strand attachment sites (red arrows). Scale bars are 30 μm.

### 5.2.4 SSP6-2c, but not SSP6-A, suppresses host defenses

Most effectors are secreted in order to suppress plant defense (Dodds & Rathjen 2010). To determine if SSP6 could suppress host defense in *N. benthamiana* we transiently expressed SSP6-A and SSP6-2c, which were subsequently infected with *P. infestans* blue 13. Specifically, GFP was infiltrated into one half of the leaf and an effector into the other half. This revealed that *P. infestans* blue 13 lesions were not significantly larger on leaf-halves expressing GFP-ΔSP-SSP6-A compared to GFP (Figure 5.4). However, we found that compared to the GFP control, GFP-ΔSP-SSP6-2c expressing leaf-halves were

significantly more susceptible to *P. infestans* blue 13. This suggested that SSP6-2c but not SSP6-A was able to suppress *N. benthamiana* defenses against the adapted pathogen *P. infestans* blue 13.

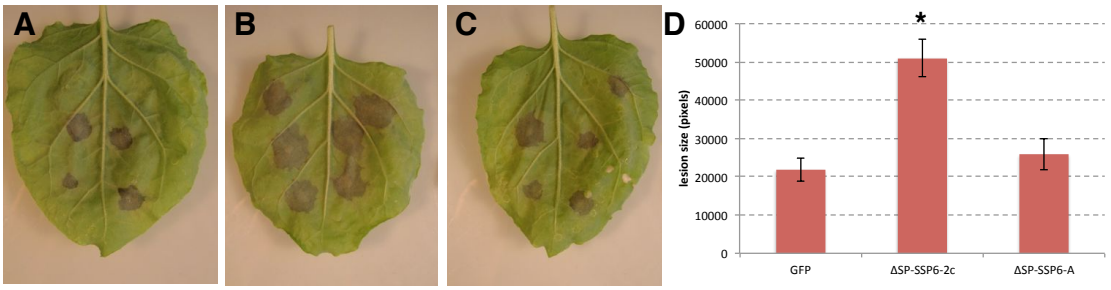



**Figure 5.4: Transient expression of SSP6-2c but not SSP6-A enhance virulence of *P. infestans* blue 13 in *N. benthamiana*.** Leaves expressing either A) Ev (GFP), B) GFP-ΔSP-SSP6-2c, or GFP-ΔSP-SSP6-A were superinfected with *P. infestans* blue 13. D) Leaves expressing GFP in one half of the leaf and either GFP-ΔSP-SSP6-A or GFP-ΔSP-SSP6-2c in the other half were infected with *P. infestans* blue 13. Lesion size was determined 7 dpi; Leaves expressing ΔSP-SSP6-A or ΔSP-SSP6-2c were normalised to GFP. Error bars denote SE, *: $p<0.01$. The experiment was repeated 3 times with a similar result.

PAMP receptors and their co-receptors are often located in, or closely associated with, the plasma membrane. Several bacterial effectors target key components of the PRR cascade to suppress PTI (Boller & Felix 2009). For example, AvrPto is myristoylated and is found in the plasma membrane where it disrupts FLS2-BAK1 interaction (L. Shan et al. 2000; L. Shan et al. 2008). As SSP6-2c localized to the plasma membrane, we hypothesized that SSP6-2c could be implicated in PTI modulation. A key output of PTI is generation of reactive oxygen species (Segonzac et al. 2011). Interestingly, transient expression in *N. benthamiana* of GFP-ΔSP-SSP6-2c but not GFP-ΔSP-SSP6-A resulted in markedly reduced generation of ROS compared to GFP control plants upon flg22 stimulation (Figure 5.5). This led us to further investigate if SSP6-2c could suppress ROS accumulation triggered by other PAMPs.

Chitin is a PAMP that is recognised by CERK1, which does not require BAK1 for signalling (Miya et al. 2007; Schwessinger et al. 2011). We tested if ΔSP-SSP6-2c or ΔSP-SSP6-A could interfere with chitin dependent signalling. Interestingly, we did not find an increase or decrease in ROS accumulation or kinetics upon chitin stimulation (Figure 5.5).

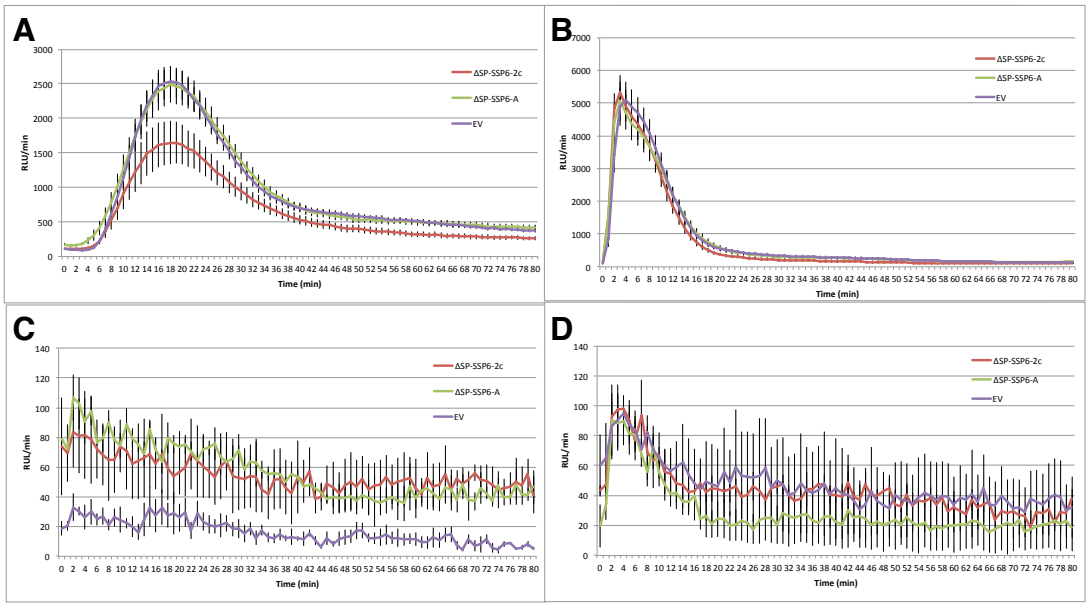



**Figure 5-5: SSP6-2c reduces flg22 but not chitin dependent ROS in *N. benthamiana*.** Transient Agro GV3101 pMP90 expression of GFP, GFP-ΔSP-SSP6-A or GFP-ΔSP-SSP6-2c followed by PAMP treatment 2 dpi with oxidative burst measurement. Leaf discs from 8 plants were treated with either A) 100nM flg22, B) 100μM Chitin or C,D) water. RLU = relative light units. Error bars denote SE. The experiment flg22 was repeated 4 times with similar result, and the chitin ROS was performed once.

### 5.2.5 SSP6-2c but not SSP6-A localizes to the *Albugo laibachii* haustorium

To determine if SSP6 was localized to the *Arabidopsis* plasma membrane we generated stable transgenic Col-0 plants harbouring either GFP-ΔSP-SSP6-A or GFP-ΔSP-SSP6-2c under a 35S promoter. Confocal microscopy of $T_2$ plants confirmed that both proteins were localized to the plasma membrane (Figure 5.6).

Pathogen infection is well known to cause dramatic reorganization of organelles and protein accumulation (Frey & Robatzek 2009). We challenged the transgenic $T_2$ plants with *AlNc14*, to determine if the distribution of SSP6-A or SSP6-2c would change upon infection. Confocal microscopy revealed that a substantial proportion of the protein was

still localized at the plasma membrane. Interestingly, we also observed that GFP-ΔSP-SSP6-2c localized around the *AlNc14* haustorium (Figure 5.6).

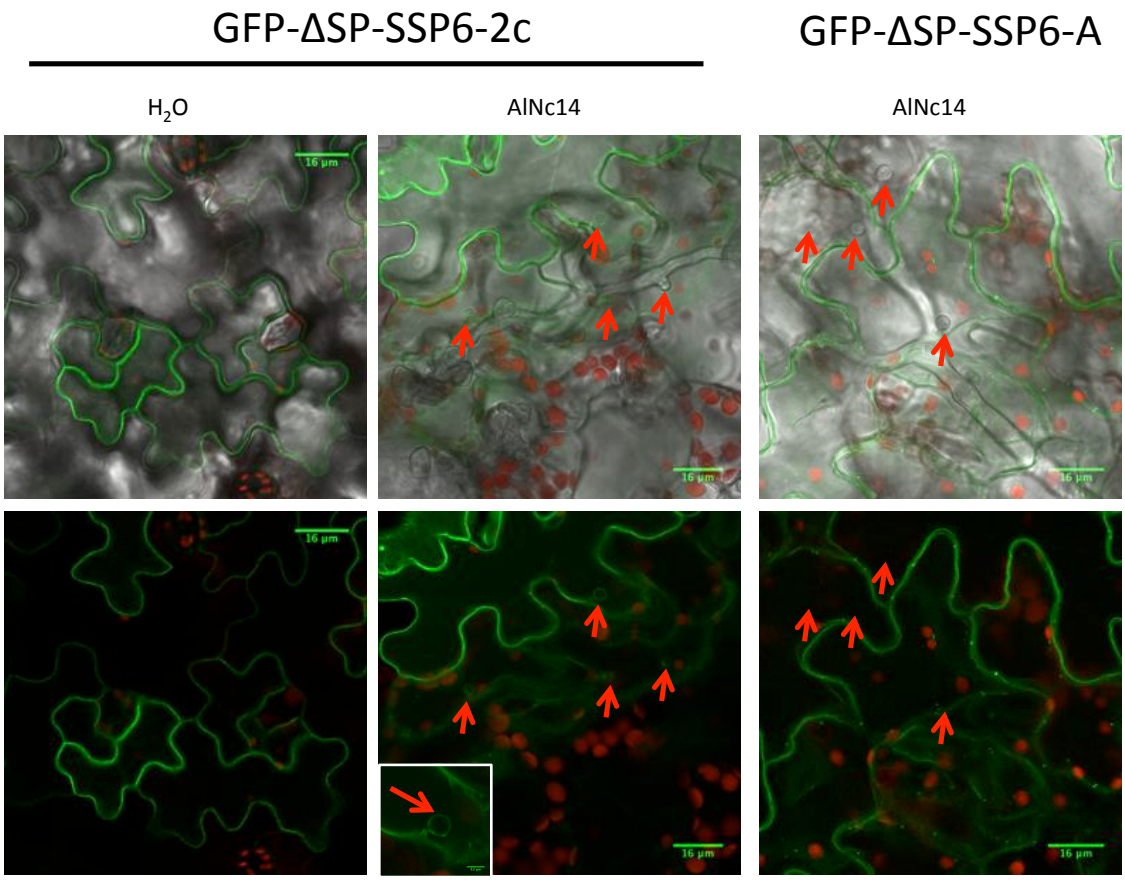



**Figure 5.6: GFP-ΔSP-SSP6-2c localizes to membranes and around *AlNc14* haustoria.** 3 week old *Arabidopsis* stable transformants expressing either GFP-ΔSP-SSP6-2c or GFP-ΔSP-SSP6-A were infected with *AlNc14* and analysed by confocal microscopy 10 dpi. Images are single confocal slices, scale bar is 16 μm. Red channel represents chlorophyll autofluorescence, green channel GFP, top panels are including DIC overlay, whereas bottom panels are without. The experiment was repeated twice with a similar result. Red arrow indicates haustoria.

## 5.3 Discussion

We found that variants of SSP6 were encoded as paralogs at a single locus in a repeat rich region of the *AlNc14* genome. The region flanking the SSP6 locus was syntenic in the closely related *Albugo* species *Ac2VRR*, but the region containing the SSP6 paralogs was absent. This suggests that the locus could be a fast evolving region of the *AlNc14* genome.

Two variants SSP6-2c and SSP6-A, were predominantly expressed and both localized to the plasma membrane. Further, SSP6-2c, but not SSP6-A, was able to enhance the virulence of *P. infestans* and suppress flg22 dependent ROS accumulation. Interestingly, we found SSP6-2c localized around *A. laibachii* haustoria in *Arabidopsis*.

### 5.3.1 SSP6 variants could be causal for host range

*Albugo candida* is able to infect a broad range of *Brassicaceae*, whereas *Albugo laibachii* is restricted to *Arabidopsis thaliana*(Thines et al. 2009). Schulze-Lefert and Panstruga (2011) hypothesized that two defense mechanisms govern host range and non-host resistance: PTI and ETI. As PRRs recognize PAMPs that are evolutionarily conserved PTI is the main force behind resistance to pathogens that do not share a co-evolutionary history with the host. In contrast, pathogens that have coevolved with their host(s) have effectors that are able to efficiently suppress host defenses, and hence resistance is governed by ETI. Thus, in their model ETI is the major contributor to resistance to co-evolved pathogens.

Since *A. laibachii* and *A. candida* are phylogenetically closely related and both are able to infect *Arabidopsis*, we hypothesize that recognition of an *A. laibachii* effector in *Brassicaceae* hosts triggers ETI and might determine non-host resistance.

As SSP6 variants are not found in the genome of *Ac2VRR* and other *A. candida* isolates, variants of SSP6 could be causal for restriction of *A. laibachii* host range to *Arabidopsis*. The hypothesis could be tested by HR assays in various *Brassicaceae* such as *B. oleracea* or *B. rapa* either via transient expression or EDV delivery of SSP6 alleles.

### 5.3.2 Expression of SSP6-2c enhances *P. infestans virulence* and suppresses ROS

We found that transient expression of SSP6-2c resulted in enhanced virulence of *P. infestans* blue 13 in *N. benthamiana.* Most effectors of plant pathogenic oomycetes also suppress PTI, as observed in *Hpa* where roughly half of the effectors tested suppressed PTI

(Stassen & Van den Ackerveken 2011; Fabro et al. 2011). Consistent with this *Hpa* Emoy2 infected leaves were found to elicit a weaker flg22 dependent ROS response in comparison to control plants (Fabro et al. 2011). Further, pretreatment with chitin or flg22 impairs growth of *Hpa.* Thus, suppression of PTI is central to oomycete fitness, although the PAMP complement of oomycetes is generally unknown.

Since SSP6-2c can suppress flg22 dependent ROS accumulation in *N. benthamiana*, we propose that this is the mechanism by which defense is suppressed and *P. infestans* virulence enhanced.

### 5.3.3 SSP6-2c may suppress PTI components around haustoria in *Arabidopsis*

Defense related proteins such as FLS2 and PEN1 (penetration 1) have been found to be closely associated with EHM around *Hpa* haustoria in *Arabidopsis* (Y.-J. Lu et al. 2012). Since FLS2 is localized around the haustorium of *Hpa* it is conceivable that other PTI components such as BAK1 could be localized around the haustorium as well to ensure triggering of PTI directly at the EHM.

In *Arabidopsis* stable transgenic lines we found that GFP-ΔSP-SSP6-A was exclusively localized at the plasma membrane upon *AlNc14* infection. In contrast, we found that SSP6-2c was localized at the plasma membrane but also around the haustorium in *AlNc14* infected plants expressing GFP-ΔSP-SSP6-2c. If we assume that components of the PTI machinery are found around the haustorium of *A. laibachii*, a hypothetical model for SSP6-2c is that it modulates a component of the PTI machinery specifically at the EHM.

We observed a strong membrane association of SSP6-2c and SSP6-A in both *N. benthamiana* and *Arabidopsis* but were unable to predict any putative membrane localization signals. For this reason we hypothesize that the variants of SSP6 are localized to the plasma membrane due to interactions with membrane bound entities. Since early events of PTI signaling occur at the plasma membrane it is most likely that these are

suppressed by SSP6-2c (Segonzac & Zipfel 2011). Presumably the membrane localization is required for the virulence function of SSP6-2c; PTI and pathology experiments on an *in planta* expressed mis-localized SSP6-2c could test this hypothesis.

### 5.3.4 Variants of SSP6 may reveal residues required for PTI suppression

Direct binding of AvrPtoB to BAK1 is required for virulence, the crystal structure of the BAK1 and AvrPtoB interaction revealed residues required for this interaction. Individual mutations of these residues resulted in the loss of interaction and virulence (W. Cheng et al. 2011).

Given that SSP6-2c can suppress flg22 dependent ROS responses but not SSP6-A, we speculate that this difference is due to a differential interaction with host proteins. The differences between SSP6-A and SSP6-2c are restricted to 7 amino acid changes and a 4 amino acid extension on SSP6-2c. These changes can be broadly divided into 3 different polymorphic regions. If we assume that each of the three regions represent a possible interaction block then there are 2 x 2 x 2 = 8 possible combinations of these blocks. Our allele-mining analysis revealed the presence of 7 SSP6 variants and therefore all but one possible combination is present in the *AlNc14* genome. Thus, the testing of these SSP6 variants will allow mapping of the region required for ROS suppression.

While, this will define the regions(s) or residues required for suppression, the biochemical mechanism by which SSP6-2c suppresses ROS would remain unknown. Further knowledge on the interactors of the different SSP6 alleles will provide important clues into the mechanism underpinning the defense suppression by SSP6-2c. To this end we are currently identifying interactors via yeast-two-hybrid screening (In collaboration with Jens Steinbrenner, Warwick UK). As SSP6 is localized to the plasma membrane, we are also performing co-immunoprecipitations followed by mass spectrometry as a complementary approach.

### 5.3.5 Conclusion and outlook

Our data point to a role for SSP6-2c in targeting a core defense regulator, which is sufficiently conserved between *Brassicaceae* and *Solanaceae* plants to enable SSP6-2c to modulate this in *N. benthamiana*. However, experiments in *Arabidopsis* are required to determine if SSP6-2c can modulate this conserved defense regulator in order to suppress FLS2 dependent ROS accumulation, and host defenses to pathogens. To this end we have generated lines of both estradiol inducible and stable 35S expressors of SSP6-A and SSP6-2c in Col-0. These will be tested with the pathogens *Hpa* and *Pst* DC3000 ΔAvrPto/ΔAvrPtoB, and also assayed for ROS production upon PAMP perception.

In conclusion, we performed a preliminary analysis of SSP6 variants, which represents a novel type of effector that is unrelated to any previously described effectors. They do not possess any known oomycete effector motifs such as RXLR, CRN or CHXC motifs suggesting uptake, but they appear to be under strong selection pressure. Their presumed function is to suppress defense via modulation of ROS signaling.

# 6. Mapping of *AlNc14* and *AcNc2* resistance loci

## 6.1 Introduction

The *Arabidopsis thaliana-Albugo* sp. pathosystem is of particular interest, as *Albugo* is a highly successful pathogen of *Arabidopsis* in nature. Based on comparison of ITS and cox2 sequences, two species of *Albugo* were identified that are parasitic on *Arabidopsis*: *Albugo laibachii* and *A. candida* (Thines et al. 2009). A survey of wild populations of *Arabidopsis* in the spring of 2006 and 2007 concluded that approximately 70% were infected with *Albugo laibachii,* formerly known as *Albugo candida* (Holub 2008). These observations correlate well with laboratory experiments, where approximately 90% of *Arabidopsis* accessions are susceptible to *Albugo laibachii* (Holub et al. 1995; Kemen et al. 2011). Initial efforts to establish the genetic foundation of the *Arabidopsis-Albugo* interaction system identified 3 resistance loci conferring resistance to *A. laibachii* Em1 (formerly *Acem1*, *A. candida* East Malling 1) (Holub et al. 1995; Borhan et al. 2001). These efforts have provided a basic genetic foundation for the pathosystem, but the precise identities of all avirulence factors in individual *Albugo* isolates and the resistances genes in *Arabidopsis* accession remains unclear. However, *Arabidopsis* Ksk-1 harbors two *Albugo laibachii* resistance loci: the dominant (*Resistance to Albugo candida 1*) *RAC1* and the semi-dominant locus *RAC3* (Borhan et al. 2001). *RAC1* is a TIR-NB-LRR, that recognizes *Alem1*, and is encoded at a single *R*-gene locus (Borhan et al. 2004).

Although the related species *Albugo candida Acem2* is able to infect *Arabidopsis,* only 10% of accessions tested has been found to be susceptible (Borhan et al. 2008). Despite a higher frequency of incompatibility on *Arabidopsis* only a single broad-spectrum

resistance gene, *WRR4* (*White Rust Resistance 4*), has been cloned (Borhan et al. 2008). Those currently known resistance genes that recognize *Albugo* sp. encode TIR-NB-LRR type of resistance genes, which require EDS1 but not PAD4 for defense signaling. In contrast, many TIR-NB-LRRs require both EDS1 and PAD4 for signaling (Aarts et al. 1998). It is conceivable that *Albugo* resistance genes function in a slightly different manner. For this reason we set out to define additional sources of resistance in *Arabidopsis.*

We used two isolates of two different species of Albugo that were recently isolated, *Albugo laibachii* Nc14 (*AlNc14*) and *Albugo candida* Nc2 (*AcNc2*) (Kemen et al. 2011). We tested the isolates on 5 different accessions and 2 segregating $F_2$ progeny of *Arabidopsis* to define the genetic basis of avirulence and resistance. We found that *AlNc14* is incompatible on Ksk-1 due to *RAC1* and *RAC3*, whereas *AcNc2* was incompatible on Col-5, Col-0, and Ksk-1 due to *WRR4*. Interestingly, we identified an additional resistance locus, *WRR5*, in Col-5.

## 6.2 Results

### 6.2.1 Resistance to *AlNc14* is determined by *RAC1* and *RAC3*.

We tested Col-0, Col-5, Ws-0, Ksk-1, and the defense-impaired mutant Col-*eds1* for resistance and susceptibility to *AlNc14*. On this limited set of ecotypes we found that inoculation with *AlNc14* resulted in formation of white blisters on green leaves on all ecotypes except for Ksk-1 (Table 6.1). Sporulation occurred at 7-10 dpi and we observed no obvious differences in susceptibility between susceptible ecotypes. Borhan *et al*. (2001) reported that Ksk-1 is resistant to the *Alem1* and the resistance is governed by two loci: *RAC1* and *RAC3*.

**Table 6.1: Resistance to *Albugo* sp. in accessions of *Arabidopsis*.** A) Five plants of each accession were infected with either AcNc2 or AlNc14. The infections wer scored 10 dpi and repeated at least twice. B) $F_2$ progeny were infected and scored 7 and 14 dpi. Plants scored as resistant had no visible phenotype, intermediate susceptible plants were characterised by strongly delayed sporulation of AlNc14. Susceptible seggregants were indistinguishable from susceptible control plants.

**A**

| Ecotype | Ksk-1 | Col-5 | Col-0 | Col-*eds1* | Ws-0 |
|---|---|---|---|---|---|
| *AcNc2* | R | R | R | S | S |
| *AlNc14* | R | S | S | S | S |

**B**

| Pathogen | Cross | Resistant | intermediate suceptible | suceptible | Ratio Observed | Ratio Predicted | $\chi^2$ | P |
|---|---|---|---|---|---|---|---|---|
| *AlNc14* | Col-0 x Ksk-1 | 45 | 12 | 3 | 45:12:3 | 12:3:1 | 0.108 | 0.947 |
| *AcNc2* | Ws-0 x Ksk-1 | 176 | n/a | 64 | 176:64 | 3:1 | 0.098 | 0.754 |

We generated a Col-0 x Ksk-1 cross and infected the $F_2$ progeny to test if these genes conferred resistance to *AlNc14*. Consistent with Borhan *et al.* (2001), we observed 3 types of interactions. Some segregants were completely resistant to *AlNc14*. The rest were either completely covered in pustules 7 dpi, thus fully susceptible, or had an intermediate pustule formation 7 dpi that slowly spread (Table 6.1 and Supplemental figure S6.1).

The delayed pustule formation was reminiscent of a heterozygous *RAC3* phenotype observed upon *Alem1* infection (Borhan et al. 2001). *RAC3* is located near the *RPP8/HRT* cluster on chromosome 5, 1.5 cM from the marker Cra-1 and before nga129. Another marker, ciw9 is located 700 kb from Cra-1 between nga129 and Cra-1. To determine presence or absence of *RAC1* we designed a linked SSLP marker based on a 400 bp insertion in *RAC1*. We used these markers to genotype the segregants, which had a delayed pustule formation. We found that all carried the Col-0 allele of *RAC1*. In contrast, none of the segregants were homozygous Col-0 at the ciw9 and nga129 loci: 5 were heterozygous, and 4 were homozygous Ksk-1. This suggested that delayed *AlNc14* pustule formation was

linked to *RAC3*, and potentially due to only one copy of the gene being present as observed for *Alem1* (Borhan et al. 2001).

Subsequently we tested the completely resistant segregants and found that they were either: homozygous for *RAC1* or *RAC3,* or heterozygous for *RAC1*, or some combination of these (Supplemental figure S6.1). Thus, we concluded that *AlNc14* avirulence on Ksk-1 was caused by the same resistance loci, *RAC1* and *RAC3,* as resistance to *Alem1* (Borhan et al. 2001).

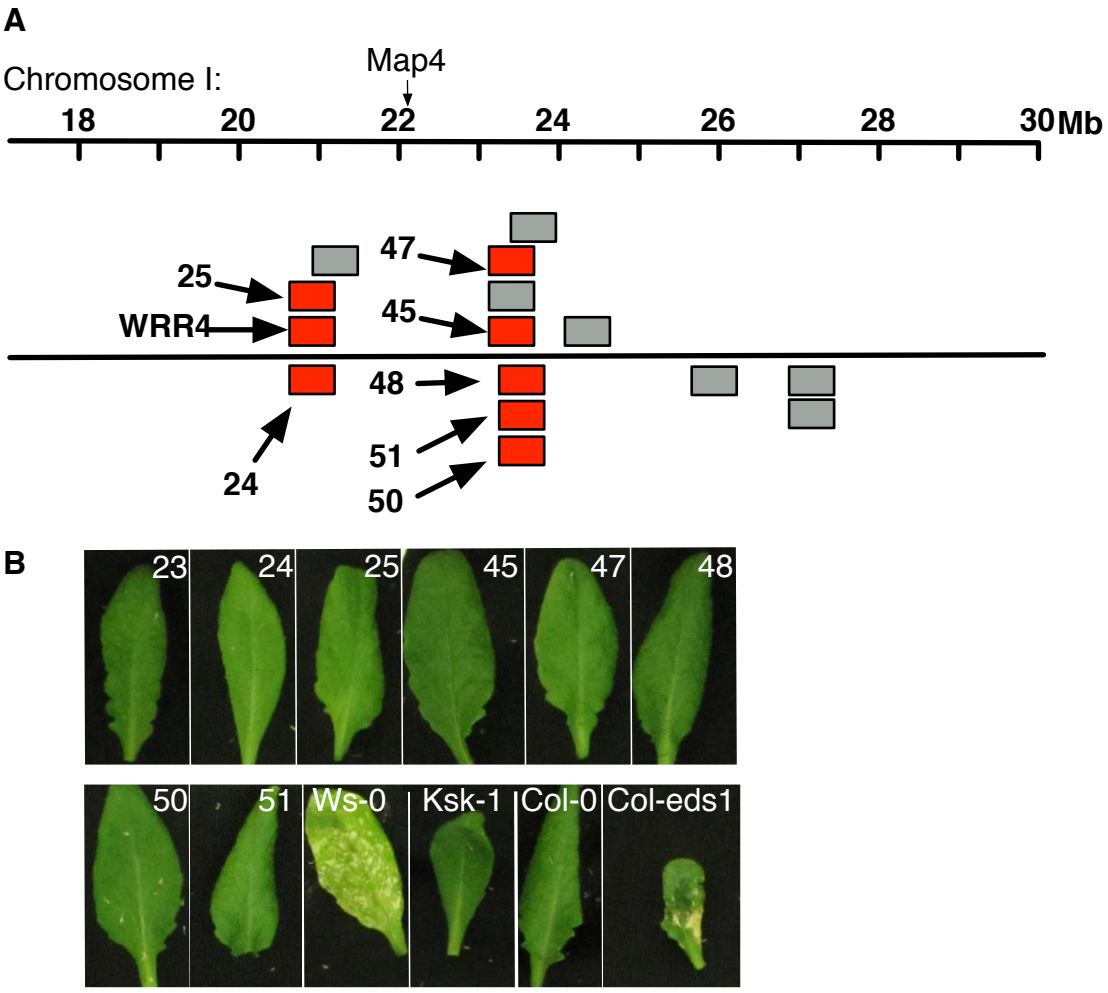



**Figure 6.1. Knockout of TIR-NB-LRRs located at the WRR4 locus does not confer sensitivity to AcNc2 in Col-0**. A) the location of the TIR-NB-LRR genes on chromosome I, relative to Map4. Knockouts tested with *AcNc2* are red and the untested grey. All ARTIC lines tested contained exonic insert T-DNAs in Col-0. Line 23 is *wrr4*. B) 8 plants of each ARTIC line was infected with AcNc2 and scored 14 dpi. The number in the top right corners denote ARTIC line number described in table S1. Sporulation of AcNc2 on Col-*eds1* was associated with strong necrosis.

### 6.2.2 Resistance to *AcNc2* is linked to *WRR4* in Ksk-1.

*Albugo candida* is a sister species of *A. laibachii* and some of the races can infect multiple *Brassica* species. *A. candida* Nc2 (*AcNc2*) ex. *Arabidopsis thaliana* can infect *Arabidopsis* and *Cardamine hirsuta* (Eric Kemen pers. comm.). The guard hypothesis proposes that R-genes have evolved to recognize modification of host proteins by invading pathogens. A single R-gene can thus recognize multiple effectors from various pathogens. These effectors may or may not be sequence related.

To test if resistance to *AcNc2* is due to *R*-genes active against *AlNc14,* we infected Col-5, Col-0, Ws-0, Col-*eds1*, and Ksk-1. We found that Ws-0 and Col-*eds1* were susceptible (Table 6.1). Interestingly, sporulation on Col-*eds1* plants was associated with extensive chlorosis surrounding the pustule (Figure 6.1). As Ksk-1 is resistant to both *AlNc14* and *AcNc2*, we crossed Ws-0 to Ksk-1 and infected the $F_2$ progeny with *AcNc2*. Resistance segregated 3:1 suggesting a single dominant R-gene locus conferred resistance (Table 6.1). We tested if resistance was linked to *RAC1* or *RAC3* but found no significant linkage (Table 6.2).

**Table 6.2. Resistance to *AcNc2* in Ksk-1 is linked to WRR4.** Susceptible $F_2$ plants from a Ws-0 x Ksk-1 cross were genotyped at loci previously associated with resistance to *Albugo* sp.

| Marker | Chrom. | linked R-gene | position | Ws-0 | Ksk-1 | Het | linkage* (cM) |
|---|---|---|---|---|---|---|---|
| map26 | **1** | RAC1 | 5.08 Mb | 15 | 5 | 15 | no linkage |
| map4 | **1** | WRR4 | 22.08 Mb | 48 | 0 | 5 | 4.95 |
| ciw8 | **5** | RAC3 | 7.5 Mb | 4 | 6 | 26 | no linkage |
| Ciw-9 | **5** | RAC3 | 17.04 Mb | 12 | 5 | 19 | no linkage |

* Linkage was established using Kosambi's function.

*WRR4* confers broad-spectrum resistance to other strains of *A. candida* and is located on the bottom of chromosome 1 in Col-0 (Borhan et al. 2008). We tested if resistance in Ksk-1 was linked to *WRR4* using the SSLP marker map4, which is located ~800 Kb from *WRR4*. Indeed, resistance was linked to this marker (Table 6.2).

### 6.2.3 Col-5 possesses two unlinked R-genes active against *AcNc2*

WRR4 is a TIR-NB-LRR that requires EDS1 for function (Borhan et al. 2008). As *WRR4* was identified in Col-0 and resistance to *AcNc2* in Col-0 was dependent on *EDS1* (Table 1), we speculated that this R-gene conferred resistance to *AcNc2*.

A collection of homozygous *R*-gene T-DNA knockouts known as the ARTIC lines were available in the lab (Lewis et al. 2010); we tested all TIR-NB-LRR knockouts available at and around the *WRR4* locus. Interestingly, none of the 8 tested *R*-gene mutants were found to be susceptible to *AcNc2*, or displayed any macroscopic phenotype such as chlorosis (Figure 6.1). This result could be explained by one of three possibilities: resistance is conferred by another *R*-gene at a different locus in Col-0, or multiple *R*-genes confer resistance. Similarly, resistance could be conferred by an untested *R*-gene at the locus for which no T-DNA knockout was available from the ARTIC collection.

To determine the cause of resistance in Col-5 we infected a Col-5 x Ws-0 $F_2$ population with *AcNc2*. Col-5 is a glabrous mutant of Col-0, which has an easily distinguishable phenotype that facilitates the identification of a successful cross and determination of linkage to *gl-1* (glabrous) located on the bottom of chromosome III in proximity of the *RPP1* locus. We did not observe any differences in the *Ac-Arabidopsis* interaction phenotypes: resistant plants were indistinguishable from Col-5, and pustules were associated with necrosis on susceptible segregants. The $F_2$ progeny segregated 15:1 for resistance to susceptibility, consistent with two dominant unlinked loci conferring resistance (Table 6.3).

To determine if the *WRR4* locus contributed to resistance in Col-5, we PCR-selected four $F_2$ segregants, which carried a Ws-0 genotype at the *WRR4* locus but were resistant to *AcNc2*, and bulked up seeds from these plants. Col-5 has a glabrous phenotype, caused by

a recessive mutation in a myb transcription factor GL1 (Oppenheimer et al. 1991). As GL1 is linked to known oomycete resistance genes such as *RPP14* (Reignault et al. 1996), we bulked up 2 $F_2$ segregants with glabrous phenotype and 2 without.

**Tabel 6.3: Resistance in Col-5 to AcNc2 is governed by two unlinked dominant loci: the *WRR4* locus and another.** 4 week old *Arabidopsis* $F_2$ plants from a Col-5 x Ws-0 cross were infected with AcNc2 and disease scored 14 dpi. Four resistant segregants, which had a Ws-0 genotype at the *WRR4* locus were selfed. The progeny ($F_3$) was tested for segregating resistance to AcNc2. nd: not determined.

| | | Observed | | Expected | | | |
|---|---|---|---|---|---|---|---|
| | | | | 1 locus (3:1) | | 2 loci (15:1) | |
| **Progeny** | **gl phenotype** | **Resistant** | **suceptible** | **$\chi^2$** | **P** | **$\chi^2$** | **P** |
| $F_2$ | segregate 1:3 | 91 | 5 | 13.16 | <0.001 | 0 | 1.000 |
| $F_3$ | segregate 1:3 | 20 | 0 | nd | nd | nd | nd |
| $F_3$ | wild type | 16 | 4 | 0.067 | 0.796 | nd | nd |
| $F_3$ | glaborous | 20 | 0 | nd | nd | nd | nd |
| $F_3$ | glaborous | 16 | 4 | 0.067 | 0.796 | nd | nd |

From these $F_2$ plants we infected the $F_3$ progeny with *AcNc2* and found that progeny of 2 $F_2$ parents did not segregate for resistance. In contrast, the two other $F_3$ progeny segregated 3:1 for resistance (Table 6.3). This suggests that there are two main resistance loci in the Col-5 genome providing resistance to *AcNc2*; one is closely linked to the *WRR4* and the other one is an unknown dominant unlinked *R*-gene locus. We call this unknown *R*-gene locus *WRR5*. It is unlikely that *WRR5* is located at the *RPP14* cluster, as progeny segregating for resistance and susceptibility originated from both glabrous and non-glabrous $F_2$ segregants (Reignault et al. 1996).

## 6.3 Discussion

In this study, we showed that Ksk-1 is resistant to *AlNc14* and *AcNc2* but for different reasons. While the *A. laibachii* isolate *AlNc14* was recognized by *RAC1* and *RAC3* similar to *Alem1*, the *Albugo candida* isolate *AcNc2* was recognized by a *R*-gene linked to the

map4 marker at the *WRR4* locus. The *R*-gene is most likely an allele of *WRR4* in Ksk-1. Resistance to *AcNc2* in Col-0 was conferred by two loci: *WRR4* and *WRR5*.

The differences in resistance could be determined by presence of a proteinaceous effector *AvrWRR4* in *A. candida*, which would be absent or sufficiently diverged to avoid recognition in *A. laibachii* strains. Alternatively, *AvrWRR4* is an essential effector conserved in all *Albugo* sp. that *A. laibachii* has evolved other effectors to suppress recognition by *WRR4*.

The latter hypothesis is supported by the ability of *AlNc14* pre-infected Col-0 plants to permit subsequent growth of *AcNc2* (Eric Kemen pers. comm.). Thus, *AlNc14* is able to suppress *WRR4* and *WRR5* mediated resistance. However, suppression of *Arabidopsis* defense by *A. laibachii* is widespread. Pre-infection with *Alem1* suppresses disease resistance to several downy mildews and powdery mildews (Cooper et al. 2008). Identifying the effector(s) that cause this broad-spectrum defense suppression and their targets will potentially shed new light on plant defense pathways.

Our current understanding of the *Albugo* sp. effector repertoire has been defined by heterologous expression. The effector complement of *A. laibachii* has been defined by the ability of a protein to enhance the virulence of *Pst* DC3000 (Kemen et al. 2011) and the effector complement of *A. candida* has also been examined in *N. benthamiana* transient necrosis assays (Links et al. 2011). While these approaches defined classes of effectors that cause a modulation of virulence, classical definition of effectors by defining *R*-genes and the cognate *Avr*-gene is missing. Recognition of *RAC1* and *RAC3* will be instrumental in defining specific effectors in *Albugo laibachii*.

Both *AlNc14* and *Alem1* are recognized by *RAC1* and *RAC3* in Ksk-1. In *Hpa,* recognition of the isolates Emoy2, Hiks1, and Waco5 by *RPP1-Nd* is correlated with *ATR1,* which is identical in sequence in these isolates. In contrast, the *Hpa* isolates Maks9, Noks1, Cala2,

Emco5, and Emwa1 are not recognized by *RPP1*-Nd because they carry highly divergent *ATR1* alleles (Rehmany et al. 2005). Thus, it is conceivable that the *A. laibachii* effectors *AvrRac1* and *AvrRac3* could be similar and/or identical in *Alem1* and *AlNc14*, but absent or highly divergent in *AcNc2* and *Acem2*, as they are not recognized by *RAC1* or *RAC3*. We anticipate that the genetics presented will aid in identifying *AvrRAC1* and *AvrRAC3* from *A. laibachii*.

*A. laibachii* is able to evade recognition in most *Arabidopsis* accessions (Kemen et al. 2011); cases of *A. candida* resistance are more frequent (Anastasia Gardiner *in prep.*). *WRR4* confers resistance to four isolates of *Albugo candida* and is present in both *Col-5* and *Nd-1*. This led Borhan *et al.* (2008) to propose that *WRR4* confers broad-spectrum resistance in *Arabidopsis* to *A. candida* isolates. Our finding strengthens this hypothesis because *WRR4* also confers resistance to *AcNc2* in both Col-5 and Ksk-1. However, this prompts the question: what molecular component in *Albugo candida* confers *WRR4* dependent avirulence?

*Solanum bulbocastanum* is highly resistant to most races of *P. infestans* (Helgeson et al. 1998). One component of this broad resistance is caused by recognition of effectors *Avrblb1* and *Avrblb2* by *Rpi-blb1* and *Rpi-blb2* respectively (Junqi Song et al. 2003; Oh et al. 2009). The function of *Avrblb2* is to prevent secretion of host proteases at the haustorial interface and thereby create a favorable environment for the oomycete (Bozkurt et al. 2011). This is similar to the bacterial effector HopAS1 that enhances virulence of *Pto* T1 on tomato, but triggers avirulence on *Arabidopsis* Col-0 plants. This relationship was reflected in the natural *P. syringae-Arabidopsis* host range. In 27 strains of *P. syringae,* those that are nonpathogenic on *Arabidopsis* carry a full-length copy of *HopAS1,* whereas the gene was truncated in strains virulent on *Arabidopsis* (K. H. Sohn et al. 2011).

Therefore, *WRR4* mediated resistance may be triggered by a single effector in all 5 *Albugo candida* races tested. The identity of the effector is unknown, but presumably it enhances fitness on other *Brassica* species, which could be shared between all the 5 *A. candida* isolates. The availability of Illumina sequencing technology has dramatically reduced sequencing cost, thus it is conceivable that the shared effector could be identified within reasonable time by comparative association genomics (de Jonge et al. 2012).

*WRR4* and *WRR5* confer resistance to *AcNc2* in Col-5. As Col-*eds1* plants supported formation of *AcNc2* pustules, associated with extensive necrosis, this suggests that *WRR5* resistance also requires *EDS1*. Thus, WRR5 is presumably a TIR-NB-LRR. Interestingly, *WRR5* recognized *AcNc2* but not *Acem2*, which could be due to recognition of an *AcNc2* specific effector (Borhan et al. 2008). Alternatively, as Borhan et al. 2008 used seedlings but we infected 4-week-old *Arabidopsis* plants, *WRR5* could be under epigenetic regulation that results in up regulation of *WRR5* post cotyledon stage. The mechanism could be comparable to LAZ2/SDG8 dependent induction of LAZ5 transcription in plants resulting in *ACD11* dependent PCD (Palma et al. 2010).

*B. rapa* and *B. juncea* are susceptible to the *Albugo candida* races *Ac7* and *Ac2v* respectively. Interestingly, *Arabidopsis WRR4* has successfully been deployed in both species and resulted in resistance to the *Albugo* sp. (Borhan et al. 2010).

In our hands, *WRR4* and *WRR5* dependent resistance appear equally strong, as segregants carrying either allele were indistinguishable from Col-5 upon infection. Therefore, if *WRR5* also confers resistance to other *Albugo* candida, it would be exciting to test if *WRR5* could be used in a similar manner. To determine the identity of *WRR5* a selection of 200 susceptible $F_2$ progeny from a Col-5 x Ws-0 cross is currently in progress. Bulk DNA from these will be generated and the causal loci identified by absence of Col-5 genotype using a custom modified next generation mapping technique (Austin et al. 2011). Our ability to

correctly identify the *WRR4* locus as a Ws-0 genotype only region will serve as internal control for the method.

In conclusion, the genetics presented provide a solid foundation for future work on cloning effectors from *AlNc14*, *AcNc2*, and also identification of the second resistance locus in Col-0: *WRR5*.

NOTE ADDED IN PROOF: *Volkan Cevik joined the lab, and has further mapped the WRR5 locus to the bottom of Chromosome 5; interestingly the locus is comprised of two resistance genes WRR5a and WRR5b, which are both required for resistance.*

# 7 Development of a high throughput screening system for assessing *Hpa* virulence on *Arabidopsis*

## 7.1 Introduction

The *Arabidopsis-Hyaloperonospora arabidopsidis* (*Hpa*) interaction is a model-system for studying plant-oomycete interactions (Slusarenko & Schlaich 2003).

The system has been instrumental in elucidating molecular details of susceptible and resistant interactions. The success of the system can in part be attributed to the availability of vast genetic resources for *Arabidopsis* and the genetic tractability of *Hpa. Hpa* does not cause disease on economically important crops, but our knowledge of *Hpa-Arabidopsis* interactions has accelerated research in economically important, yet genetically less tractable, systems. One of the current bottlenecks in the field is the lack of high throughput screening methods to accurately quantify the virulence of an *Hpa* infection.

The lifecycle can be divided into discrete phases, which begin with infection of *Arabidopsis* by an oospore (sexual progeny) germinating in the soil or a conidiospore (asexual progeny) landing on the leaf (Slusarenko & Schlaich 2003). A successful infection results in growth of intercellular hyphae throughout the plant. Spherical feeding structures, haustoria, protrude from the hyphae into plant cells. Approximately 5-7 days post infection sexual oospores form by fertilization of the female oogonium by male antheridium. However the first visible symptoms of infection appear after 3-4 days when conidiophores carrying conidiospores emerge through open stomata. Conidiospores are asexual progeny that are dispersed by a slingshot-twist-mechanism upon drying and loss of turgor in the conidiophore (Slusarenko & Schlaich 2003). Some conidiospores may land on a leaf where they germinate and begin a round of asexual infection. In contrast, oospores are lodged within the plant leaf and released into the soil when the leaf rots.

In the wild resistance and susceptibility to *Hpa* differ between accessions; genetic dissection of 11 *Arabidopsis* accessions identified 13 loci, *RPP1-RPP12* and *RPP39,* governing resistance to 7 *Hpa* isolates (Holub et al. 1994; Holub 2006). Effort in several labs led to cloning of multiple R-genes such as *RPP1*, *RPP4*, *RPP5*, and *RPP13* (Parker et al. 1997; Bittner-Eddy et al. 2000; Botella et al. 1998). Sexual crosses between different *Hpa* races and association mapping led to the identification of avr-determinants such as *ATR1*, *ATR5*, *ATR13* and ATR39-1 (Rehmany et al. 2005; Bailey et al. 2011; R. Allen et al. 2004; Goritschnig et al. 2012). A common denominator for the above-mentioned screens is the requirement for distinct susceptibility or resistance phenotypes.

Another type of screen was recently pioneered by Fabro and colleagues; they took advantage of the newly sequenced *Hpa* genome to conduct a screen using bacterial delivery of 63 effectors (one at a time) into 12 ecotypes (Fabro et al. 2011; Baxter et al. 2010) via type III secretion from *Pseudomonas syringae*. Interestingly, this work showed that avr activity of effectors is the exception rather than the rule. Normally individual effectors enhance virulence on a given accession, but the extent may vary. This is also consistent with *Hpa*-Arabidopsis interactions: *Hpa* Waco9 on Ws-0 is highly infectious, whereas it is less virulent on Col-0 (Holub et al. 1994). Likewise, *Hpa* Noco2 is highly virulent on Col-0 but not Tsu-1. The loci that contribute to differences in virulence are unknown; investigation has been hindered by the lack of a quantitative high-throughput *Hpa* quantitation systems.

Current methods for quantifying *Hpa* infections report: the number of conidiophores per leaf or plant, or the number of conidiospores per g (fw) or plant (Jambunathan et al. 2001; Fabro et al. 2011; H. S. Kim 2002). These methods rely on counting translucent objects with low contrast that could lead to errors. Further, counting low contrast objects is inherently difficult, and requires much time to assure adequate certainty.

Trypan blue is routinely used to stain conidiophores and enhance contrast, thus reducing quantification errors. However, the dye is non-specific and also stains dead plant cells, intercellular hyphae, oospores and veins (Hofius et al. 2009). Further, the procedure is time consuming and highly toxic; trypan blue has been labeled as a carcinogen (IARC 1999). For these reasons, different techniques based on other dyes that are more specific and enhance contrast are preferable.

Sugar binding lectins have been used extensively to characterize oomycete cells (Carzaniga et al. 2001; Wawra, Bain, et al. 2012b); interestingly, the carbohydrate composition of the *Hpa* conidia is markedly different from the germtube, appresoria and ECM of *Hpa* (Carzaniga et al. 2001). For example, Concanavalin A (Con-A) is a α-manose binding lectin that strongly labels ECM, germ tube and appresoria, and to a lesser extent conidia. Similarly, wheatgerm agglutinin (WGA) and *Griffonia simplicifolia* agglutinin-II bind N-acetylglucosamine groups, which are structural components found in chitin. Other lectins only bind conidia such as *Bauhinia purpurea agglutinin.* Although the lectins bind specifically to various parts of the oomycete, they are expensive and are unstable.

Chitin is a major constituent of fungal cell walls, whereas the oomycete cell wall consists primarily of cellulose and glucans (Bartnicki-Garcia 1968). Uvitex 2B is a fluorescent brightener that stains fungi and algae through presumed binding to chitin and cellulose (Koch & Pimsler 1987; Wachsmuth 1988). It has been extensively used to stain plant-pathogenic fungi (Chen et al. 2010; Diagne et al. 2011). The main driving forces behind the widespread use of uvitex 2B are the dye is non-toxic, cheap, and highly photo-stable (Wachsmuth 1988). Given the versatility of uvitex 2B we tested the stain on *Hpa*-infected *Arabidopsis* leaves.

Here we report that uvitex 2B readily stains aerial *Hpa* structures and devise a simple protocol. Our staining protocol shows that uvitex 2B is more suitable, for staining conidiophores and conidiospores, than trypan blue. We demonstrate that automated detection of uvitex 2B stained conidiophores relative to the leaf size is correlated with the number of condiophores per $cm^2$. We call this the uvitex 2B method. Importantly, this method distances the experimenter from data acquisition and dramatically decreases sampling time. We independently verify the uvitex 2B method by confirming that 300 µM BTH significantly reduces *Hpa* virulence. Although the method is developed for the *Hyaloperonospora-Arabidopsis* patho-system, preliminary experiments show that the uvitex 2B method could be extended to other plant-pathogen systems.

## 7.2 Results

### 7.2.1 ConA and uvitex 2B stain *Hpa* conidiophores and condiospores

To establish a high-throughput semi automated quantification system for assaying *Hpa* infection, we tested if the two sugar binding dyes, ConA, and uvitex 2B, could label *Hpa* conidiophores.

Consistent with previous reports, we found that ConA-FITC bound specifically spores and also conidiophores (Carzaniga et al. 2001). In addition, we noticed a strong auto fluorescent signal from trichomes within the detection range of FITC, which would impede accurate quantification with automated scripts (Figure 7.1).

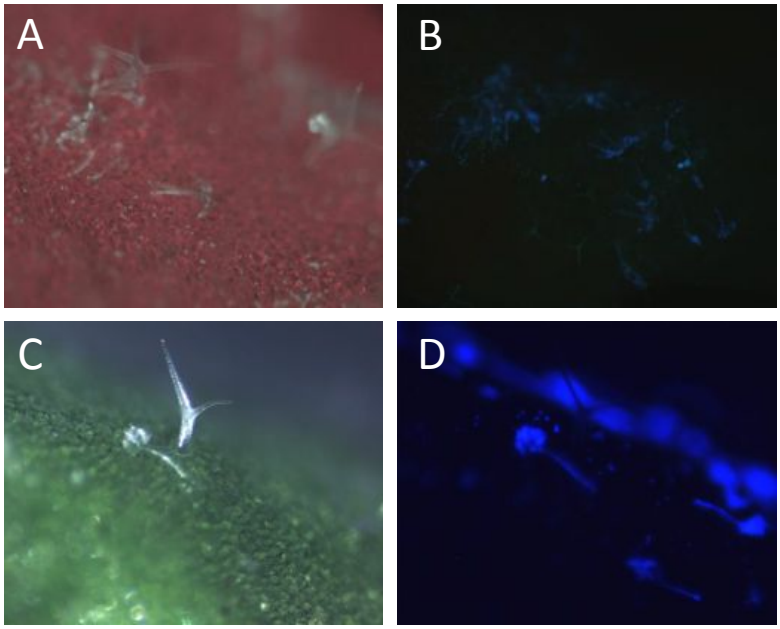



**Figure 7.1: ConA and uvitex 2B stain conidiophores.**
4 week old plants were infected with 10e4 spores/ml *Hpa* Noco2. 6dpi leaves were stained with A) ConA-FITC or B-D) uvitex 2B. Panels A,B and D are epiflourescence pictures, whereas C) is bright field image.

Incubation with uvitex 2B produced a strong specific signal from conidiophores and to a lesser extent spores without any significant background signal when visualized with a UV long pass filter (Figure 7.1). Importantly, as we image uvitex 2B-stained aerial *Hpa* structures on the leaf surface we have a dramatically increased focal depth compared to transmitted bright field visualization (Figure 7.1); this feature is beneficial for automated counting. For these reasons, we chose to establish a growth quantitation method based on uvitex 2B staining of *Hpa*.

### 7.2.2 The uvitex 2B method accurately quantifies *Hpa* infection

Our method can be divided into discrete steps: (i) a standardized infection, (ii) data acquisition using uvitex 2B, and (iii) statistical analyses (Figure 7.2). Three-week old plants are infected with $10^4$ spores/ml, and incubated at high humidity for 6 days. 3 leaves from the same rosette are sampled and subjected to the uvitex 2B method.

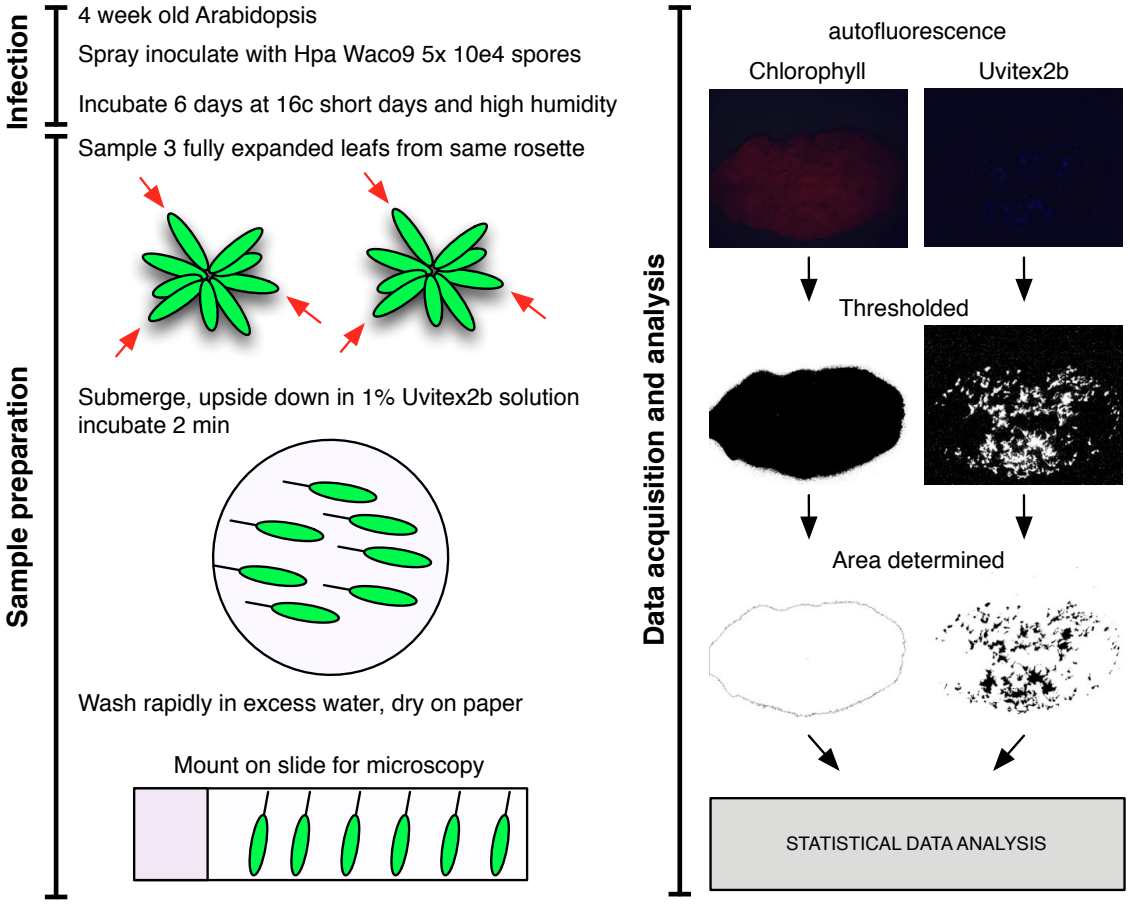



**Figure 7.2: Summary of the uvitex 2B protocol for staining *Hpa* conidiophores.** Consult main text for a detailed description.

The output of the uvitex 2B method is a measure of the percentage of the entire leaf that is covered by *Hpa* conidiophores. In brief, two fluorescent images are acquired: one of uvitex 2B stained *Hpa* structures, and another of chlorophyll autofluorescence. The pictures are converted to binary images, where all fluorescence above background is considered as a specific signal. Subsequently, the size of fluorescent clusters is determined using ImageJ (Schneider et al. 2012). Fluorescent clusters smaller than 20 pixels in size are considered background noise. The infection is determined as a ratio of the uvitex 2B stained area relative to the chlorophyll autofluorescence area.

To test if the uvitex 2B method could accurately determine and quantify the *Hpa* infection, we infected various mutants with known outcomes. We infected the following mutants and natural ecotypes to test for increases and decreases in *Hpa* virulence. In the Col-0

background, we chose the slightly more resistant mutant Col-*arf9-1* (auxin response factor 9-1), and as a more susceptible mutant we used Col-0 expressing 35S:AFB1 (auxin signaling F-box protein 1) (Robert-Seilaniantz, Maclean, et al. 2011b). In addition, we included Col-*ein5-1* (ethylen insensitive 5-1) that is insensitive to ethylene (Olmedo et al. 2006) and should elicit a stronger SA dependent response. Thus, we would predict that Col-*ein5-1* should be more resistant to *Hpa*. To cover the entire range of *Hpa* possible outcomes, we tested the resistant ecotype Ler-0, and the highly susceptible mutant Ws-*eds1*.

We found that compared to Col-0: *arf9-1* and *ein5-1* supported less growth of *Hpa* Waco9 revealed as a smaller area of uvitex 2B staining per leaf area. The incompatible reaction of Ler-0 to *Hpa* Waco9 exemplified the specificity of uvitex 2B staining: we observed no conidiophores and observed no uvitex 2B staining on Ler-0 leaves (Figure 7.3). Further, both methods revealed that Col-0 plants expressing *35S:AFB1* are more susceptible than wild type, and that Ws-*eds1* and plants are the most susceptible. This is consistent with previous reports of infection with *Hpa* Noco2 (Robert-Seilaniantz, Maclean, et al. 2011b; Parker et al. 1996) and suggests that gain and loss of resistance in these mutants is independent of the *Hpa* strain.

To verify the uvitex 2B method was able to accurately determine the level of infection consistent with previous publications, we independently counted the number of conidiophores/cm$^2$. We used the same leaves that previously had been subjected to the uvitex 2B method.

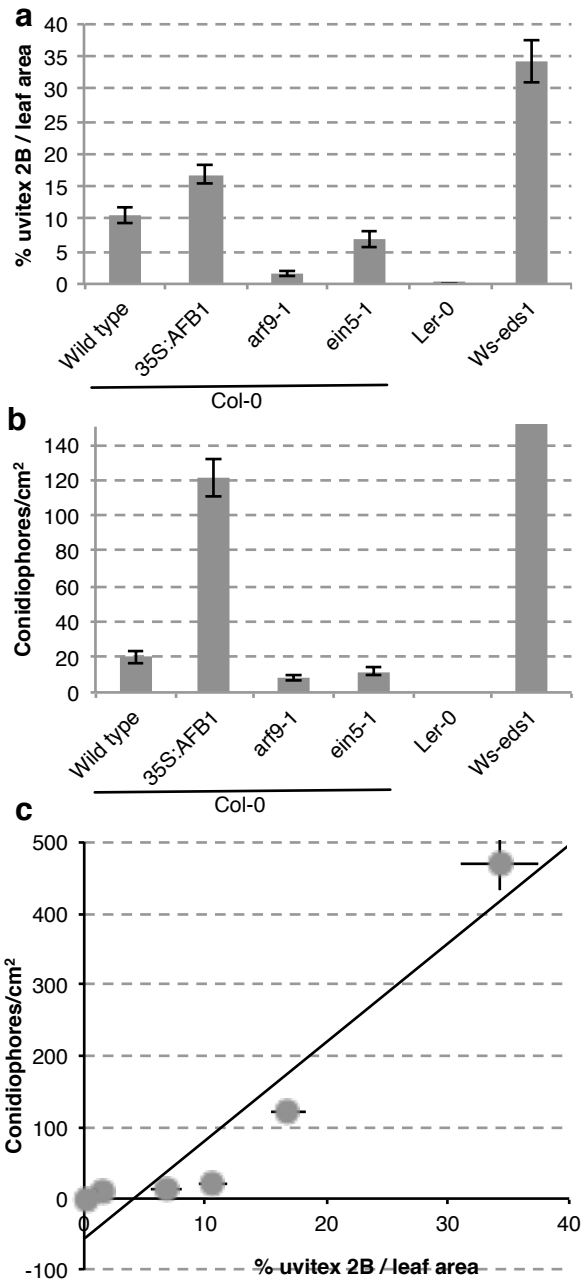



**Figure 7.3: The uvitex 2B method correctly determine the amount of asexual *Hpa* propagules.** 3 week old plants were infected with 5x10e4 spores/ml and the infection scored at 6 dpi. First by the uvitex 2B method A), and subsequently the leaves were stained with trypan blue and B) the number of conidiophores/cm2 determined. C) shows the correlation of the uvitex 2B method with trypan blue staining. The experiments were repeated 3 times with similar results. Error bars denote 2xSE.

As observed for the uvitex 2B staining, we found that Ws-*eds1* plants were highly susceptible and had an average of 344±33 conidiophores/cm$^2$. We found no conidiophores on Ler-0 plants. The Col-0 mutant *35S:AFB1* showed increased growth of *Hpa* Waco9, whereas *arf9-1* and *ein5-1* supported less growth. Further, the uvitex 2B method revealed that Col-*ein5*-1 was more resistant, which we verified by counting trypan blue stained conidiophores.

To test if the methods were comparable, we plotted the average uvitex 2B staining/leaf area versus the number of conidiophores/cm$^2$ for each mutant. We performed a linear correlation analysis to establish if the two measures could be considered equal. The analysis showed that they correlated well ($r^2$ = 0.905, Figure 7.3). The experiment therefore provides a convincing example that the uvitex 2B method does not compromise accuracy of conidiophore counting.

### 7.2.3 *Hpa* infection plateaus at different time points

During cultivation of *Hpa* Waco9 in the lab, we noticed that conidiophores were observable in large quantities by eye 1 to 2 days earlier on Ws-0 plants than on Col-0 plants. Therefore *Hpa* Waco9 could sporulate earlier on Ws-0 plants, or, at an increased frequency. To test this we performed a time course experiment.

We found that the amount of *Hpa* Waco9 conidiophores increased dramatically between day 5 and 6 on both Col-0 and Ws-0 plants (Figure 7.4). However, the rates of conidiophore production were different on Col-0 and Ws-0 plants. The apparent success of *Hpa* Waco9 on Ws-0 plants could be due to loss of fitness on Col-0, perhaps due to aborted conidiophore production. To our knowledge this is the first growth curve of *Hpa*.

The growth curve enabled us to determine if the dynamic range of the uvitex 2B method could be saturated. In a natural highly susceptible interaction (Ws-0 infected with *Hpa*

Waco9) the resolution of the acquired pictures, which is the basis for the automatic quantitation of pathogen growth, could be limiting. Seven days after infection there was no detectable increase in sporulation of *Hpa* Waco9 on both Col-0 and Ws-0. Approximately 50% of the Ws-0 leaf area was covered with conidiophores compared to approximately 10% for Col-0.

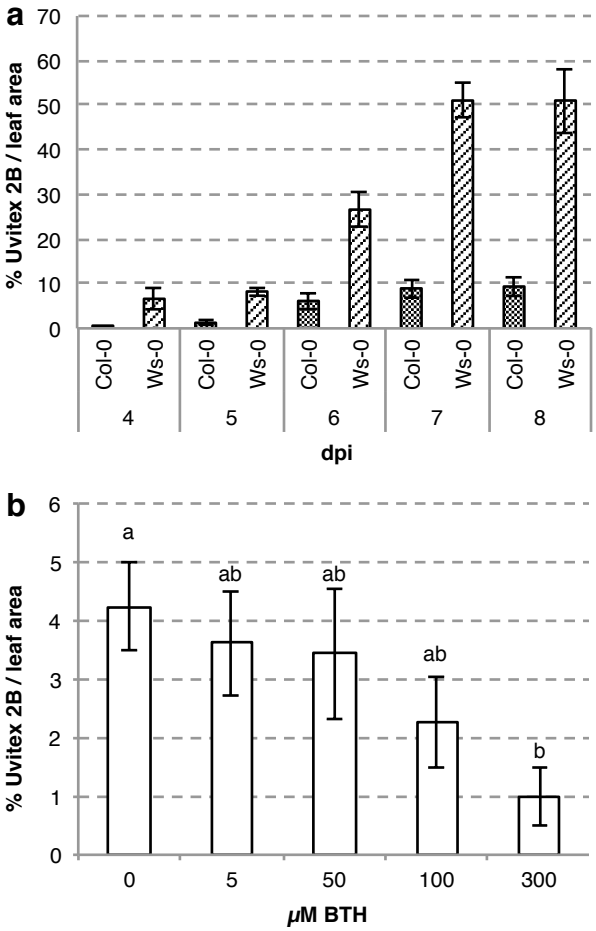



**Figure 7.4: Permutation of host result in altered *Hpa* Waco9 infection.** (a) Growth curve of *Hpa* Waco9 on 3 week old Col-0 and Ws-0 plants. (b) BTH pretreatment 3 days before infection with *Hpa* Waco9 result in increased resistance of Col-0 plants. The infection was scored at 6 dpi (days post infection). The experiments were repeated three times with similar results. Means with identical letter are not significantly different (Tukey test); Error bars denote standard error of the mean.

We interpret this as 7 days post infection represents the peak of sporulation, but we cannot rule out the possibility that the uvitex 2B method has been saturated locally. When leaves are heavily infected with *Hpa*, dense clusters of conidiophores form that are unresolvable

by the uvitex 2B method. This may lead to an underestimation of the overall level of infection. However, new patches of conidiophores did not form on Col-0 plants at later time points. Thus, we conclude the dynamic range of the uvitex 2B method was sufficient to assess moderate and highly susceptible interactions.

### 7.2.4 Uvitex 2B method confirms 300µM BTH induces resistance to *Hpa*

We wanted to independently confirm our initial finding that the uvitex 2B method could correctly determine infection level at various levels of increased resistance. BTH is a commonly used chemical that induces systemic acquired resistance in plants; previously 300 µM BTH has been applied to induce SAR (systemic acquired resistance) in *Arabidopsis* (Lawton et al. 1996). We used benzothiadiazole (BTH) to induce resistance in Col-0 plants to *Hpa* Waco9. Specifically, we wanted to test if quantifying infection with uvitex 2B would allow us to characterize any quantitative effects of varying concentrations of BTH.

We found that mock treatment or 5 µM, 50 µM and 100 µM BTH did not significantly affect *Hpa* Waco9 growth. While statistically insignificant, we did observe a slight decrease in *Hpa* Waco9 growth upon treatment with 100 µM BTH. As expected application of 300 µM BTH resulted in strongly restricted *Hpa* Waco9 growth (Figure 7.4). Thus, treatment of Col-0 plants with BTH independently confirms previous findings, and provides an additional example of the versatility of the uvitex 2B method.

### 7.2.5 Uvitex 2B stains Phytophthora infestans and powdery mildew

We tested if uvitex 2B would stain other plant-pathogens, due to the success of establishing an *Hpa* infection quantifying system. Two pathogens that are hampered by a

lack of rapid quantification methods are: powdery mildews i.e. *Blumeria graminis* and *P. infestans*. For example, lesion size of *P. infestans* infections on *N. benthamiana* is measured by ruler, or, by determining the area of infection, which is more time consuming (Bozkurt et al. 2011). For this reason we drop inoculated 4-week-old detached *N. benthamiana* leaves with a *P. infestans* 88069 zoospore suspension. Necrotrophic lesions were visible by eye 6 dpi; we stained the infected leaves using the uvitex 2B protocol developed for *Hpa* on *Arabidopsis*. Interestingly, we found that uvitex 2B readily stained aerial *P. infestans* mycelium with diminishing background fluorescence (Figure 7.5). To determine if uvitex 2B could stain aerial parts of powdery mildews we obtained an *Arabidopsis* Col-0 leaf heavily infected with powdery mildew (morpho-type similar to *Golovinomyces orontii*). We stained this leaf using the uvitex 2B protocol and found that the aerial mycelium readily was stained (Figure 7.5).

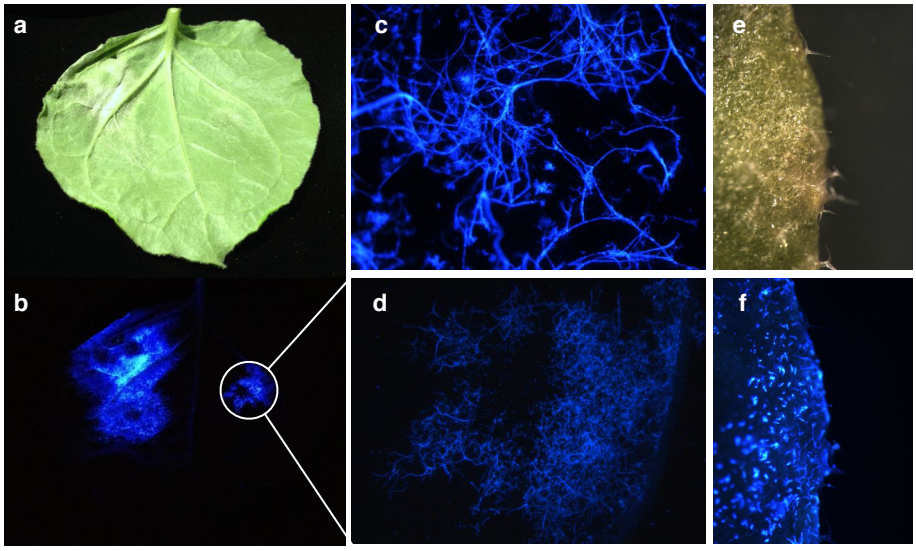



**Figure 7.5: Uvitex 2B stain aerial structures of *P. infestans* and Powdery mildew.** Detached *N. benthaminana* infected with *P. infestans* 88069 6 dpi brightfield (a) and uvitex 2B stained (b). c and d) are higher magnifications of (b). Powdery mildew infection of Col-0 plants (e) bright field and (f) uvitex 2B stained.

## 7.3 Discussion

We devised a simple protocol for infection quantification, with few steps, involving a non-toxic chemical that is cheap and requires standard equipment present in most plant biology

laboratories (Figure 2). The protocol is based on uvitex 2B, which specifically stains aerial *Hpa* structures and dramatically increases the Z-depth compared to transmitted bright field visualization. Since uvitex 2B specifically stains *Hpa* structures with diminishing background staining, we were able to automate the quantification process.

We verified the uvitex 2B method by showing that the uvitex 2B results correlated with the number of trypan blue stained conidiophores. Interestingly, trypan blue staining did not stain all conidiophores, a feature we did not observe on uvitex 2B stained leaves.

We additionally tested the uvitex 2B method by performing two experiments with known outcomes. First we performed a time course of a compatible interaction; consistent with expectations Ws-0 plants are more susceptible than Col-0 to *Hpa* Waco9. We noticed that the uvitex 2B method potentially underestimated the amount of sporulation in highly susceptible interactions. However, quantitative screens for differences in high susceptibility on cotyledons or mature plants have not been studied extensively. For example, a recent screen of T-DNA mutants of highly connected *Arabidopsis* proteins, which interact with candidate bacterial and oomycete effectors classifies cotyledons with more than 15 *Hpa* conidiophores in a 15+ category (Mukhtar et al. 2011). Similarly, only limited conidiophore counts have been performed on highly susceptible mutant adult plants (H. S. Kim 2002; Donofrio & Delaney 2001). In these instances the uvitex 2B method will give a rapid indication whether or not manual counting of highly susceptible individuals is justified.

While, the uvitex 2B method potentially underestimates the level of sporulation with extremely susceptible interactions, it excels at moderately high to low levels of infection. In moderate susceptible/resistance *Hpa-Arabidopsis* interactions, we were able to confirm that Col-0, *arf9-1* and *ein5-1* are more resistant to *Hpa* Waco9, compared to 35S:AFB1

that is more susceptible. Further, we treated plants with increasing amounts of BTH and quantitatively validated that 300 µM BTH significantly suppresses *Hpa* growth (Lawton et al. 1996). Thus, the uvitex 2B method accurately determines low to moderately high levels of infection. In conclusion, the uvitex 2B method is an attractive alternative for assessing susceptibility differences in the range from moderately susceptible to highly resistant.

**How does the uvitex 2B method compare to other methods?**

In principle there are multiple methods for quantifying *Hpa* growth, but only those that measure either sexual or asexual progeny reflect success on a given host. Two common measures are: conidiophores per plant, and conidiospores per plant or per gram fresh weight (McDowell et al. 2011; Fabro et al. 2011; Robert-Seilaniantz, Maclean, et al. 2011b). The outcomes of these different measures are comparable (Jorge Badel, unpublished results). We believe that the uvitex 2B method provides a compelling alternative to these methods for the following reasons. First of all, the automated uvitex 2B method adds extra reliability to data, as it distances the experimenter further from the sampling. Second, mutants and ecotypes can vary extensively in size and morphology, which could influence the results if no normalization is performed. For example, ecotypes with smaller leaves will due to the less available leaf surface be scored as more resistant per leaf compared to plants with more leaf area. To circumvent the problem of different plant sizes, experimenters normalize the amount of spores to plant fresh weight. However, normalizing to fresh weight will skew data, as fresh weight is inversely correlated with level of infection due to water loss from the infected tissue (our unpublished observations). For this reason *Pseudomonas* growth curves are also normalized to area (K. H. Sohn et al. 2007b; Zhou et al. 1998). Similarly, the uvitex 2B method is a relative measure of the infected area to entire area of the leaf. Thus, uvitex 2B analysis does not require

assumptions about plant size(s). Third, the automated uvitex 2B method is rapid: compared to traditional methods it is at least 5 fold faster on average. In our hands counting around 350 conidiophores on a mature *Ws-eds1* leaf takes approximately 20-40 minutes versus 30-60 seconds to assess level of infection with the uvitex 2B method.

As an example, performing an experiment where 6 different mutant lines are compared to wild type with 18 replicate measurements (in total 126 data points) takes 84 minutes on average to sample. The benefits are two-fold: More replicates of a given experiment can be performed in the same time. This will lead to a more thorough sampling of the plant-*Hpa* infection distribution, and thus lead to more reliable assessments of differentials in virulence. Further, as the uvitex 2B method enables rapid quantification of pathogen growth, quantitative screens with *Hpa* are now within reach. This could lead to discovery of novel components of the *Arabidopsis* defense, as most defense mutants have been identified in other types of screens (Petersen et al. 2000; Glazebrook et al. 1996).

We anticipate that the dye will be extensively employed, even without automating the sampling process, for the following reasons. Firstly, uvitex 2B clearly labels pathogenic structures and enhances the contrast, which results in an increased depth resolution for these structures (Figure 7.1). This will allow inexperienced experimenters to readily identify them. Second, uvitex 2B will help correctly quantifying weak infections, where some conidiophores might escape detection even by an experienced experimenter. Thus, for manual counting of conidiophores the dye will speed up the process, reduce health hazards and counting errors due to the increased contrast and Z-depth. It is currently used extensively in our lab for this purpose.

Given the apparent success of the uvitex 2B method to quantify *Hpa* growth, we tested if other pathogens could be stained in a similar manner. We found that uvitex 2B readily stains *Phytophthora* sp. and powdery mildews. Therefore, we envisage that similar

automated quantitation assays could be developed for these pathogens; thus opening the potential for a more rapid and less biased quantification of infection.

In conclusion, the uvitex 2B method is a novel application of a well-established fluorescent compound. By developing custom scripts, we can speed up the data acquisition and distanced the experimenter further from the sampling. We propose the use of the uvitex 2B method will advance studies in *Hpa-Arabidopsis* biology, as well as other plant/pathogen interactions.

# 8 General discussion

The question of how pathogens create an environment they can inhabit and produce offspring, either sexually or asexually, is the major challenge in the field of plant-microbe interactions. We focused on the *Arabidopsis-Albugo* system in this study and especially two aspects: what proteinaceous effectors do *Albugo laibachii* employ to suppress host defenses, and where and how they function inside the plant cell. And secondly, is resistance to *AlNc14* and *AcNc2* governed by the same *R*-genes as *Alem1* and *Acem2?*

## 8.1 Resistance to *Albugo* sp. is governed by multiple R-loci

We found that race *AlNc14* was avirulent on Ksk-1 due to recognition by the same *R*-genes, *RAC1* and *RAC3* as were reported for *Alem1* (Borhan et al. 2001). While, the identity of *RAC3* is unknown, *RAC1* encodes a TIR-NB-LRR, thus *AvrRac1* is present in both isolates (Borhan et al. 2004). The AvrRac1 protein could be similar in sequence in *AlNc14* and *Alem1*, as observed for avirulent alleles of ATR1 in *Hpa* Emoy2, Waco5, Maks9, and Maks9 (Rehmany et al. 2005).

In our characterization of resistances to *AcNc2*, we found that an additional resistance locus was present in Col-5. Volkan Cevik, who recently joined as a post doc in the lab had concurrently identified an additional resistance locus in Col-5 in his previous lab, which is active against another *A. candida* race *Acem2*. He mapped this locus and cloned *WRR5a* and *WRR5b,* which both are required for full resistance (Volkan Cevik, personal communication). Interestingly *WRR5a* and *WRR5b* also confer resistance to *AcNc2* suggesting that the two loci in Col-5 conferring resistance to *AcNc2* are *WRR4* and *WRR5a/b.* For this reason further work on mapping this *R*-gene will not be undertaken. However, the existence of multiple resistance genes confirms that *AcNc2* is less adapted to growth on *Arabidopsis* compared to *AlNc14.* This is consistent with the observation that

*AcNc2* is able to infect *Cardamine hirsuta* without causing trailing necrosis upon infection and thus might be more specialized for growth on *C. hirsuta* (Eric Kemen pers. comm.). In contrast to previously characterized *R*-genes active against *Albugo* sp. the necrotic resistance does not signal through *EDS1* (Chapter 7); therefore it could represent a novel source of quantitative resistance to *A. candida* species.

## 8.2 Do oomycetes have chitin in the cell wall?

We found that *Hpa* was stained by uvitex 2B, which is surprising, since uvitex 2B has been reported to stain algae and fungi strongly, and also bind to chitin (Koch & Pimsler 1987; Coleman et al. 1989). However, biochemical fractionation of the cell wall of various oomycetes suggests that it consists primarily of β-form glucans (Aronson et al. 1967; Bartnicki-Garcia 1968). Consistent with this, a histochemical study on spores and cysts of *Phytophthora cinnamomi* detected binding of ConA and soybean agglutinin to the cell surface, whereas wheat germ agglutinin did not bind. This suggests the presence of cellulose, but not chitin on the cell surface (Bacic et al. 1985). Further, ConA strongly labels the *Hpa* cell wall indicative of β-glucans in the cell wall (Carzaniga et al. 2001). For these reasons uvitex 2B may bind to other compounds than chitin in the cell wall.

Interestingly, the chemically related compound calcofluor white MR2 has been found to bind the β-form of polysaccharide polymers (Maeda & Ishida 1967). Thus, it is conceivable that uvitex 2B, like calcoflour white MR2, could bind β-forms of polysaccharide polymers. Thus, uvitex 2B could bind to β-form glucans in *Hpa* and *Phytophthora* cell walls.

While it seems most likely that uvitex 2B binds β-glucans in the cell wall, we cannot rule out the possibility of some binding to chitin. Carzaniga and colleagues (2001) found that *Hpa* conidia also bound WGA. This could be indicative of N-acetylglucosamine

containing glycoproteins, but could also suggest some chitin in the cell wall. A recent report determined that chitin is present on the surface of sporangiophores, sporangia and hyphal cell walls on the closely related oomycete *Plasmopara viticola* (Werner et al. 2002). Chitin was also found lining the hyphal cell wall of *P. ultimum* (Chérif et al. 1993). Similarly, noncrystalline chitosaccharides were found on the surface of *Aphanomyces euteiches* and application of the chitin synthase inhibitor Nikkomycin Z resulted in loss of cell wall integrity (Badreddine et al. 2008). Likewise, Guerriero and colleagues (Guerriero et al. 2010) identified SmCHS1 (*Saprolegnia monoica* chitin synthase 1) and SmCHS2 that are expressed at the hyphal tip of *Saprolegnia monoica* and function as a chitin synthase. Interestingly, both *Hpa* and *Phytophothora* have homologs of chitin synthases. As deletions of CHS genes can result in dramatic loss of pathogenicity, these genes present appealing targets for deletion analyses, which potentially could implicate chitin biosynthesis in the biology of *Peronosporales*. Deletion mutants could potentially also reveal insights into uvitex 2B binding to *Hpa* and *Phytophthora*. In addition, virulence assays on *Arabidopsis* mutants impaired in chitin perception, e.g. *cerk1* mutants, could reveal if an *Hpa* infection was associated with chitin perception.

## 8.3 Has functional screening of oomycete effectors been biased?

Most oomycete effectors were initially identified based on an avirulence function, but recent screening efforts suggest that effectors primarily suppress host defenses (Hein et al. 2009; Bozkurt et al. 2012; Rivas 2011). Two systematic studies of the potential virulence function of the predicted RXLR effectorome of *P. sojae* and *Hpa* support this notion (Fabro et al. 2011; Wang et al. 2011). For the tested *Hpa* RXLR effector candidates defense suppression was correlated with suppression of PTI; reduction of virulence caused by effectors was rare (Fabro et al. 2011). The key finding from testing of the RXLRs from

*P. sojae* was that most could suppress PCD triggered by BAX or INF1 (PTI), while some triggered PCD alone (Wang et al. 2011).

However, both screening approaches were slightly biased, as a positive outcome would imply a role for modulation of defense by the effector, whereas a negative result would be inconclusive. Fabro et al. (2011) do partially take this into account as they include 4 NC proteins, which are predicted not to alter virulence, in the screen. However NC2 enhances virulence on 3/12 ecotypes tested, thus underlining the requirement for follow up experiments with other pathogens such as *Hpa* in this case. Further, eukaryotic effectors were assayed using the fitness of phytopathogenic bacteria as output. Thus, only effectors that have a beneficial function in bacterial pathogenesis would be identified.

Obvious physiological differences between oomycetes and bacteria exist, and thus effectors that have functions specifically required for oomycete pathogenesis should exist. For example, oomycetes create haustoria, which are projected through the plant cell wall and are in contact with the EHM presumed to be a host derived membrane (Caillaud, Piquerez & Jones 2012a; Y.-J. Lu et al. 2012).

Penetration of host cells has been extensively studied in powdery mildews, such as *Blumeria graminis* f.sp. *hordei* (*Bgh*) a non-host of *Arabidopsis* that is unable to penetrate the *Arabidopsis* cell wall to create haustoria. However, *Arabidopsis* penetration (pen) mutants allow *Bgh* to penetrate the epidermal cell wall. Two pen-mutants *pen1* and *syp122* encode for syntaxin proteins that are required for vesicle transport (Collins et al. 2003). Kwon et al (2008) found that the PEN1 disease resistance primarily functions through a SNAP33 adaptor and a subset of vesicle-associated membrane proteins (VAMPs). Interestingly, VAMP721 and VAMP722 were required for resistance to the haustorial forming pathogens *Hpa*, *Bgh* and *G. orontii*, but not the bacterial pathogen *Pst* DC3000 (Kwon et al. 2008). A *vamp721/vamp722* double mutant is lethal, which Kwon and

coworkers attribute to loss of an essential secretory pathway. However, a *pen1/syp122* double mutant is also lethal. Surprisingly, the lethality of a *pen1/syp122* double mutant is dependent on *EDS1* (Ziguo Zhang et al. 2007). This suggests that the PEN1/SYP122 secretion pathway could be guarded by TIR-NB-LRRs. Such a scenario seems likely, as VAMP721 and VAMP722 were implicated in resistance to haustorial-forming pathogen and thus represent potential effector targets. Therefore, this PEN1/SYP122 dependent secretory pathway is one, presumably out of many, which are not required for bacterial virulence and thus unlikely targets of bacterial effectors. Therefore, it is likely oomycete effectors exist, which would not be identified in screens using bacterial virulence as output. However, there are further differences in virulence strategies employed by different pathogens, as illustrated by van Damme et al. (2005) who conducted a screen for gain of resistance to *Hpa* Cala2 and *Hpa* Waco9 in Ler-*eds1*-2 plants. Interestingly, the downy mildew resistant (dmr) mutants dmr1, dmr2, and dmr6, which are resistant to *Hpa* isolates, still show normal susceptibility to *Pst* DC3000 and the powdery mildew *Golovinomyces orontiii* (van Damme et al. 2005; Stuttmann et al. 2011).

Similarly, a screen for powdery mildew resistance (pmr) revealed that pmr5, pmr6 and mlo cause resistance to *Erysiphe cichoracearum* and *Golovinomyces orontiii,* but resistance to *Pst* DC3000 and *Hpa* was unaltered (Vogel et al. 2004; Vogel 2002; Consonni et al. 2006). These examples suggest that different pathways are specifically required for virulence of some pathogens, and thus only represent credible virulence targets for effectors of these pathogens.

The dmr-, pmr- and pen-screens relied on clear phenotypes (van Damme et al. 2005; Collins et al. 2003; Vogel & Somerville 2000). While such screenings revealed unique pathways and susceptibility factors, subtle quantitative phenotypes could have been missed. The primary reason for a lack of large-scale virulence screens for subtle

phenotypes is the lack of a high throughput screening methods. Although there is no silver bullet our uvitex 2B method provides a method, which yields reliable results and significantly reduces screening time. Therefore we envision it is feasible to quantitatively screen large populations of *Arabidopsis* mutants for differential susceptibility or resistance to *Hpa* isolates.

Similarly, high throughput semi-quantitative screens where the enhanced contrast of the uvitex 2B stained conidiophores would also allow identification of differential susceptibility to *Hpa* by eye, which normally would not be possible due to lack of contrast. As an example the Salk homozygous T-DNA collection could be screened upon infection with various *Hpa* isolates in this manner (Alonso et al. 2003). To identify direct virulence targets of effectors, transgenic *Arabidopsis* lines carrying candidate effectors could be tested with this method.

## 8.4 What types of effectors exist and has prediction been biased?

Oomycete effector prediction has primarily been based on data from avirulence proteins (R. Allen et al. 2004; Goritschnig et al. 2012; Armstrong et al. 2005; Dou et al. 2008). A defining characteristic for these is that the genes occur in expanded families usually without homology to conserved genes. The proteins have an N-terminal signal peptide and modular structure with an N-terminal delivery domain (i.e. RXLR or CRN) and a C-terminal effector domain that harbors the function (Schornack et al. 2009; Kamoun 2006). These characteristics have been the corner stone of effector prediction in other oomycete pathogens. Subsequent testing of effectors identified by these criteria has shown that some do have virulence functions, thus confirming predictions.

In *Albugo laibachii* we found that RXLRs are probably not an evolutionary important class of effectors by performing permutation experiments (Kemen et al. 2011). Instead proteins with an N-terminal CHXC motif were overrepresented in the secretome. Although the

CHXCs are different from the RXLRs they had a similar modular structure with a conserved N-terminus and a diverged C-terminus. The N-terminal CHXC domain was found to facilitate translocation into the plant cell. This required the CHXC motif, as mutations reduced delivery (Kemen et al. 2011). Likewise, we found that CHXCs contributed to virulence in EDV assays (Chapter 3, Kemen et al. 2011). Therefore, suggesting these to be bona fide effectors adhering to the aforementioned principles. However do other types of effectors exist, which may not have a defined uptake motif?

Initial characterization of the exportosome in *Plasmodium falciparum* predicted PTEX effectors as having a PEXEL motif preceding a signal peptide (Maier et al. 2008).

Interestingly, evidence has accumulated suggesting that PEXEL-negative (PNEP), which also lack signal peptides exist. The PNEPs skeleton-binding protein 1 (SBP1), and the membrane associated histidine-rich protein 1 (MAHRP1), and ring-exported protein 2 (REX2) all lack a signal peptide but share a defining feature, which is the presence of a single transmembrane domain (Spielmann & Gilberger 2010). However, they may be secreted via conventional secretory pathways, as ER intermediates are detected. In contrast, the heme detoxification protein (HDP) contains no hydrophobic stretches that could suggest a TM domain or signal peptide. Interestingly, HDP secretion is Brefeldin A-insensitive and HDP is thus secreted in an unconventional manner (Spielmann & Gilberger 2010).

Given effectors in *P. falciparum* are secreted into the host cell without known translocation motifs or signal peptides similar mechanisms may exist in phytopathogenic oomycetes. For this reason, it is likely that the effectorome of phytopathogenic oomycetes has been underestimated and predictions skewed. Our data suggests that this indeed could be the case. Contrary to the conventional knowledge that most effectors do not have homology to proteins of known function (Goritschnig et al. 2012; Rehmany et al. 2005; Armstrong et al.

2005; Dou et al. 2008) we found that CHXC1, which encodes a 684 amino acid protein with homology to a HECT E3 ligase, is able to suppress host defenses when expressed *in planta* (Chapter 4).

Interestingly HECT E3 ligases are a significantly expanded family of proteins in oomycetes, which in some cases carry an N-terminal signal peptide indicative of secretion into the host cell. We propose that HECT E3 ligases play similar roles in other oomycetes and may even have diversified functions as several cytosolic and secreted HECT E3 ligases gene families are expanded in other oomycetes. This type of HECT E3 ligases may be a core type of effector, as they are found in all oomycetes. The possibility that HECT E3 ligases are virulence factors in pathogenic oomycetes is unprecedented and a major finding of this thesis.

## 8.5 What is the function of CHXC1 *in planta*?

Recognition of PAMPs comprises the first layer of plant defense; responses triggered by recognition of PAMPs from pathogens such as oomycetes, fungi and bacteria overlap to some extent (Zipfel 2008). Genetic and network analyses of *Arabidopsis* mutants have shown that signaling events often converge on a limited numbers of genes (hubs) downstream of activation (Tsuda et al. 2009; Mutwil et al. 2009; Aarts et al. 1998). For example, the resistance mutant *eds1* is more susceptible to fungi, oomycetes and bacteria (Aarts et al. 1998).

We found that delivery of ΔSP-CHXC1, but not ΔSP-CHXC1(C651A) via EDV resulted in enhanced growth of *Pst* DC3000 Δ*AvrPto*/Δ*AvrPtoB*. As AvrPto and AvrPtoB are two key effectors of *Pst* DC3000 virulence that suppress PTI (He et al. 2004), this points to a potential role for CHXC1 in PTI suppression. However, it seems likely that CHXC1, AvrPto, and AvrPtoB have different modes of action, because AvrPto and

AvrPtoB are found at the plasma membrane but CHXC1 is nuclear. Therefore, CHXC1 may target defense components other than PRRs that are involved in PTI in *Arabidopsis*.

We now have the appropriate tools in *Arabidopsis* to test if the enhanced virulence of *Hpa* (Chapter 4) is dependent on PTI suppression. For example, upon stimulus with flg22, elf18 and chitin in estradiol-induced Col-0 plants expressing either ΔSP-CHXC1 or ΔSP-CHXC1(C651A), suppression of ROS accumulation, MAPK activation and marker gene expression could be tested. In addition, infection experiments with *Hpa* Emco5, *Pst* DC3000 Δ*AvrPto*/Δ*AvrPtoB*, *AlNc14*, *P. capsici* could be performed. These will lead to a deeper understanding of CHXC1 function and dissect whether virulence is dependent on suppression of PTI.

While CHXC1 may suppress PTI it could suppress other components of *Arabidopsis* defense*. A. laibachii* has been reported to partially suppress *lsd1* induced run away cell death and suppress ATR1-WsB resistance conferred by RPP1-Nd (Cooper et al. 2008). In addition, *A. laibachii* is able to partially suppress *acd11* dependent cell death (our unpublished results). The histone methyltransferase LAZ2/SDG8 (Lazarus 2 or SET (Su(var)3-9, E(z) and Trithorax-conserved) DOMAIN GROUP 8) suppresses *acd11* (Accelerated cell death 11) dependent cell death via epigenetic transcriptional regulation of the ACD11 guarding R-protein LAZ5 (Palma et al. 2010). An intriguing possibility is that CHXC1 interacts with LAZ2/SDG8 in the nucleus and either inhibits proper signaling or targets it for degradation via the ubiquitin proteasomal pathway.

In a pioneering study Mukhtar et al. (2011) performed a systematic yeast-two-hybrid-interaction screen of 83 effectors from *P. syringae*, and *Hpa* against three classes of *Arabidopsis* immune system proteins, and ~8000 other Arabidopsis

proteins. Significantly, some proteins, termed hubs, are highly connected to both *Arabidopsis* proteins and pathogen effectors. The effect of removing these hubs from the network resulted in altered disease susceptibility. Thus, knowledge of interacting proteins can lead to a clearer picture of the function of the given effector.

For this reason, we initiated a yeast-two-hybrid screen to identify CHXC1 interacting proteins (in collaboration with J. Steinbrenner UC Warwick). In addition, co-immuno-precipitations of interacting proteins of ΔSP-CHXC1 and ΔSP-CHXC1(C651A) could be identified by mass spectrometry. These two orthogonal approaches will identify interactors that represent potential substrates of CHXC1, but also aid in identifying the components that may be required for ubiquitination *in vitro.* Analyses of T-DNA knockout and over-expressors of genes whose corresponding proteins interact with CHXC1 will determine their functional relevance in an *Albugo-Arabidopsis* interaction.

Transcription occurs in the nucleus and therefore transcription factors have to be imported into the nucleus. Upon pathogen attack a rapid and dynamic change in pathogen-responsive gene transcription occurs. In many cases the efficiency of the host response is determined by changes in nuclear/cytoplasmic distributions of interacting signal transducers or transcription factors (García & Parker 2009).

A systematic localization-study of *Hpa* effectors revealed that 66% are nuclear localized and 50% of these are exclusively localized inside the nucleus (Caillaud, Piquerez, Fabro, et al. 2012b). The fact that many effectors are translocated into the host cell nucleus suggests that modulation of transcriptional responses or direct interaction with chromatin is a common strategy employed by pathogens to promote host susceptibility (Rivas 2011).

We found that CHXC1 is localized to the nucleus in both *Arabidopsis* and *N. benthamiana*. Since the 105 kDa GFP-ΔSP-CHXC1 fusion construct is larger than the size limit for proteins to freely diffuse into the nucleus, we speculate that nuclear uptake of the protein is an active process.

The best described pathway for uptake of proteins is through the nuclear pore complex (NPC), which facilitates most nucleocytoplasmic transport of proteins. Proteins are guided through the NPC by importins that bind NLS sequences encoded in the protein. Proteins without NLS can move through the NPC, either via interaction with NLS-containing proteins or by directly binding the components of the NPC. While all possible NLS sequences have not been elucidated the sequence usually consists of positively charged amino acids (Lange et al. 2006; Kosugi et al. 2008). Thus, the region of CHXC1 with positively charged amino acids $K_{164}$KFSQRQRGAQRRK$L_{178}$ could function as a mono-partite NLS.

The distinct nuclear localization of CHXC1 enables assignment of function to this compartment. We hypothesize nuclear localization of CHXC1 is required for defense suppression. Future experiments on plants expressing CHXC1ΔNLS, assuming this is excluded from the nucleus, could determine if localization is required for defense suppression.

While CHXC1 could target host components located within the nucleus, another possibility is that CHXC1 targets components of the nuclear import and export machinery for degradation. The machinery has been implicated in defense predominantly by elegant work in the lab of Xin Li. Mutations in MOS6 (MODIFIER OF SNC1, 6), MOS3, and MOS7 were isolated in a screen for suppressors of *snc1* (Palma et al. 2005; Y. T. Cheng et al. 2009; Y Zhang 2005). They encode an importin α3, and yeast homologs of the nucleoporins Nup96 and Nup88 respectively. Interestingly, in

some cases mutations in importins and nucleoporins resulted in varying degrees of enhanced disease susceptibility to virulent pathogens, but also loss of *R*-gene mediated resistance to avirulent pathogens (Palma et al. 2005; Y. T. Cheng et al. 2009; Y Zhang 2005; Wiermer et al. 2012). This leads us to speculate that α-importins and components of the NPC could represent the targets of CHXC1. In this scenario, these components are modified by CHXC1 and as a side effect CHXC1 is imported into the nucleus. Determining the defense suppressing properties of a GFP-NES-ΔSP-CHXC1 construct, which allows CHXC1 to enter the nucleus but not accumulate due to the nuclear export signal, could test this hypothesis.

We have currently **not** shown that CHXC1 has E3 ligase activity, and strategies to further test for catalytic activity include: complementation of a yeast *Rps5* mutant, optimizing *in vitro* CHXC1 assay conditions, and performing CoIPs with epitope tagged CHXC1 to determine if this is ubiquitinated dependent on cys651. These strategies have been discussed extensively in chapter 4.

As we observed that CHXC1 cys651 is required for virulence of *Pst* DC3000 and *Hpa* Noco2, this could indicate that CHXC1 could possess HECT E3 ligase activity. However, cys651 could also be important for proper folding, stability, or interaction with other proteins. The experiments proposed would shed light on this interesting effector, which has homologs in other oomycetes.

### 8.6 Are SSP6 variants causal for *A. laibachii* host range and does SSP6-2c suppress PTI via interaction with BAK1?

We identified SSP6 as a candidate effector with different 7 variants; EST mining revealed that two alleles SSP6-2c and SSP6-A were predominantly expressed in *AlNc14*. Through a synteny analysis we determined that the *A. laibachii* locus containing alleles of SSP6 was absent from the closely related species *A. candida* (Chapter 5). Recently, Sohn and

coworkers (2012) described that the *Pto* T1 effector HopAS1 is causal for avirulence on *Arabidopsis* and absent from *P. syringae* strains virulent on *Arabidopsis.*

Since the SSP6 locus is absent from *A. candida* variants of SSP6 could be causal for *A. laibachii* host range. The hypothesis could be tested by expression of SSP6 alleles in other *Brassicaceae* species (e.g. *B. rapa* the host of *Ac2VRR* (Borhan et al. 2004)) to determine if they trigger HR.

We studied SSP6-A and SSP6-2c and found that both localized to the plasma membrane, but upon *AlNc14* infection SSP6-2c also localized to the EHM. This was correlated with the ability of SSP6-2c, but not SSP6-A, to enhance the virulence of *P. infestans* and suppress flg22 dependent ROS in *N. benthamiana.*

Interestingly, SSP6-2c could suppress flg22 but not chitin dependent ROS accumulation, which suggests that SSP6-2c targets specific components upstream of the ROS generating machinery (Segonzac & Zipfel 2011). Flg22 perception by FLS2 triggers PTI-signaling via a direct interaction with the auxiliary membrane bound kinase BAK1 and presumably also other unidentified components(Schwessinger et al. 2011; Roux et al. 2011). In addition other PRRs, for example EFR, trigger PTI via BAK1 (Schwessinger et al. 2011). As oomycetes do not produce bacterial flagellin, it is unlikely that SSP6-2c targets FLS2 directly to suppress the flg22 triggered PTI. Instead we favor a model where SSP6-2c modulates or hinders proper function of downstream signaling components shared between multiple PRRs i.e. BAK1 which is required for flg22 but not chitin dependent PTI. Such a mechanism would explain why SSP6-2c can suppress flg22 dependent PTI and presumably also would suppress PTI triggered by *A. laibachii* derived PAMPs or other PAMPs in general. Since the *N. benthamiana* homolog of BAK1, NbSERK3, is required for basal resistance to *Phytophthora* sp (Chaparro-Garcia et al. 2011) this model also explains the enhanced virulence of *P. infestans* blue 13 on *N. benthamiana* leaves expressing SSP6-2c.

## 8.7 What are *Albugo* sp. avirulence genes?

We found that CHXCs, CRNs and SSPs from *A. laibachii* were delivered to the plant cell and were able to enhance the virulence of *Pst* DC3000 (Kemen et al. 2011, Chapter 3). In Chapter 4 we studied CHXC1 in more detail and found, consistent with EDV experiments, that expression of CHXC1 *in planta* enhanced the virulence of *Hpa* Noco2. Likewise expression of SSP6-2c in *N. benthamiana* enhanced the virulence of *P. infestans* (Chapter 5). Thus, these proteins function as predicted for effectors.

However, it is currently unknown if these effectors, and more broadly effector classes, are causal for host range e.g. by causing avirulence on a given host accessions. To determine if these effector classes contain avr proteins three complementary approaches could be pursued: one relying on classical genetics, another on mutagenesis, and a third on a comparative population genomics effector screening.

Firstly to genetically define *A. laibachii avr*-genes a sexual cross between isolates with different virulence profiles could be generated. The resulting $F_2$ progeny should be used to inoculate a host accession, resistant to one of the parents, and test for segregating *avr-R* activity. The identity of the causal *avr*-gene could be identified through map based cloning on the virulent $F_2$ progeny. We found that *AlNc14* is resistant to HR-5 but susceptible to *Alem1* (Kemen et al. 2011). For this reason it is feasible that a cross between these isolates could lead to discovery of an *A. laibachii avr*-gene.

Alternatively, *avr*-genes could potentially be identified via mutagenesis of *A. laibachii*. The sexual progeny derived from a mutagenized *A. laibachii* parent would be screened for gain of virulence on plants, which are resistant to the non-mutagenized *A. laibachii*. In this case we predict that the cognate *avr*-gene is mutagenized in the virulent $F_2$ progeny. To identify the causal *avr*-gene whole genome sequencing of multiple individual mutants

should reveal gene(s) that are mutated in all individual mutants. Otherwise map based cloning could be applied.

These two approaches might identify *avr*-genes, but they could also identify factors, which could be epistatic to the *avr*-gene; for example, proteins required for avr protein delivery, maturation etc. Therefore further experiments are required to elucidate if the putative *avr*-gene is causal avirulence. The ideal experiment is to express the recognized allele of the avr-gene in the virulent *A. laibachii* parent and test for gain of avirulence. However we are unable to transform *A. laibachii.* Therefore tests of both alleles of the putative *avr*-gene with the cognate *R*-gene *in planta* would substantiate the claim that it is sufficient for host range (Bailey et al. 2011; R. Allen et al. 2004; Rehmany et al. 2005; K. H. Sohn et al. 2007b).

The two previously mentioned strategies for identifying an avirulence gene require sexual propagation. A successful cross between *A. candida* race 2 and race 7 has been reported (Adhikari et al. 2003). However, the infection of *Arabidopsis* plants with *A. laibachii* oospores requires optimization in our hands. To circumvent the requirement for sexual reproduction, CHXCs, SSPs, CRNs or other candidate effectors could be screened for appearance of HR by transient co-expressed with known *R*-genes in *N. benthamiana* (Krasileva et al. 2010; Bos et al. 2006). The possibility of success of such an approach has doubled with our identification of a second *R*-gene, *WRR5*, active against *A. candida* (Volkan Cevick pers. Comm., Chapter 7). Thus, the approach could be applied with the *R*-genes *WRR4* and *WRR5* to identify *AvrWRR4* and *AvrWRR5* from *AcNc2*. Likewise, *AvrRac1* could be potentially be identified by transiently expressing *A. laibachii* candidate effectors with RAC1 in *N. benthamiana*.

The third strategy exploits *in planta* expression. Assuming that an *avr*-gene is identical or highly similar within avirulent isolates and very dissimilar or absent from virulent isolates,

then the *avr*-gene could be identified through a comparative population genomics approach. Such a strategy has been successfully applied to identify Ave1 (Avirulence on Ve1 tomato) from *Verticillium dahliae* and *Verticillium albo-atrum* (de Jonge et al. 2012). In brief, the proteomes of multiple *Albugo* isolates with varying virulence profiles on host accessions would be screened for such genes.

To determine if these candidate *avr*-genes are causal for avirulence in the given host accession they could be delivered to the plant cell and tested for recognition either by EDV or using a GUS eclipse co-bombardment assay on resistant plants (Bailey et al. 2011; R. Allen et al. 2004; Rehmany et al. 2005; K. H. Sohn et al. 2007b). In case the cognate *R* gene is known *N. benthamiana* transient co-expression could be employed. This approach is currently being pursued by Oliver Furzer in the lab.

## 8.8 Proof and understanding of oomycete effector translocation into plant cells is a bottleneck

To date most papers on oomycete effectors assume that the effector is delivered into the host cell; translocation is rarely tested. The main reason is a scarcity of proper methods. Currently, generation of oomycete transformants is technically challenging or impossible (Schornack et al. 2010; Kemen et al. 2011; Mcleod et al. 2008; Judelson et al. 1991). In cases where transformants expressing fluorescent effector fusions have been generated signal accumulation has been demonstrated around the haustorium but not conclusively in the cytosol (Whisson et al. 2007; Gilroy et al. 2011). For these reasons, convincing examples of oomycete effector delivery into host cells are currently scarce.

Ariane Kemen in the lab has developed a series of excellent, but technically challenging, assays to determine translocation of effectors into the host cell. In a natural *AlNc14* infection of Col-THO leaves SSP6 was immuno-localized using specific antibodies.

Interestingly, SSP6 was exclusively localized around the *AlNc14* haustrium presumably in the EHMx and around the EHM (Ariane Kemen personal comm.). Thus, the immuno-localization is consistent with transgenic experiments in *Arabidopsis* and *N. benthamiana*, where we find that SSP6-2c but not SSP6-A, is localized around the *AlNc14* haustorium and can suppress ROS dependent PTI.

However, an important issue is which side of the EHM SSP6 variants bind. While, we found SSP6-A and SSP6-2c bind strongly to the inner side of the plasma membrane, it may also bind the EHM layer facing into the EHMx. However, this binding may be weaker due to either lack of protein interaction partners or a asymmetric distribution of phospholipids in the membrane bilayer (Quinn 2002).

Unfortunately our CHXC1 antibody also recognized endogenous *Arabidopsis* proteins hindering it use in immuno-localizations.

### 8.9 *P. capsici* expressing CHXC1 become more virulent on *Arabidopsis*

An intriguing observation initially reported by Cooper et al. (2008) is that *Alem1* can suppress ETI triggered by *RPP1* recognition of ATR1$^{Emoy2}$ and similarly suppress RPW8 mediated resistance to powdery mildew fungus (morphotype similar to *Erysiphe cruciferarum*). Likewise, *P. capsici* is normally unable to infect *Arabidopsis*, but preinfection with *AlNc14* suppresses host defenses against *P. capsici* allowing it to colonize the leaf (Ariane Kemen pers. comm.).

Remarkably, *P. capsici* transformants expressing CHXC1 driven by the *ham34* promoter in pNC2 gain the ability to infect *Arabidopsis* Col-0 and Col-Tho leaves (Ariane Kemen pers. comm.). In contrast *P. capsici* transformants carrying pNC2 empty vector are avirulent on *Arabidopsis* (Ariane Kemen personal comm.). We interpret these important results as CHXC1 elevating virulence of *P. capsici* on *Arabidopsis* and is presumably secreted from *P. capsici* into the host cell.

We favor a model where CHXC1 is expressed in *AlNc14* and secreted into the host cell nucleus, where it suppresses host defenses dependent on catalytic activity of the HECT E3 ligase. However, we cannot exclude the possibility that CHXC1 blocks proper defense signaling dependent on cys651 and mimics a HECT E3 ligase but is without enzymatic function.

## 8.10 Concluding remarks

The work of the SLJ *Albugo*-group has over the last 4 years dramatically expanded our knowledge on *Albugo* biology. The main findings in this thesis are: (i) Contrary to expectations, which initially led us astray, RXLRs are not a major class of effectors in *Albugo laibachii*. Instead we identified CHXCs as a plausible group of effectors and presumably major virulence factors in *Albugo* sp. (Kemen et al. 2011). Further, (ii) we have shown that CHXC1 a secreted protein with homology to a HECT E3 ligase, which requires cys651 for defense suppression. Secreted HECT E3 ligases are widespread in pathogenic oomycetes. Hopefully, the characterization of CHXC1 will lead to a paradigm shift, where it becomes generally accepted in the oomycete community that in addition to fast evolving highly redundant effectors, a set of core effectors exist, which may be essential for virulence.

## 9. References


Aarts, N. et al., 1998. Different requirements for EDS1 and NDR1 by disease resistance genes define at least two R gene-mediated signaling pathways in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, 95(17), pp.10306–10311.

Abramovitch, R.B. et al., 2006. Type III effector AvrPtoB requires intrinsic E3 ubiquitin ligase activity to suppress plant cell death and immunity. *Proceedings of the National Academy of Sciences of the United States of America*, 103(8), pp.2851–2856.

Adhikari, T.B. et al., 2003. Genetic and molecular analyses in crosses of race 2 and race 7 of albugo Candida. *Phytopathology*, 93(8), pp.959–965.

Allen, R. et al., 2004. Host-parasite coevolutionary conflict between Arabidopsis and downy mildew. *Science (New York, NY)*, 306(5703), p.1957.

Allen, R.L. et al., 2008. Natural variation reveals key amino acids in a downy mildew effector that alters recognition specificity by an Arabidopsis resistance gene. *Molecular plant pathology*, 9(4), pp.511–523.

Alonso, J.M. et al., 2003. Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science (New York, NY)*, 301(5633), pp.653–657.

Anderson, D.M. & Frank, D.W., 2012. Five mechanisms of manipulation by bacterial effectors: a ubiquitous theme. *PLoS pathogens*, 8(8), p.e1002823.

Armstrong, M.R. et al., 2005. An ancestral oomycete locus contains late blight avirulence gene Avr3a, encoding a protein that is recognized in the host cytoplasm. *Proceedings of the National Academy of Sciences of the United States of America*, 102(21),

pp.7766–7771.

Arnold, K. et al., 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics (Oxford, England)*, 22(2), pp.195–201.

Aronson, J.M., Cooper, B.A. & Fuller, M.S., 1967. Glucans of oomycete cell walls. *Science (New York, NY)*, 155(3760), pp.332–335.

Austin, R.S. et al., 2011. Next-generation mapping of Arabidopsis genes. *The Plant Journal*, 67(4), pp.715–725.

Bacic, A., Williams, M.L. & Clarke, A.E., 1985. Studies on the cell surface of zoospores and cysts of the fungus Phytophthora cinnamomi: nature of the surface saccharides as determined by quantitative lectin binding studies. *Journal of Histochemistry & Cytochemistry*, 33(5), pp.384–388.

Badreddine, I. et al., 2008. Cell Wall Chitosaccharides Are Essential Components and Exposed Patterns of the Phytopathogenic Oomycete Aphanomyces euteiches. *Eukaryotic Cell*, 7(11), pp.1980–1993.

Bailey, K. et al., 2011. Molecular Cloning of ATR5 Emoy2from Hyaloperonospora arabidopsidis, an Avirulence Determinant That Triggers RPP5-Mediated Defense in Arabidopsis. *Molecular plant-microbe interactions : MPMI*, 24(7), pp.827–838.

Bartnicki-Garcia, S., 1968. Cell wall chemistry, morphogenesis, and taxonomy of fungi. *Annual review of microbiology*, 22, pp.87–108.

Baxter, L. et al., 2010. Signatures of adaptation to obligate biotrophy in the Hyaloperonospora arabidopsidis genome. *Science (New York, NY)*, 330(6010),

pp.1549–1551.

Beakes, G.W. & Sekimoto, S., 2009. *The evolutionary phylogeny of oomyetes - Insights gained from studies of holocarpic parasite of algae and invertebrates* K. Lamour & S. Kamoun, eds., Hoboken, NJ, USA: John Wiley & Sons, Inc.

Belkhadir, Y. et al., 2004. Arabidopsis RIN4 negatively regulates disease resistance mediated by RPS2 and RPM1 downstream or independent of the NDR1 signal modulator and is not required for the virulence functions of bacterial type III effectors AvrRpt2 or AvrRpm1. *The Plant cell*, 16(10), pp.2822–2835.

Beck, M. et al., 2012. The INs and OUTs of pattern recognition receptors at the cell surface. *Current Opinion in Plant Biology*, 15(4), pp.367–374.

Bendtsen, J.D. et al., 2004. Improved prediction of signal peptides: SignalP 3.0. *Journal of molecular biology*, 340(4), pp.783–795.

Bernoux, M. et al., 2011. Structural and functional analysis of a plant resistance protein TIR domain reveals interfaces for self-association, signaling, and autoregulation. *Cell Host and Microbe*, 9(3), pp.200–211.

Bhattacharjee, S. et al., 2008. Maurer's clefts of Plasmodium falciparum are secretory organelles that concentrate virulence protein reporters for delivery to the host erythrocyte. *Blood*, 111(4), pp.2418–2426.

Bittner-Eddy, P.D. et al., 2000. RPP13 is a simple locus in Arabidopsis thaliana for alleles that specify downy mildew resistance to different avirulence determinants in Peronospora parasitica. *The Plant journal : for cell and molecular biology*, 21(2), pp.177–188.

Block, A. & Alfano, J.R., 2011. Plant targets for Pseudomonas syringae type III effectors: virulence targets or guarded decoys? *Current Opinion in Microbiology*, 14(1), pp.39–46.

Boddey, J.A. et al., 2010. An aspartyl protease directs malaria effector proteins to the host cell. *Nature*, 463(7281), pp.627–631.

Boller, T. & Felix, G., 2009. A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annual review of plant biology*, 60, pp.379–406.

Bonardi, V. et al., 2012. A new eye on NLR proteins: focused on clarity or diffused by complexity? *Current opinion in immunology*, pp.1–10.

Bonardi, V. et al., 2011. Expanded functions for a family of plant intracellular immune receptors beyond specific recognition of pathogen effectors. *Proceedings of the National Academy of Sciences of the United States of America*.

Borhan, M.H. et al., 2004. The arabidopsis TIR-NB-LRR gene RAC1 confers resistance to Albugo candida (white rust) and is dependent on EDS1 but not PAD4. *Molecular plant-microbe interactions : MPMI*, 17(7), pp.711–719.

Borhan, M.H. et al., 2001. White rust (Albugo candida) resistance loci on three Arabidopsis chromosomes are closely linked to downy mildew (Peronospora parasitica) resistance loci. *Molecular plant pathology*, 2(2), pp.87–95.

Borhan, M.H. et al., 2008. WRR4 encodes a TIR-NB-LRR protein that confers broad-spectrum white rust resistance in Arabidopsis thaliana to four physiological races of Albugo candida. *Molecular plant-microbe interactions : MPMI*, 21(6), pp.757–768.

Borhan, M.H. et al., 2010. WRR4, a broad-spectrum TIR-NB-LRR gene from Arabidopsis thaliana that confers white rust resistance in transgenic oilseed brassica crops. *Molecular plant pathology*, 11(2), pp.283–291.

Bos, J.I.B. et al., 2009. Distinct amino acids of the Phytophthora infestans effector AVR3a condition activation of R3a hypersensitivity and suppression of cell death. *Molecular plant-microbe interactions : MPMI*, 22(3), pp.269–281.

Bos, J.I.B. et al., 2010. Phytophthora infestans effector AVR3a is essential for virulence and manipulates plant immunity by stabilizing host E3 ligase CMPG1. *Proceedings of the National Academy of Sciences of the United States of America*, 107(21), pp.9909–9914.

Bos, J.I.B. et al., 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 : for cell and molecular biology*, 48(2), pp.165–176.

Botella, M.A. et al., 1998. Three genes of the Arabidopsis RPP1 complex resistance locus recognize distinct Peronospora parasitica avirulence determinants. *The Plant cell*, 10(11), pp.1847–1860.

Boutemy, L.S. et al., 2011. Structures of Phytophthora RXLR effector proteins: a conserved but adaptable fold underpins functional diversity. *Journal of Biological Chemistry*.

Bozkurt, T.O. et al., 2012. Oomycetes, effectors, and all that jazz. *Current Opinion in Plant Biology*, pp.1–10.

Bozkurt, T.O. et al., 2011. Phytophthora infestans effector AVRblb2 prevents secretion of a plant immune protease at the haustorial interface. *Proceedings of the National Academy of Sciences*, 108(51), pp.20832–20837.

Burgess, R.R., 2009. *Refolding Solubilized Inclusion Body Proteins* 1st ed, Elsevier Inc.

Cabral, A. et al., 2012. Nontoxic Nep1-like proteins of the downy mildew pathogen Hyaloperonospora arabidopsidis: repression of necrosis-inducing activity by a surface-exposed region. *Molecular plant-microbe interactions : MPMI*, 25(5), pp.697–708.

Caillaud, M.-C., Piquerez, S.J.M. & Jones, J.D.G., 2012a. Characterization of the membrane-associated HaRxL17 Hpa effector candidate. *Plant Signaling & Behavior*, 7(1), pp.145–149.

Caillaud, M.-C., Piquerez, S.J.M., Fabro, G., et al., 2012b. Subcellular localization of the Hpa RxLR effector repertoire identifies a tonoplast-associated protein HaRxL17 that confers enhanced plant susceptibility. *The Plant journal : for cell and molecular biology*, 69(2), pp.252–265.

Carzaniga, R., Bowyer, P. & O'Connell, R., 2001. Production of extracellular matrices during development of infection structures by the downy mildew Peronospora parasitica. *New Phytologist*, 149(1), pp.83–93.

Century, K.S. et al., 1997. NDR1, a pathogen-induced component required for Arabidopsis disease resistance. *Science (New York, NY)*, 278(5345), pp.1963–1965.

Chang JH, et al. 2000. avrPto enhances growth and necrosis caused by Pseudomonas syringae pv.tomato in tomato lines lacking either Pto or Prf. Mol Plant Microbe Interact. 13(5):568-71.

Chang, J.H. et al., 2005. A high-throughput, near-saturating screen for type III effector genes from Pseudomonas syringae. *Proceedings of the National Academy of Sciences of the United States of America*, 102(7), pp.2549–2554.

Chang, H.H. et al., 2008. N-terminal processing of proteins exported by malaria parasites. *Molecular and biochemical parasitology*, 160(2), pp.107–115.

Chaparro-Garcia, A. et al., 2011. The Receptor-Like Kinase SERK3/BAK1 Is Required for Basal Resistance against the Late Blight Pathogen Phytophthora infestans in Nicotiana benthamiana. *PLoS ONE*, 6(1), p.e16608.

Chen, X. et al., 2010. An eQTL Analysis of Partial Resistance to Puccinia hordei in Barley R. Allaby, ed. *PLoS ONE*, 5(1), p.e8598.

Cheng, W. et al., 2011. Structural Analysis of Pseudomonas syringae AvrPtoB Bound to Host BAK1 Reveals Two Similar Kinase-Interacting Domains in a Type III Effector. *Cell Host and Microbe*, pp.1–30.

Cheng, Y.T. & Li, X., 2012. Ubiquitination in NB-LRR-mediated immunity. *Current Opinion in Plant Biology*, pp.1–8.

Cheng, Y.T. et al., 2009. Nuclear Pore Complex Component MOS7/Nup88 Is Required for Innate Immunity and Nuclear Accumulation of Defense Regulators in Arabidopsis. *THE PLANT CELL ONLINE*, 21(8), pp.2503–2516.

Cheung, F. et al., 2008. Analysis of the Pythium ultimum transcriptome using Sanger and Pyrosequencing approaches. *BMC genomics*, 9, p.542.

Chérif, M., Benhamou, N. & Bélanger, R.R., 1993. Occurrence of cellulose and chitin in the hyphal walls of Pythium ultimum: a comparative study with other plant pathogenic fungi. *Canadian journal of microbiology*, 39(2), pp.213–222.

Chisholm, S.T. et al., 2005. Molecular characterization of proteolytic cleavage sites of the Pseudomonas syringae effector AvrRpt2. *Proceedings of the National Academy of Sciences of the United States of America*, 102(6), pp.2087–2092.

Chisholm, S.T. et al., 2006. Host-microbe interactions: shaping the evolution of the plant immune response. *Cell*, 124(4), pp.803–814.

Choi, Y. et al., 2008. Evidence for uncharted biodiversity in the Albugo candida complex, with the description of a new species. *Mycological research*, 112(11), pp.1327–1334.

Choi, Y.-J., Hong, S.-B. & Shin, H.-D., 2006. Genetic diversity within the Albugo candida complex (Peronosporales, Oomycota) inferred from phylogenetic analysis of ITS rDNA and COX2 mtDNA sequences. *MOLECULAR PHYLOGENETICS AND EVOLUTION*, 40(2), pp.400–409.

Chou, S. et al., 2011. Hyaloperonospora arabidopsidis ATR1 effector is a repeat protein with distributed recognition surfaces. *Proceedings of the National Academy of Sciences*, 108(32), pp.13323–13328.

Clough, S.J. & Bent, A.F., 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *The Plant journal : for cell and molecular biology*, 16(6), pp.735–743.

Coleman, T. et al., 1989. New fluorescence assay for the quantitation of fungi. *Journal of clinical microbiology*, 27(9), pp.2003–2007.

Collins, N.C. et al., 2003. SNARE-protein-mediated disease resistance at the plant cell wall. *Nature*, 425(6961), pp.973–977.

Consonni, C. et al., 2006. Conserved requirement for a plant host cell protein in powdery

mildew pathogenesis. *Nature Genetics*, 38(6), pp.716–720.

Constantinescu, O. & Fatehi, J., 2002. Peronospora-like fungi (Chromista, Peronosporales) parasitic on Brassicaceae and related hosts. *Nova Hedwigia*, 74(3), pp.291–338.

Cooper, A.J. et al., 2008. Basic compatibility of Albugo candida in Arabidopsis thaliana and Brassica juncea causes broad-spectrum suppression of innate immunity. *Molecular plant-microbe interactions : MPMI*, 21(6), pp.745–756.

Dangl, J.L. & Jones, J.D., 2001. Plant pathogens and integrated defence responses to infection. *Nature*, 411(6839), pp.826–833.

de Castro, E. et al., 2006. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Research*, 34(Web Server), pp.W362–W365.

de Jonge, R. & Thomma, B.P.H.J., 2009. Fungal LysM effectors: extinguishers of host immunity? *Trends in Microbiology*, 17(4), pp.151–157.

de Jonge, R. et al., 2010. Conserved Fungal LysM Effector Ecp6 Prevents Chitin-Triggered Immunity in Plants. *Science (New York, NY)*, 329(5994), pp.953–955.

de Jonge, R. et al., 2012. Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens uncovered by genome and RNA sequencing. *Proceedings of the National Academy of Sciences*, 109(13), pp.5110–5115.

Deshaies, R.J. & Joazeiro, C.A.P., 2009. RING Domain E3 Ubiquitin Ligases. *Annual Review of Biochemistry*, 78(1), pp.399–434.

Diagne, N. et al., 2011. Uvitex2B: a rapid and efficient stain for detection of arbuscular

mycorrhizal fungi within plant roots. *Mycorrhiza*.

Dodds, P.N. & Rathjen, J.P., 2010. Plant immunity: towards an integrated view of plant–pathogen interactions. *Nature reviews Genetics*, 11(8), pp.539–548.

Dodds, P.N. et al., 2006. Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proceedings of the National Academy of Sciences of the United States of America*, 103(23), pp.8888–8893.

Dong, S. et al., 2011. Phytophthora sojae Avirulence Effector Avr3b is a Secreted NADH and ADP-ribose Pyrophosphorylase that Modulates Plant Immunity D. Mackey, ed. *PLoS pathogens*, 7(11), p.e1002353.

Donofrio, N.M. & Delaney, T.P., 2001. Abnormal callose response phenotype and hypersusceptibility to Peronospoara parasitica in defence-compromised arabidopsis nim1-1 and salicylate hydroxylase-expressing plants. *Molecular plant-microbe interactions : MPMI*, 14(4), pp.439–450.

Dou, D. et al., 2008. RXLR-mediated entry of Phytophthora sojae effector Avr1b into soybean cells does not require pathogen-encoded machinery. *The Plant cell*, 20(7), pp.1930–1947.

Dunning, F.M. et al., 2007. Identification and mutational analysis of Arabidopsis FLS2 leucine-rich repeat domain residues that contribute to flagellin perception. *The Plant cell*, 19(10), pp.3297–3313.

Eddy, S.R., 2011. Accelerated Profile HMM Searches. *PLoS computational biology*, 7(10), p.e1002195.

Ellis, J.G. & Dodds, P.N., 2011. Showdown at the RXLR motif: Serious differences of opinion in how effector proteins from filamentous eukaryotic pathogens enter plant cells. *Proceedings of the National Academy of Sciences of the United States of America*, 108(35), pp.14381–14382.

Fabro, G. et al., 2011. Multiple Candidate Effectors from the Oomycete Pathogen Hyaloperonospora arabidopsidis Suppress Host Plant Immunity F. M. Ausubel, ed. *PLoS pathogens*, 7(11), p.e1002348.

Falk, A. et al., 1999. EDS1, an essential component of R gene-mediated disease resistance in Arabidopsis has homology to eukaryotic lipases. *Proceedings of the National Academy of Sciences of the United States of America*, 96(6), pp.3292–3297.

Feys, B.J. et al., 2005. Arabidopsis SENESCENCE-ASSOCIATED GENE101 stabilizes and signals within an ENHANCED DISEASE SUSCEPTIBILITY1 complex in plant innate immunity. *The Plant cell*, 17(9), pp.2601–2613.

Frey, N.F.D. & Robatzek, S., 2009. Trafficking vesicles: pro or contra pathogens? *Current Opinion in Plant Biology*, 12(4), pp.437–443.

Gan, P.H.P. et al., 2010. Lipid binding activities of flax rust AvrM and AvrL567 effectors. *Plant Signaling & Behavior*, 5(10), pp.1272–1275.

Gao, Z. et al., 2011. Plant intracellular innate immune receptor Resistance to Pseudomonas syringae pv. maculicola 1 (RPM1) is activated at, and functions on, the plasma membrane. *Proceedings of the National Academy of Sciences*, 108(18), pp.7619–7624.

García, A.V. & Parker, J.E., 2009. Heaven's Gate: nuclear accessibility and activities of plant immune regulators. *Trends in plant science*, 14(9), pp.479–487.

Gijzen, M. & Nürnberger, T., 2006. Nep1-like proteins from plant pathogens: recruitment

and diversification of the NPP1 domain across taxa. *Phytochemistry*, 67(16), pp.1800–1807.

Gilroy, E.M. et al., 2011. CMPG1-dependent cell death follows perception of diverse pathogen elicitors at the host plasma membrane and is suppressed by Phytophthora infestans RXLR effector AVR3a. *The New phytologist*.

Gimenez-Ibanez, S. et al., 2009. AvrPtoB targets the LysM receptor kinase CERK1 to promote bacterial virulence on plants. *Current biology : CB*, 19(5), pp.423–429.

Glazebrook, J., 2005. Contrasting Mechanisms of Defense Against Biotrophic and Necrotrophic Pathogens. *Annual review of phytopathology*, 43(1), pp.205–227.

Glazebrook, J., Rogers, E.E. & Ausubel, F.M., 1996. Isolation of Arabidopsis mutants with enhanced disease susceptibility by direct screening. *Genetics*, 143(2), pp.973–982.

Gómez-Gómez, L., Felix, G. & Boller, T., 1999. A single locus determines sensitivity to bacterial flagellin in Arabidopsis thaliana. *The Plant journal : for cell and molecular biology*, 18(3), pp.277–284.

Gómez-Gómez, L., Bauer, Z. & Boller, T., 2001. Both the extracellular leucine-rich repeat domain and the kinase activity of FSL2 are required for flagellin binding and signaling in Arabidopsis. *The Plant cell*, 13(5), pp.1155–1163.

González-Lamothe, R. et al., 2006. The U-box protein CMPG1 is required for efficient activation of defense mechanisms triggered by multiple resistance genes in tobacco and tomato. *The Plant cell*, 18(4), pp.1067–1083.

Goritschnig, S. et al., 2012. Computational Prediction and Molecular Characterization of an Oomycete Effector and the Cognate Arabidopsis Resistance Gene D. S. Guttman, ed. *PLoS genetics*, 8(2), p.e1002502.

Goritschnig, S., Zhang, Y. & Li, X., 2007. The ubiquitin pathway is required for innate immunity in Arabidopsis. *The Plant Journal*, 49(3), pp.540–551.

Grouffaud, S. et al., 2008. Plasmodium falciparum and Hyaloperonospora parasitica effector translocation motifs are functional in Phytophthora infestans. *Microbiology (Reading, England)*, 154(Pt 12), pp.3743–3751.

Göhre, V. et al., 2008. Plant Pattern-Recognition Receptor FLS2 Is Directed for Degradation by the Bacterial Ubiquitin Ligase AvrPtoB. *Current Biology*, 18(23), pp.1824–1832.

Gray, W.M. et al., 2001. Auxin regulates SCF(TIR1)-dependent degradation of AUX/IAA proteins. *Nature*, 414(6861), pp.271–276.

Gräslund, S. et al., 2008. Protein production and purification. *Nature Methods*, 5(2), pp.135–146.

Guerriero, G. et al., 2010. Chitin Synthases from Saprolegnia Are Involved in Tip Growth and Represent a Potential Target for Anti-Oomycete Drugs S. Kamoun, ed. *PLoS pathogens*, 6(8), p.e1001070.

Guo, Y.L. et al., 2011. Genome-Wide Comparison of Nucleotide-Binding Site-Leucine-Rich Repeat-Encoding Genes in Arabidopsis. *PLANT PHYSIOLOGY*, 157(2), pp.757–769.

Guttman, D.S. et al., 2002. A functional screen for the type III (Hrp) secretome of the plant pathogen Pseudomonas syringae. *Science (New York, NY)*, 295(5560), pp.1722–1726.

Haas, B.J. et al., 2009. Genome sequence and analysis of the Irish potato famine pathogen Phytophthora infestans. *Nature*, 461(7262), pp.393–398.

Hahn, M. et al., 1997. A putative amino acid transporter is specifically expressed in haustoria of the rust fungus Uromyces fabae. *Molecular plant-microbe interactions : MPMI*, 10(4), pp.438–445.

He, S.Y., Nomura, K. & Whittam, T.S., 2004. Type III protein secretion mechanism in mammalian and plant pathogens. *Biochimica et biophysica acta*, 1694(1-3), pp.181–206.

Heese, A. et al., 2007. The receptor-like kinase SERK3/BAK1 is a central regulator of innate immunity in plants. *Proceedings of the National Academy of Sciences of the United States of America*, 104(29), pp.12217–12222.

Hein, I. et al., 2009. The zig-zag-zig in oomycete-plant interactions. *Molecular plant pathology*, 10(4), pp.547–562.

Helgeson, J. et al., 1998. Somatic hybrids between Solanum bulbocastanum and potato: a new source of resistance to late blight. *TAG Theoretical and Applied Genetics*, 96(6), pp.738–742.

Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, Harrison T, et al. (2004) A host-targeting signal in virulence proteins reveals a secretome in malarial infection. Science 306: 1934–1937.

Hilton, J.L., Kearney, P.C. & Ames, B.N., 1965. Mode of action of the herbicide, 3-amino-1,2,4-triazole(amitrole): inhibition of an enzyme of histidine biosynthesis. *Archives of biochemistry and biophysics*, 112(3), pp.544–547.

Hinsch, M. & Staskawicz, B., 1996. Identification of a new Arabidopsis disease resistance locus, RPs4, and cloning of the corresponding avirulence gene, avrRps4, from Pseudomonas syringae pv. pisi. *Molecular plant-microbe interactions : MPMI*, 9(1),

pp.55–61.

Hofius, D. et al., 2009. Autophagic Components Contribute to Hypersensitive Cell Death in Arabidopsis. *Cell*, 137(4), pp.773–783.

Holub, E., 2006. Evolution of parasitic symbioses between plants and filamentous microorganisms. *Current Opinion in Plant Biology*, 9(4), pp.397–405.

Holub, E., 2008. Natural history of Arabidopsis thaliana and oomycete symbioses. *Eur J Plant Pathol*, 122, pp.91–109.

Holub, E., Beynon, J. & Crute, I., 1994. Phenotypic and genotypic characterization of interactions between isolates of Peronospora parasitica and accessions of Arabidopsis thaliana. *Molecular plant-microbe interactions*, 7(2), pp.223–239.

Holub, E.B. et al., 1995. Phenotypic and genotypic variation in the interaction between Arabidopsis thaliana and Albugo candida. *Molecular plant-microbe interactions : MPMI*, 8(6), pp.916–928.

Hwangbo, K. et al., 2010. Rapid and simple method for DNA extraction from plant and algal species suitable for PCR amplification using a chelating resin Chelex 100. *Plant Biotechnology Reports*, 4(1), pp.49–52.

IARC, 1999. *International Agency for Research on Cancer mono- graphs on the evaluation of carcinogenic risks to humans*,

Innes, R.W. et al., 1993. Molecular analysis of avirulence gene avrRpt2 and identification of a putative regulatory sequence common to all known Pseudomonas syringae avirulence genes. *Journal of bacteriology*, 175(15), pp.4859–4869.

Iizasa, E., Mitsutomi, M. & Nagano, Y., 2010. Direct Binding of a Plant LysM Receptor-

like Kinase, LysM RLK1/CERK1, to Chitin in Vitro. *Journal of Biological Chemistry*, 285(5), pp.2996–3004.

Jambunathan, N., Siani, J.M. & McNellis, T.W., 2001. A humidity-sensitive Arabidopsis copine mutant exhibits precocious cell death and increased disease resistance. *The Plant cell*, 13(10), pp.2225–2240.

Janjusevic, R., 2006. A Bacterial Inhibitor of Host Programmed Cell Death Defenses Is an E3 Ubiquitin Ligase. *Science (New York, NY)*, 311(5758), pp.222–226.

Jiang, R.H.Y. et al., 2008. RXLR effector reservoir in two Phytophthora species is dominated by a single rapidly evolving superfamily with more than 700 members. *Proceedings of the National Academy of Sciences*, 105(12), pp.4874–4879.

Jiang, R.H.Y. et al., 2008. RXLR effector reservoir in two Phytophthora species is dominated by a single rapidly evolving superfamily with more than 700 members. *Proceedings of the National Academy of Sciences*, 105(12), pp.4874–4879.

Jin, Q. et al., 2003. Type III protein secretion in Pseudomonas syringae. *Microbes and infection / Institut Pasteur*, 5(4), pp.301–310.

Jirage, D., T. L. Tootle, et al. (1999). "Arabidopsis thaliana PAD4 encodes a lipase- like gene that is important for salicylic acid signaling." Proc Natl Acad Sci U S A 96(23): 13583-8.

Jones, J.D.G. & Dangl, J.L., 2006. The plant immune system. *Nature*, 444(7117), pp.323–329.

Judelson, H.S., Tyler, B.M. & Michelmore, R.W., 1991. Transformation of the oomycete pathogen, Phytophthora infestans. *Molecular plant-microbe interactions : MPMI*, 4(6),

pp.602–607.

Kaku, H. et al., 2006. Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor. *Proceedings of the National Academy of Sciences of the United States of America*, 103(29), pp.11086–11091.

Kale, S.D. & Tyler, B.M., 2011. Entry of oomycete and fungal effectors into plant and animal host cells. *Cellular microbiology*, pp.no–no.

Kale, S.D. et al., 2010. External Lipid PI3P Mediates Entry of Eukaryotic Pathogen Effectors into Plant and Animal Host Cells. *Cell*, 142(2), pp.284–295.

Kamoun, S., 2006. A catalogue of the effector secretome of plant pathogenic oomycetes. *Annual review of phytopathology*, 44, pp.41–60.

Kamoun, S. et al., 1997. A gene encoding a protein elicitor of Phytophthora infestans is down-regulated during infection of potato. *Molecular plant-microbe interactions : MPMI*, 10(1), pp.13–20.

Kaschani, F. & van der Hoorn, R.A.L., 2011. A model of the C14-EPIC complex indicates hotspots for a protease-inhibitor arms race in the oomycete-potato interaction. *Plant Signaling & Behavior*, 6(1), pp.109–112.

Kaschani, F. et al., 2010. An Effector-Targeted Protease Contributes to Defense against Phytophthora infestans and Is under Diversifying Selection in Natural Hosts. *PLANT PHYSIOLOGY*, 154(4), pp.1794–1804.

Kelley, B.S. et al., 2010. A Secreted Effector Protein (SNE1) from Phytophthora infestans is a Broadly Acting Suppressor of Programmed Cell Death. *The Plant journal : for cell and molecular biology*.

Kemen, E. & Jones, J.D.G., 2012. Obligate biotroph parasitism: can we link genomes to lifestyles? *Trends in plant science*, 17(8), pp.448–457.

Kemen, E. et al., 2011. Gene Gain and Loss during Evolution of Obligate Parasitism in the White Rust Pathogen of Arabidopsis thaliana. *PLoS Biology*, 9(7), p.e1001094.

Kerscher, O., Felberbaum, R. & Hochstrasser, M., 2006. Modification of Proteins by Ubiquitin and Ubiquitin-Like Proteins. *Annual review of cell and developmental biology*, 22(1), pp.159–180.

Kim, H.C. & Huibregtse, J.M., 2009. Polyubiquitination by HECT E3s and the Determinants of Chain Type Specificity. *Molecular and cellular biology*, 29(12), pp.3307–3318.

Kim, H.C. et al., 2011. Structure and function of a HECT domain ubiquitin-binding site. *Nature Publishing Group*, 12(4), pp.334–341.

Kim, H.S., 2002. Arabidopsis SON1 Is an F-Box Protein That Regulates a Novel Induced Defense Response Independent of Both Salicylic Acid and Systemic Acquired Resistance. *THE PLANT CELL ONLINE*, 14(7), pp.1469–1482.

Knepper, C., Savory, E.A. & Day, B., 2011. Arabidopsis NDR1 Is an Integrin-Like Protein with a Role in Fluid Loss and Plasma Membrane-Cell Wall Adhesion. *PLANT PHYSIOLOGY*, 156(1), pp.286–300.

Koch, H. & Pimsler, M., 1987. Evaluation of Uvitex 2B: a nonspecific fluorescent stain for detecting and identifying fungi and algae in tissue. *Lab. Med*, 18, pp.603–606.

Koning-Ward, T.F. de et al., 2009. A newly discovered protein export machine in malaria parasites. *Nature*, 459(7249), pp.945–949.

Kosugi, S. et al., 2008. Six Classes of Nuclear Localization Signals Specific to Different Binding Grooves of Importin. *Journal of Biological Chemistry*, 284(1), pp.478–485.

Kraft, E., 2005. Genome Analysis and Functional Characterization of the E2 and RING-Type E3 Ligase Ubiquitination Enzymes of Arabidopsis. *PLANT PHYSIOLOGY*, 139(4), pp.1597–1611.

Krasileva, K.V. et al., 2011. Global Analysis of Arabidopsis/Downy Mildew Interactions Reveals Prevalence of Incomplete Resistance and Rapid Evolution of Pathogen Recognition D. J. Kliebenstein, ed. *PLoS ONE*, 6(12), p.e28765.

Krasileva, K.V., Dahlbeck, D. & Staskawicz, B.J., 2010. Activation of an Arabidopsis Resistance Protein Is Specified by the in Planta Association of Its Leucine-Rich Repeat Domain with the Cognate Oomycete Effector. *THE PLANT CELL ONLINE*, pp.1–16.

Kwon, C. et al., 2008. Co-option of a default secretory pathway for plant immune responses. *Nature*, 451(7180), pp.835–840.

Lam, E., Kato, N. & Lawton, M., 2001. Programmed cell death, mitochondria and the plant hypersensitive response. *Nature*, 411(6839), pp.848–853.

Lange, A. et al., 2006. Classical Nuclear Localization Signals: Definition, Function, and Interaction with Importin. *Journal of Biological Chemistry*, 282(8), pp.5101–5105.

Lawton, K.A. et al., 1996. Benzothiadiazole induces disease resistance in Arabidopsis by activation of the systemic acquired resistance signal transduction pathway. *The Plant journal : for cell and molecular biology*, 10(1), pp.71–82.

Lee, D.H. & Goldberg, A.L., 1998. Proteasome inhibitors: valuable new tools for cell

biologists. *Trends in cell biology*, 8(10), pp.397–403.

Lee, S.-W. et al., 2009. A type I-secreted, sulfated peptide triggers XA21-mediated innate immunity. *Science (New York, NY)*, 326(5954), pp.850–853.

Leonelli, L. et al., 2011. Structural elucidation and functional characterization of the Hyaloperonospora arabidopsidis effector protein ATR13. *PLoS pathogens*, 7(12), p.e1002428.

Levesque, C.A. et al., 2010. Genome sequence of the necrotrophic plant pathogen Pythium ultimum reveals original pathogenicity mechanisms and effector repertoire. *Genome biology*, 11(7), p.R73.

Lewis, J.D. et al., 2010. Allele-specific virulence attenuation of the Pseudomonas syringae HopZ1a type III effector via the Arabidopsis ZAR1 resistance protein. *PLoS genetics*, 6(4), p.e1000894.

Li, R. et al., 2010. The sequence and de novo assembly of the giant panda genome. *Nature*, 463(7279), pp.311–317.

Li, Y. et al., 2008. Control of final seed and organ size by the DA1 gene family in Arabidopsis thaliana. *Genes & development*, 22(10), pp.1331–1336.

Lin, D., Diao, J. & Zhou, D., 2011. Biochemical and Structural Studies of a HECT-like Ubiquitin Ligase from Escherichia coli O157: H7. *Journal of Biological Chemistry*.

Lin, D.Y.-W., Diao, J. & Chen, J., 2012. Crystal structures of two bacterial HECT-like E3 ligases in complex with a human E2 reveal atomic details of pathogen-host interactions. *Proceedings of the National Academy of Sciences*.

Lindeberg, M., Cunnac, S. & Collmer, A., 2012. Pseudomonas syringae type III effector repertoires: last words in endless arguments. *Trends in Microbiology*, 20(4), pp.199–208.

Links, M.G. et al., 2011. De novo sequence assembly of Albugo candida reveals a small genome relative to other biotrophic oomycetes. *BMC genomics*, 12(1), p.503.

Liu, J. et al., 2009. RIN4 functions with plasma membrane H+-ATPases to regulate stomatal apertures during pathogen attack. *PLoS Biology*, 7(6), p.e1000139.

Liu, T. et al., 2012. Chitin-induced dimerization activates a plant immune receptor. *Science (New York, NY)*, 336(6085), pp.1160–1164.

Loh YT, Zhou J, Martin GB. 1998. The myristylation motif of Pto is not required for disease resistance. Mol Plant Microbe Interact. 11(6):572-6.

Lu, D. et al., 2010. A receptor-like cytoplasmic kinase, BIK1, associates with a flagellin receptor complex to initiate plant innate immunity. *Proceedings of the National Academy of Sciences*, 107(1), pp.496–501.

Lu, D. et al., 2011. Direct ubiquitination of pattern recognition receptor FLS2 attenuates plant innate immunity. *Science (New York, NY)*, 332(6036), pp.1439–1442.

Lu, Y.-J. et al., 2012. Patterns of plant subcellular responses to successful oomycete infections reveal differences in host cell reprogramming and endocytic trafficking. *Cellular microbiology*, 14(5), pp.682–697.

Mackey D, Holt BF 3rd, Wiig A, Dangl JL., 2002. RIN4 interacts with Pseudomonas syringae type III effector molecules and is required for RPM1-mediated resistance in Arabidopsis. *Cell*. 108(6):743-54.

Mackey, D. et al., 2003. Arabidopsis RIN4 is a target of the type III virulence effector AvrRpt2 and modulates RPS2-mediated resistance. *Cell*, 112(3), pp.379–389.

Maeda, H. & Ishida, N., 1967. Specificity of binding of hexopyranosyl polysaccharides with fluorescent brightener. *Journal of biochemistry*, 62(2), pp.276–278.

Maekawa, T., Cheng, W., et al., 2011a. Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal functional module for triggering cell death. *Cell Host and Microbe*, 9(3), pp.187–199.

Maekawa, T., Kufer, T.A. & Schulze-Lefert, P., 2011b. NLR functions in plant and animal immune systems: so far and yet so close. *Nature Immunology*, 12(9), pp.817–826.

Maier, A.G. et al., 2008. Exported proteins required for virulence and rigidity of Plasmodium falciparum-infected human erythrocytes. *Cell*, 134(1), pp.48–61.

Manning, V.A. & Ciuffetti, L.M., 2005. Localization of Ptr ToxA Produced by Pyrenophora tritici-repentis Reveals Protein Import into Wheat Mesophyll Cells. *The Plant cell*, 17(11), pp.3203–3212.

McCann, H.C. et al., 2012. Identification of innate immunity elicitors using molecular signatures of natural selection. *Proceedings of the National Academy of Sciences*, 109(11), pp.4215–4220.

McDowell, J.M. et al., 2000. Downy mildew (Peronospora parasitica) resistance genes in Arabidopsis vary in functional requirements for NDR1, EDS1, NPR1 and salicylic acid accumulation. *The Plant journal : for cell and molecular biology*, 22(6), pp.523–529.

McDowell, J.M. et al., 2011. Propagation, storage, and assays with Hyaloperonospora

arabidopsidis: A model oomycete pathogen of Arabidopsis. *Methods in molecular biology (Clifton, NJ)*, 712, pp.137–151.

Mcleod, A. et al., 2008. Toward improvements of oomycete transformation protocols. *The Journal of eukaryotic microbiology*, 55(2), pp.103–109.

Mentlak, T.A. et al., 2012. Effector-mediated suppression of chitin-triggered immunity by magnaporthe oryzae is necessary for rice blast disease. *THE PLANT CELL ONLINE*, 24(1), pp.322–335.

Mittal, S. & Davis, K.R., 1995. Role of the phytotoxin coronatine in the infection of Arabidopsis thaliana by Pseudomonas syringae pv. tomato. *Molecular plant-microbe interactions : MPMI*, 8(1), pp.165–171.

Miya, A. et al., 2007. CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in Arabidopsis. *Proceedings of the National Academy of Sciences*, 104(49), pp.19613–19618.

Monaghan, J. & Zipfel, C., 2012. Plant pattern recognition receptor complexes at the plasma membrane. *Current Opinion in Plant Biology*, 15(4), pp.349–357.

Monaghan, J. et al., 2009. Two Prp19-Like U-Box Proteins in the MOS4-Associated Complex Play Redundant Roles in Plant Innate Immunity J. L. Dangl, ed. *PLoS pathogens*, 5(7), p.e1000526.

Moraes TF, Spreter T, Strynadka NC 2008. Piecing together the type III injectisome of bacterial pathogens. Curr Opin Struct Biol 18: 258–266.

Mucyn TS, et al. 2006. The tomato NBARC-LRR protein Prf interacts with Pto kinase in vivo to regulate specific plant immunity. Plant Cell. 18(10):2792-806.

Mukhtar, M.S. et al., 2011. Independently evolved virulence effectors converge onto hubs in a plant immune system network. *Science (New York, NY)*, 333(6042), pp.596–601.

Mutwil, M. et al., 2009. Assembly of an Interactive Correlation Network for the Arabidopsis Genome Using a Novel Heuristic Clustering Algorithm. *PLANT PHYSIOLOGY*, 152(1), pp.29–43.

Narusaka, M. et al., 2009. RRS1 and RPS4 provide a dual Resistance-gene system against fungal and bacterial pathogens. *The Plant journal : for cell and molecular biology*, 60(2), pp.218–226.

Nemri, A. et al., 2010. Genome-wide survey of Arabidopsis natural variation in downy mildew resistance using combined association and linkage mapping. *Proceedings of the National Academy of Sciences*, 107(22), pp.10302–10307.

Nguyen Ba, A.N. et al., 2009. NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction. *BMC Bioinformatics*, 10(1), p.202.

Oh, S.K. et al., 2009. In Planta Expression Screens of Phytophthora infestans RXLR Effectors Reveal Diverse Phenotypes, Including Activation of the Solanum bulbocastanum Disease Resistance Protein Rpi-blb2. *THE PLANT CELL ONLINE*, 21(9), pp.2928–2947.

Oliva, R. et al., 2010. Recent developments in effector biology of filamentous plant pathogens. *Cellular microbiology*, 12(6), pp.705–715.

Olmedo, G. et al., 2006. ETHYLENE-INSENSITIVE5 encodes a 5“-->3” exoribonuclease required for regulation of the EIN3-targeting F-box proteins EBF1/2. *Proceedings of the National Academy of Sciences of the United States of America*, 103(36), pp.13286–

13293.

Oppenheimer, D.G. et al., 1991. A myb gene required for leaf trichome differentiation in Arabidopsis is expressed in stipules. *Cell*, 67(3), pp.483–493.

Orbach, M.J. et al., 2000. A telomeric avirulence gene determines efficacy for the rice blast resistance gene Pi-ta. *The Plant cell*, 12(11), pp.2019–2032.

Palma, K. et al., 2010. Autoimmunity in Arabidopsis acd11 is mediated by epigenetic regulation of an immune receptor. *PLoS pathogens*, 6(10), p.e1001137.

Palma, K., Zhang, Y. & Li, X., 2005. An importin alpha homolog, MOS6, plays an important role in plant innate immunity. *Current biology : CB*, 15(12), pp.1129–1135.

Parker, J.E. et al., 1996. Characterization of eds1, a mutation in Arabidopsis suppressing resistance to Peronospora parasitica specified by several different RPP genes. *The Plant cell*, 8(11), pp.2033–2046.

Parker, J.E. et al., 1997. The Arabidopsis downy mildew resistance gene RPP5 shares similarity to the toll and interleukin-1 receptors with N and L6. *The Plant cell*, 9(6), pp.879–894.

Petersen, M. et al., 2000. Arabidopsis map kinase 4 negatively regulates systemic acquired resistance. *Cell*, 103(7), pp.1111–1120.

Piscatelli, H. et al., 2011. The EHEC Type III Effector NleL Is an E3 Ubiquitin Ligase That Modulates Pedestal Formation F. G. van der Goot, ed. *PLoS ONE*, 6(4), p.e19331.

Punta, M. et al., 2011. The Pfam protein families database. *Nucleic Acids Research*,

40(D1), pp.D290–D301.

Quinn, P.J., 2002. Plasma membrane phospholipid asymmetry. *Sub-cellular biochemistry*, 36, pp.39–60.

Qutob, D. et al., 2006. Phytotoxicity and Innate Immune Responses Induced by Nep1-Like Proteins. *THE PLANT CELL ONLINE*, 18(12), pp.3721–3744.

Raffaele, S. & Kamoun, S., 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. *Nature reviews Microbiology*, 10(6), pp.417–430.

Rafiqi, M. et al., 2010. Internalization of flax rust avirulence proteins into flax and tobacco cells can occur in the absence of the pathogen. *The Plant cell*, 22(6), pp.2017–2032.

Rehmany, A. et al., 2005. Differential recognition of highly divergent downy mildew avirulence gene alleles by RPP1 resistance genes from two Arabidopsis lines. *THE PLANT CELL ONLINE*, 17(6), p.1839.

Reignault, P. et al., 1996. Four Arabidopsis RPP loci controlling resistance to the Noco2 isolate of Peronospora parasitica map to regions known to contain other RPP recognition specificities. *Molecular plant-microbe interactions : MPMI*, 9(6), pp.464–473.

Rietman, H., 2011. Putting the Phytophthora infestans genome sequence at work; multiple novel avirulence and potato resistance gene candidates revealed. *Wagnigen PhD thesis*, pp.1–169.

Rivas, S., 2011. A plethora of virulence strategies hidden behind nuclear targeting of microbial effectors. pp.1–20.

Robert-Seilaniantz, A., Grant, M. & Jones, J.D.G., 2011a. Hormone Crosstalk in Plant Disease and Defense: More Than Just JASMONATE-SALICYLATE Antagonism. *Annual review of phytopathology*, 49(1), pp.317–343.

Robert-Seilaniantz, A., Maclean, D., et al., 2011b. The microRNA miR393 redirects secondary metabolite biosynthesis away from camalexin and towards glucosinolates. *The Plant journal : for cell and molecular biology*.

Rosebrock TR, et al. 2007. A bacterial E3 ubiquitin ligase targets a host protein kinase to disrupt plant immunity. Nature. 448(7151):370-4.

Roux, M. et al., 2011. The Arabidopsis Leucine-Rich Repeat Receptor-Like Kinases BAK1/SERK3 and BKK1/SERK4 Are Required for Innate Immunity to Hemibiotrophic and Biotrophic Pathogens. *THE PLANT CELL ONLINE*, 23(6), pp.2440–2455.

Sainsbury, F., Thuenemann, E.C. & Lomonossoff, G.P., 2009. pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnology Journal*, 7(7), pp.682–693.

Santner, A. & Estelle, M., 2010. The ubiquitin-proteasome system regulates plant hormone signaling. *The Plant Journal*, 61(6), pp.1029–1040.

Schneider, C.A., Rasband, W.S. & Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7), pp.671–675.

Schornack, S. et al., 2010. Ancient class of translocated oomycete effectors targets the host nucleus. *Proceedings of the National Academy of Sciences of the United States of America*, 107(40), pp.17421–17426.

Schornack, S. et al., 2009. Ten things to know about oomycete effectors. *Molecular plant pathology*, 10(6), pp.795–803.

Schwessinger, B. et al., 2011. Phosphorylation-dependent differential regulation of plant growth, cell death, and innate immunity by the regulatory receptor-like kinase BAK1. *PLoS genetics*, 7(4), p.e1002046.

Schulze-Lefert, P. & Panstruga, R., 2011. A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. *Trends in plant science*, 16(3), pp.117–125.

Segonzac, C. & Zipfel, C., 2011. Activation of plant pattern-recognition receptors by bacteria. *Current Opinion in Microbiology*, 14(1), pp.54–61.

Segonzac, C. et al., 2011. Hierarchy and roles of PAMP-induced responses in Nicotiana benthamiana. *PLANT PHYSIOLOGY*.

Shan, L. et al., 2008. Bacterial effectors target the common signaling partner BAK1 to disrupt multiple MAMP receptor-signaling complexes and impede plant immunity. *Cell Host and Microbe*, 4(1), pp.17–27.

Shan, L. et al., 2000. The Pseudomonas AvrPto protein is differentially recognized by tomato and tobacco and is localized to the plant plasma membrane. *THE PLANT CELL ONLINE*, 12(12), p.2323.

Shan, W. et al., 2004. The Avr1b locus of Phytophthora sojae encodes an elicitor and a regulator required for avirulence on soybean plants carrying resistance gene Rps1b. *Molecular plant-microbe interactions : MPMI*, 17(4), pp.394–403.

Shang, Y. et al. 2006. RAR1, a central player in plant immunity, is targeted by Pseudomonas syringae effector AvrB. Proc Natl Acad Sci U S A. 103(50):19200-5.

Shirano, Y., 2002. A Gain-of-Function Mutation in an Arabidopsis Toll Interleukin1 Receptor-Nucleotide Binding Site-Leucine-Rich Repeat Type R Gene Triggers Defense Responses and Results in Enhanced Disease Resistance. *THE PLANT CELL ONLINE*, 14(12), pp.3149–3162.

Slusarenko, A.J. & Schlaich, N.L., 2003. Downy mildew of Arabidopsis thaliana caused by Hyaloperonospora parasitica (formerly Peronospora parasitica). *Molecular plant pathology*, 4(3), pp.159–170.

Sohn, K. et al., 2007a. The downy mildew effector proteins ATR1 and ATR13 promote disease susceptibility in Arabidopsis thaliana. *THE PLANT CELL ONLINE*, 19(12), p.4077.

Sohn, K.H. et al., 2011. HopAS1 recognition significantly contributes to Arabidopsis nonhost resistance to Pseudomonas syringae pathogens. *New Phytologist*, 193(1), pp.58–66.

Sohn, K.H. et al., 2007b. The downy mildew effector proteins ATR1 and ATR13 promote disease susceptibility in Arabidopsis thaliana. *The Plant cell*, 19(12), pp.4077–4090.

Sohn, K.H., Zhang, Y. & Jones, J.D.G., 2009. The Pseudomonas syringae effector protein, AvrRPS4, requires in planta processing and the KRVY domain to function. *The Plant journal : for cell and molecular biology*, 57(6), pp.1079–1091.

Sommer, D.D. et al., 2007. Minimus: a fast, lightweight genome assembler. *BMC Bioinformatics*, 8(1), p.64.

Song, Jing et al., 2009. Apoplastic effectors secreted by two unrelated eukaryotic plant pathogens target the tomato defense protease Rcr3. *Proceedings of the National Academy of Sciences*, 106(5), pp.1654–1659.

Song, Junqi et al., 2003. Gene RB cloned from Solanum bulbocastanum confers broad spectrum resistance to potato late blight. *Proceedings of the National Academy of Sciences of the United States of America*, 100(16), pp.9128–9133.

Song, W.Y. et al., 1995. A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21. *Science (New York, NY)*, 270(5243), pp.1804–1806.

Soylu, S., 2004. Ultrastructural characterisation of the host-pathogen interface in white blister-infected Arabidopsis leaves. *Mycopathologia*, 158(4), pp.457–464.

Soylu, S. et al., 2003. Ultrastructural characterisation of interactions between Arabidopsis thaliana and Albugo candida. *Physiological and Molecular Plant Pathology*, 63(4), pp.201–211.

Spanu, P.D., 2012. The Genomics of Obligate (and Nonobligate) Biotrophs. *Annual review of phytopathology*.

Spielmann, T. & Gilberger, T.-W., 2010. Protein export in malaria parasites: do multiple export motifs add up to multiple export pathways? *Trends in Parasitology*, 26(1), pp.6–10.

Spoel, S.H. & Dong, X., 2012. How do plants achieve immunity? Defence without specialized immune cells. *Nature Reviews Immunology*, 12(2), pp.89–100.

Stassen, J.H.M. & Van den Ackerveken, G., 2011. How do oomycete effectors interfere with plant life? *Current Opinion in Plant Biology*, 14(4), pp.407–414.

Stavrinides, J., Ma, W. & Guttman, D.S., 2006. Terminal reassortment drives the quantum evolution of type III effectors in bacterial pathogens. *PLoS pathogens*, 2(10), p.e104.

Stukenbrock, E.H. & McDonald, B.A., 2009. Population genetics of fungal and oomycete effectors involved in gene-for-gene interactions. *Molecular plant-microbe interactions : MPMI*, 22(4), pp.371–380.

Stuttmann, J. et al., 2011. Perturbation of Arabidopsis Amino Acid Metabolism Causes Incompatibility with the Adapted Biotrophic Pathogen Hyaloperonospora arabidopsidis. *THE PLANT CELL ONLINE*, 23(7), pp.2788–2803.

Swiderski, M.R., Birker, D. & Jones, J.D.G., 2009. The TIR domain of TIR-NB-LRR resistance proteins is a signaling domain involved in cell death induction. *Molecular plant-microbe interactions : MPMI*, 22(2), pp.157–165.

Tameling, W.I.L. et al., 2002. The tomato R gene products I-2 and MI-1 are functional ATP binding proteins with ATPase activity. *The Plant cell*, 14(11), pp.2929–2939.

Tao, H. et al., 2010. Purifying natively folded proteins from inclusion bodies using sarkosyl, Triton X-100, and CHAPS. *BioTechniques*, 48(1), pp.61–64.

Thines, M. & Kamoun, S., 2010. Oomycete-plant coevolution: recent advances and future prospects. *Current Opinion in Plant Biology*.

Thines, M. et al., 2009. A new species of Albugo parasitic to Arabidopsis thaliana reveals new evolutionary patterns in white blister rusts (Albuginaceae).

Tian, M. et al., 2011. 454 Genome sequencing of Pseudoperonospora cubensis reveals effector proteins with a QXLR translocation motif. *Molecular plant-microbe interactions : MPMI*, 24(5), pp.543–553.

Tian, M. et al., 2007. A Phytophthora infestans cystatin-like protein targets a novel tomato papain-like apoplastic protease. *PLANT PHYSIOLOGY*, 143(1), pp.364–377.

Torto, T.A. et al., 2003. EST mining and functional expression assays identify extracellular effector proteins from the plant pathogen Phytophthora. *Genome research*, 13(7), pp.1675–1685.

Trujillo, M. et al., 2008. Negative Regulation of PAMP-Triggered Immunity by an E3 Ubiquitin Ligase Triplet in Arabidopsis. *Current Biology*, 18(18), pp.1396–1401.

Tsuda, K. et al., 2009. Network properties of robust immunity in plants. *PLoS genetics*, 5(12), p.e1000772.

Tyler, B.M. et al., 2006. Phytophthora genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science (New York, NY)*, 313(5791), pp.1261–1266.

Tzfira, T., Vaidya, M. & Citovsky, V., 2004. Involvement of targeted proteolysis in plant genetic transformation by Agrobacterium. *Nature*, 431(7004), pp.87–92.

Van Criekinge, W. & Beyaert, R., 1999. Yeast Two-Hybrid: State of the Art. *Biological procedures online*, 2, pp.1–38.

van Damme, M. et al., 2005. Identification of arabidopsis loci required for susceptibility to the downy mildew pathogen Hyaloperonospora parasitica. *Molecular plant-microbe interactions : MPMI*, 18(6), pp.583–592.

van Damme, M. et al., 2012. The Irish Potato Famine Pathogen Phytophthora infestans Translocates the CRN8 Kinase into Host Plant Cells B. J. Howlett, ed. *PLoS pathogens*, 8(8), p.e1002875.

van den Burg, H.A. et al., 2006. Cladosporium fulvum Avr4 protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection. *Molecular plant-microbe interactions : MPMI*, 19(12), pp.1420–1430.

van der Hoorn RA, Kamoun S. 2008. From Guard to Decoy: a new model for perception of plant pathogen effectors. Plant Cell. 20(8):2009-17

van West, P. et al., 2010. The putative RxLR effector protein SpHtp1 from the fish pathogenic oomycete Saprolegnia parasitica is translocated into fish cells. *FEMS Microbiology Letters*, 310(2), pp.127–137.

Verdecia, M.A. et al., 2003. Conformational flexibility underlies ubiquitin ligation mediated by the WWP1 HECT domain E3 ligase. *Molecular cell*, 11(1), pp.249–259.

Vierstra, R.D., 2009. The ubiquitin–26S proteasome system at the nexus of plant biology. *Nature reviews Molecular cell biology*, 10(6), pp.385–397.

Vleeshouwers, V.G.A.A. et al., 2011. Understanding and exploiting late blight resistance in the age of effectors. *Annual review of phytopathology*, 49, pp.507–531.

Voegele, R.T. & Mendgen, K.W., 2011. Nutrient uptake in rust fungi: how sweet is parasitic life? *Euphytica*, 179(1), pp.41–55.

Voegele, R.T. et al., 2006. Cloning and characterization of a novel invertase from the obligate biotroph Uromyces fabae and analysis of expression patterns of host and pathogen invertases in the course of infection. *Molecular plant-microbe interactions : MPMI*, 19(6), pp.625–634.

Voegele, R.T. et al., 2001. The role of haustoria in sugar supply during infection of broad bean by the rust fungus Uromyces fabae. *Proceedings of the National Academy of*

*Sciences of the United States of America*, 98(14), pp.8133–8138.

Vogel, J. & Somerville, S., 2000. Isolation and characterization of powdery mildew-resistant Arabidopsis mutants. *Proceedings of the National Academy of Sciences of the United States of America*, 97(4), pp.1897–1902.

Vogel, J.P., 2002. PMR6, a Pectate Lyase-Like Gene Required for Powdery Mildew Susceptibility in Arabidopsis. *THE PLANT CELL ONLINE*, 14(9), pp.2095–2106.

Vogel, J.P. et al., 2004. Mutations in PMR5 result in powdery mildew resistance and altered cell wall composition. *The Plant Journal*, 40(6), pp.968–978.

Voges, D., Zwickl, P. & Baumeister, W., 1999. The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annual Review of Biochemistry*, 68, pp.1015–1068.

Voglmayr, H. & Riethmüller, A., 2006. Phylogenetic relationships of Albugo species (white blister rusts) based on LSU rDNA sequence and oospore data. *Mycological research*, 110(Pt 1), pp.75–85.

Wachsmuth, E.D., 1988. A comparison of the highly selective fluorescence staining of fungi in tissue sections with Uvitex 2B and Calcofluor White M2R. *The Histochemical journal*, 20(4), pp.215–221.

Walls, D. & Loughran, S.T., 2011. Tagging recombinant proteins to enhance solubility and aid purification. *Methods in molecular biology (Clifton, NJ)*, 681, pp.151–175.

Wang, Q. et al., 2011. Transcriptional Programming and Functional Interactions within the Phytophthora sojae RXLR Effector Repertoire. *THE PLANT CELL ONLINE*, pp.1–24.

Wawra, S., Agacan, M., et al., 2012a. The avirulence protein 3a (AVR3a) from the potato pathogen Phytophthora infestans, forms homodimers through its predicted translocation region and does not specifically bind phospholipids. *Journal of Biological Chemistry*.

Wawra, S., Bain, J., et al., 2012b. Host-targeting protein 1 (SpHtp1) from the oomycete Saprolegnia parasitica translocates specifically into fish cells in a tyrosine-O-sulphate-dependent manner. *Proceedings of the National Academy of Sciences*.

Werner, S. et al., 2002. Chitin synthesis during in planta growth and asexual propagation of the cellulosic oomycete and obligate biotrophic grapevine pathogen Plasmopara viticola. *FEMS Microbiology Letters*, 208(2), pp.169–173.

Whisson, S.C. et al., 2007. A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature*, 450(7166), pp.115–118.

Wiermer, M., Feys, B.J. & Parker, J.E., 2005. Plant immunity: the EDS1 regulatory node. *Current Opinion in Plant Biology*, 8(4), pp.383–389.

Wiermer, M. et al., 2012. Putative members of the Arabidopsis Nup107-160 nuclear pore sub-complex contribute to pathogen defense. *The Plant Journal*, 70(5), pp.796–808.

Williams, S.J. et al., 2011. An autoactive mutant of the M flax rust resistance protein has a preference for binding ATP, whereas wild-type M protein binds ADP. *Molecular plant-microbe interactions : MPMI*, 24(8), pp.897–906.

Willmann, R. et al., 2011. Arabidopsis lysin-motif proteins LYM1 LYM3 CERK1 mediate bacterial peptidoglycan sensing and immunity to bacterial infection. *Proceedings of the National Academy of Sciences*.

Win, J. et al., 2007. Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. *The Plant cell*, 19(8), pp.2349–2369.

Wirthmueller, L. et al., 2007. Nuclear accumulation of the Arabidopsis immune receptor RPS4 is necessary for triggering EDS1-dependent defense. *Current Biology*, 17(23), pp.2023–2029.

Xiang, T. et al., 2008. Pseudomonas syringae effector AvrPto blocks innate immunity by targeting receptor kinases. *Current Biology*, 18(1), pp.74–80.

Xin, D.-W. et al., 2012. Functional Analysis of NopM, a Novel E3 Ubiquitin Ligase (NEL) Domain Effector of Rhizobium sp. Strain NGR234 J. Chang, ed. *PLoS pathogens*, 8(5), p.e1002707.

Yaeno, T. et al., 2011. Phosphatidylinositol monophosphate-binding interface in the oomycete RXLR effector AVR3a is required for its stability in host cells to modulate plant immunity. *Proceedings of the National Academy of Sciences of the United States of America*.

Yang, Z., 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Computer applications in the biosciences : CABIOS*, 13(5), pp.555–556.

Yoshida, A. et al., 2012. Reduction of glucose uptake through inhibition of hexose transporters and enhancement of their endocytosis by methylglyoxal in Saccharomyces cerevisiae. *Multiple values selected*, 287(1), pp.701–711.

Zerbino, D.R. & Birney, E., 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome research*, 18(5), pp.821–829.

Zhang, Y, 2005. A Putative Nucleoporin 96 Is Required for Both Basal Defense and

Constitutive Resistance Responses Mediated by suppressor of npr1-1,constitutive 1. *THE PLANT CELL ONLINE*, 17(4), pp.1306–1316.

Zhang, Ying et al., 2006. The inflammation-associated Salmonella SopA is a HECT-like E3 ubiquitin ligase. *Molecular microbiology*, 62(3), pp.786–793.

Zhang, Yuelin et al., 2003. A gain-of-function mutation in a plant disease resistance gene leads to constitutive activation of downstream signal transduction pathways in suppressor of npr1-1, constitutive 1. *The Plant cell*, 15(11), pp.2636–2646.

Zhang, Ziguo et al., 2007. A SNARE-protein has opposing functions in penetration resistance and defence signalling pathways. *The Plant journal : for cell and molecular biology*, 49(2), pp.302–312.

Zheng, X.-Y. et al., 2012. Coronatine Promotes Pseudomonas syringae Virulence in Plants by Activating a Signaling Cascade that Inhibits Salicylic Acid Accumulation. *Cell Host and Microbe*, 11(6), pp.587–596.

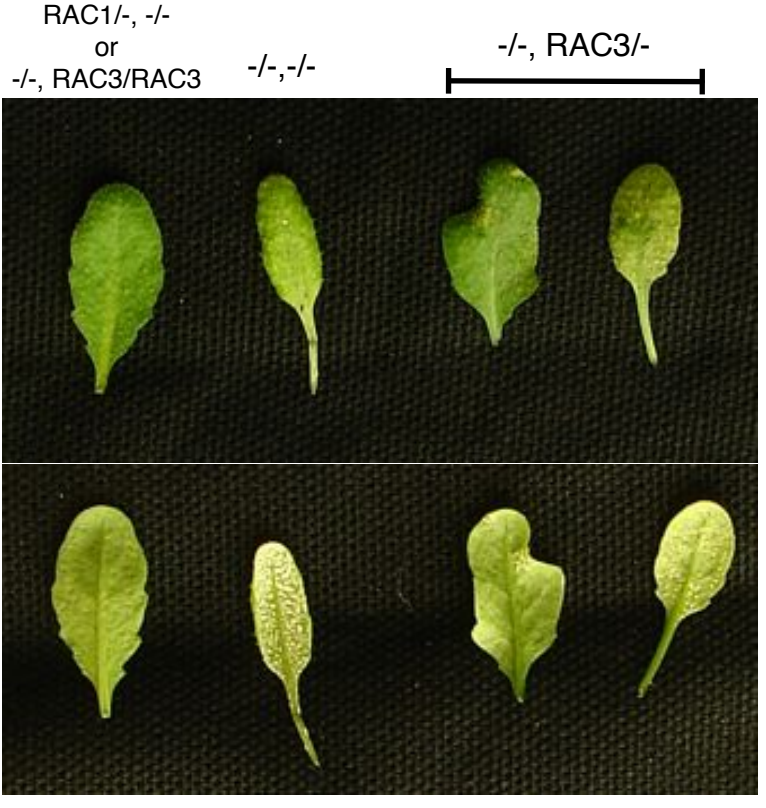
Zhou, N. et al., 1998. PAD4 functions upstream from salicylic acid to control defense responses in Arabidopsis. *The Plant cell*, 10(6), pp.1021–1030.

Zipfel, C., 2008. Pattern-recognition receptors in plant innate immunity. *Current opinion in immunology*, 20(1), pp.10–16.

Zipfel, C. et al., 2004. Bacterial disease resistance in Arabidopsis through flagellin perception. *Nature*, 428(6984), pp.764–767.

Zipfel, C. et al., 2006. Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts Agrobacterium-mediated transformation. *Cell*, 125(4), pp.749–760.

# 10. Appendix




**Supplemental figure S1: *AlNc14* is recognised by *RAC1* and *RAC3*.** F2 segregants from a Ksk-1 x Col-0 cross were infected with *AlNc14.* Pictures were taken 7 dpi. Note that RAC3 heterozygotes are partially resistant to *AlNc14*. '-' denote Col-0 genotype at locus.

**Supplemental table S6.1. Resistance to *AlNc2* is dependent on *EDS1* in Col-0.** 4 week old plants were infected with *AcNc2* and scored 14 dpi. Refer to figure 6.1 for specifics.

| Lab Code | Locus | Alt Name | insertion line | Interaction |
|---|---|---|---|---|
| 23 | At1g56510 | WRR4 | SALK_148037C | R |
| 24 | At1g56520 | | SALK_111589C | R |
| 25 | At1g56540 | | SAIL_205_B06 | R |
| 45 | At1g63730 | | SALK_087810C | R |
| 47 | At1g63750 | | SALK_022493C | R |
| 48 | At1g63860 | | SALK_033050C | R |
| 50 | At1g63880 | RLM1 | SALK_110393 | R |
| 51 | At1g64070 | RLM1 | SALK_042846C | R |
| Col-eds1 | At3g48090 | EDS1 | RNAi construct | S |

| Organism | Proteome downloaded from | File | Source |
|---|---|---|---|
| Arabidopsis lyrata | ftp://ftp.ensemblgenomes.org/pub/plants/release-15/fasta/arabidopsis_lyrata/pep/ | Arabidopsis_lyrata.v.1.0.15.pep.abinitio.fa | EBI - Ensemble Plants release 15 |
| Arabidopsis thaliana | ftp://ftp.ensemblgenomes.org/pub/plants/release-15/fasta/arabidopsis_thaliana/pep/ | Arabidopsis_thaliana.TAIR10.15.pep.all.fa | EBI - Ensemble Plants release 15 |
| Chlamydomonas reinhardtii | ftp://ftp.ensemblgenomes.org/pub/plants/release-15/fasta/chlamydomonas_reinhardtii/pep/ | Chlamydomonas_reinhardtii.v3.0.15.pep.all.fa | EBI - Ensemble Plants release 15 |
| *Physcomitrella patens* | ftp://ftp.ensemblgenomes.org/pub/plants/release-15/fasta/physcomitrella_patens/pep/ | Physcomitrella_patens.ASM242v1.15.pep.all.fa | EBI - Ensemble Plants release 15 |
| **Fungi** | | | |
| *Puccinia graminis* | ftp://ftp.ensemblgenomes.org/pub/fungi/release-15/fasta/puccinia_graminis/pep/ | Puccinia_graminis.ASM14992v1.15.pep.all.fa | EBI - Ensemble Fungi release 15 |
| *Magnaporthe oryzae* | ftp://ftp.ensemblgenomes.org/pub/fungi/release-15/fasta/magnaporthe_oryzae/pep/ | Magnaporthe_oryzae.MG8.15.pep.all.fa | EBI - Ensemble Fungi release 15 |
| *Saccharomyces cerevisiae* | ftp://ftp.ensemblgenomes.org/pub/fungi/release-15/fasta/saccharomyces_cerevisiae/pep/ | Saccharomyces_cerevisiae.SacCer_Apr2011.15.pep.all.fa | EBI - Ensemble Fungi release 15 |
| *Schizosaccharomyces pombe* | ftp://ftp.ensemblgenomes.org/pub/fungi/release-15/fasta/schizosaccharomyces_pombe/pep/ | Schizosaccharomyces_pombe.ASM294v1.15.pep.all.fa | EBI - Ensemble Fungi release 15 |
| *Ustilago maydis* | ftp://ftp.ensemblgenomes.org/pub/fungi/release-15/fasta/ustilago_maydis/pep/ | Ustilago_maydis.UM1.15.pep.all.fa | EBI - Ensemble Fungi release 15 |
| ***Protists*** | | | |
| *Dictyostelium discoideum* | ftp://ftp.ensemblgenomes.org/pub/protists/release-15/fasta/dictyostelium_discoideum/pep/ | Dictyostelium_discoideum.dictybase.01.15.pep.all.fa | EBI - Ensemble Protist release 15 |
| *Thalassiosira pseudonana* | ftp://ftp.ensemblgenomes.org/pub/protists/release-15/fasta/thalassiosira_pseudonana/pep/ | Thalassiosira_pseudonana.ASM14940v1.15.pep.all.fa | EBI - Ensemble Protist release 15 |
| *Toxoplasma gondii* | ftp://ftp.ensemblgenomes.org/pub/protists/release-15/fasta/toxoplasma_gondii/pep/ | Toxoplasma_gondii.JCVI-tgg-v1.0.15.pep.all.fa | EBI - Ensemble Protist release 15 |
| *Plasmodium falciparum* | ftp://ftp.ensemblgenomes.org/pub/protists/release-15/fasta/plasmodium_falciparum/pep/ | Plasmodium_falciparum.ASM276v1.15.pep.all.fa | EBI - Ensemble Protist release 15 |
| ***Protists - oomycetes*** | | | |
| *Albugo laibachii* | ftp://ftp.ensemblgenomes.org/pub/protists/release-15/fasta/albugo_laibachii/pep/ | Albugo_laibachii.ENA1.15.pep.all.fa.gz | EBI - Ensemble Protist release 15 |
| *Albugo candida Nc2* | In house genome | AcNc2_prot.aa2012.fasta | Anastasia Gardiner unpublished |
| *Hyaloperonospora arabidopsidis* | In house genome | hpa_v8.3_transcript_3.faa | Naveed Ishaque unpublished |
| *Phytophthora infestans* | ftp://ftp.ensemblgenomes.org/pub/protists/release-15/fasta/phytophthora_infestans/pep/ | Phytophthora_infestans.ASM14294v1.15.pep.all.fa | EBI - Ensemble Protist release 15 |
| *Phytophthora ramorum* | ftp://ftp.ensemblgenomes.org/pub/protists/release-15/fasta/phytophthora_ramorum/pep/ | Phytophthora_ramorum.ASM14973v1.15.pep.all.fa | EBI - Ensemble Protist release 15 |
| *Phytophthora sojae* | ftp://ftp.ensemblgenomes.org/pub/protists/release-15/fasta/phytophthora_sojae/pep/ | Phytophthora_sojae.ASM14975v1.15.pep.all.fa | EBI - Ensemble Protist release 15 |
| *Phytophthora capsici* | http://genome.jgi-psf.org/Phyca11/download/ | Phyca11_filtered_proteins.fasta | JGI - v11.10 |
| *Pythium ultimum* | ftp://ftp.ensemblgenomes.org/pub/protists/release-15/fasta/pythium_ultimum/pep/ | Pythium_ultimum.pug.15.pep.all.fa | EBI - Ensemble Protist release 15 |
| *saprolegnia parasitica* | http://www.broadinstitute.org/annotation/genome/Saprolegnia_parasitica/MultiHome.html | saprolegnia_parasitica_cbs_223.65_1_proteins.fasta | EBI - Ensemble Protist release 15 |

OPEN ACCESS Freely available online



# Gene Gain and Loss during Evolution of Obligate Parasitism in the White Rust Pathogen of *Arabidopsis thaliana*

**Eric Kemen[1], Anastasia Gardiner[1], Torsten Schultz-Larsen[1], Ariane C. Kemen[1], Alexi L. Balmuth[1,2], Alexandre Robert-Seilaniantz[1], Kate Bailey[1], Eric Holub[3], David J. Studholme[4], Dan MacLean[1], Jonathan D. G. Jones[1]***

**1** The Sainsbury Laboratory, Norwich Research Park, Norwich, United Kingdom, **2** The GenePool, The University of Edinburgh, Edinburgh, United Kingdom, **3** School of Life Sciences, University of Warwick, Wellesbourne Campus, United Kingdom, **4** School of Biosciences, University of Exeter, Exeter, United Kingdom

## Abstract

Biotrophic eukaryotic plant pathogens require a living host for their growth and form an intimate haustorial interface with parasitized cells. Evolution to biotrophy occurred independently in fungal rusts and powdery mildews, and in oomycete white rusts and downy mildews. Biotroph evolution and molecular mechanisms of biotrophy are poorly understood. It has been proposed, but not shown, that obligate biotrophy results from (i) reduced selection for maintenance of biosynthetic pathways and (ii) gain of mechanisms to evade host recognition or suppress host defence. Here we use Illumina sequencing to define the genome, transcriptome, and gene models for the obligate biotroph oomycete and *Arabidopsis* parasite, *Albugo laibachii*. *A. laibachii* is a member of the Chromalveolata, which incorporates Heterokonts (containing the oomycetes), Apicomplexa (which includes human parasites like *Plasmodium falciparum* and *Toxoplasma gondii*), and four other taxa. From comparisons with other oomycete plant pathogens and other chromalveolates, we reveal independent loss of molybdenum-cofactor-requiring enzymes in downy mildews, white rusts, and the malaria parasite *P. falciparum*. Biotrophy also requires "effectors" to suppress host defence; we reveal RXLR and Crinkler effectors shared with other oomycetes, and also discover and verify a novel class of effectors, the "CHXCs", by showing effector delivery and effector functionality. Our findings suggest that evolution to progressively more intimate association between host and parasite results in reduced selection for retention of certain biosynthetic pathways, and particularly reduced selection for retention of molybdopterin-requiring biosynthetic pathways. These mechanisms are not only relevant to plant pathogenic oomycetes but also to human pathogens within the Chromalveolata.

**Citation:** Kemen E, Gardiner A, Schultz-Larsen T, Kemen AC, Balmuth AL, et al. (2011) Gene Gain and Loss during Evolution of Obligate Parasitism in the White Rust Pathogen of *Arabidopsis thaliana*. PLoS Biol 9(7): e1001094. doi:10.1371/journal.pbio.1001094

**Academic Editor:** Frederick M. Ausubel, Massachusetts General Hospital, Harvard Medical School, United States of America

**Received** October 27, 2010; **Accepted** May 10, 2011; **Published** July 5, 2011



**Funding:** The authors thank European Research Council Advanced Investigator grant 233376 (ALBUGON), Gatsby Foundation, BBSRC, DFG (KE 1509/1-1), DASTI (Danish Agency for Science, Technology and Innovation) for funding. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

**Abbreviations:** KOG, core eukaryotic orthologous group; LTR, long terminal repeat; Nc14, Norwich 14; Pst, *Pseudomonas syringae* pv. tomato; TM, transmembrane.

* E-mail: jonathan.jones@tsl.ac.uk

## Introduction

For more than 150 years, attempts to culture downy mildews, powdery mildews, and rusts on artificial nutrient media have been unsuccessful. The terms obligate parasitism and obligate biotrophy are used to denote organisms that live in such an obligatory association with living hosts [1,2]. Recent research on the obligate biotroph powdery mildew fungus *Blumeria graminis* or downy mildew oomycete *Hyaloperonospora arabidopsidis* reveals a close correlation between the biotrophic life style and massive gene losses in primary and secondary metabolism [3,4]. Obligate biotrophs form an intimate haustorial interface with parasitized cells. Haustoria are differentiated intercellular hyphae, but little is known about their functionality and evolution beyond their involvement in nutrient uptake [5,6].

The obligate biotroph oomycete *Albugo laibachii* is a member of the Chromalveolata, which incorporates Dinophyta, Ciliophora, Heterokonts (containing the oomycetes), Haptophyta, Cryptophyta, and Apicomplexa (which includes human parasites like *Plasmodium falciparum* and *Toxoplasma gondii* [7,8]).

Within the oomycetes, *A. laibachii* belongs to a lineage known as peronosporalean, which includes the hemibiotrophic pathogen of potato *Phytophthora infestans* [9] and the necrotroph pathogen *Pythium ultimum* [10]. Within this lineage, obligate biotrophy evolved twice independently in white blister rusts (Albuginales) and downy mildews (part of the Peronosporaceae) [11]. The downy mildew pathogen *H. arabidopsidis* and *A. laibachii* are both pathogens of the model plant *Arabidopsis thaliana* [12]. While both show similar infection structures within the host [13,14], *A. laibachii* releases motile zoospores from asexual spores and sexual oospores, while *H. arabidopsidis* lacks all motile stages [4,15]. Both pathogens are regularly found to co-infect plants and sporulate on the same leaf [16].

## Author Summary

Plant pathogens that cannot grow except on their hosts are called obligate biotrophs. How such biotrophy evolves is poorly understood. In this study, we sequenced the genome of the obligate biotroph white rust pathogen (*Albugo laibachii*, Oomycota) of *Arabidopsis*. From comparisons with other oomycete plant pathogens, diatoms, and the human pathogen *Plasmodium falciparum*, we reveal a loss of important metabolic enzymes. We also reveal the appearance of defence-suppressing "effectors", some carrying motifs known from other oomycete effectors, and discover and experimentally verify a novel class of effectors that share a CHXC motif within 50 amino acids of the signal peptide cleavage site. Obligate biotrophy involves an intimate association within host cells at the haustorial interface (where the parasite penetrates the host cell's cell wall), where nutrients are acquired from the host and effectors are delivered to the host. We found that *A. laibachii*, like *Hyaloperonospora arabidopsidis* and *Plasmodium falciparum*, lacks molybdopterin-requiring biosynthetic pathways, suggesting relaxed selection for retention of, or even selection against, this pathway. We propose that when defence suppression becomes sufficiently effective, hosts become such a reliable source of nutrients that a free-living phase can be lost. These mechanisms leading to obligate biotrophy and host specificity are relevant not only to plant pathogenic oomycetes but also to human pathogens.


A remarkable consequence of infection by *Albugo* sp. is enhanced host plant susceptibility to other parasites to which the host is resistant in the absence of *Albugo* infection, and also impairment of cell death mechanisms [16]. *Albugo* sp. infect 63 genera and 241 species [17], including economically important *Brassica rapa* (canola), *B. juncea* (oilseed mustard), and *B. oleracea* (cabbage family vegetables) [18,19]. Recent analysis of oomycete evolutionary history [11] suggest that *Albugo* is more closely related to necrotrophs such as *Pythium* than to downy mildews, and thus provides a unique system to study the evolution and consequences of biotrophy, and to identify new defence-suppressing effectors and their host targets.

## Results/Discussion

### *A. laibachii* Isolates

Since prolonged culture of pathogen strains can result in genetic changes [20], we sequenced a fresh highly virulent field isolate of *A. laibachii*. The strain was selected from a heavily infected *Ar. thaliana* field plot (Norwich, United Kingdom) [21], and strains were single zoospore purified. Isolate Norwich 14 (Nc14) was determined as *A. laibachii* [19] and used for further analyses. In contrast to Nc14, *A. laibachii* isolate Em1 (formerly Acem1, *A. candida* East Malling 1 [19]) is an established *Albugo* strain that was collected 15 y ago [16,22,23], and we resequenced this strain. Both strains show identical ITS (internal transcribed spacer of ribosomal RNAs) and COX2 (cytochrome C oxidase subunit II) sequences. To ensure that sequence differences observed between these strains are of biological relevance not just the result of background mutations, we tested the host range for both isolates on 126 *Ar. thaliana* accessions and identified 12 that show resistance to only one of the *A. laibachii* isolates (Table S1). Nc14 is virulent on more accessions than the Em1 isolate is (Table 1).

### Illumina Genome Sequencing, Assembly, and Quality Assessment

The *A. laibachii* Nc14 genome was sequenced using Illumina 76-bp paired reads with ~240-fold coverage (Figure 1). In order to assemble the diploid heterozygous genome, an assembly pipeline was developed using Velvet [24] as primary assembler and Minimus [25] as meta-assembler (Figure S1). Short read assembly programs are sensitive to heterozygous positions depending on read depth and kmer-length. Reads not aligning to bacterial or plant sequence in public databases were used to estimate the genome size as ~37 Mbp. Using the estimated genome size, 50% of the resulting assembly is contained in 164 contigs with an N50 of 56.5 kbp. A comparative analysis of contig size classes versus frequency indicates that 90% of the assembled genome shows a high degree of continuity in only 585 contigs, while 10% of the genome is fragmented in 3,231 contigs (Figure 2A). Read depth indicates that this 10% of the genome shows elevated levels of nucleotide coverage that are likely to comprise unresolved repeats (Figure 2B). Aligning Illumina cDNA reads from different stages of infection to reveal transcriptionally active regions in the assembly shows that few transcripts arise from the unresolved repetitive regions of the genome (Figure 2D), suggesting that the gene space of a genome can be reliably defined using Illumina-only approaches. A CEGMA [26] analysis revealed a high degree of completeness of assembly of core eukaryotic genes, as well as a continuity within the core genes comparable to high-quality Sanger read assemblies (Figure S2; Table S2). We designed 32 primer pairs for regions between 0.6 and 5 kb based on our assembly (Table S3). Thirty-one genomic regions could be amplified and were Sanger sequenced from both ends. All PCR products had the predicted size, and sequences showed 100% identity to the genome assembly.

The mitochondrial draft genome was assembled in a separate attempt because of its high repeat content and therefore higher coverage compared to the core genome. The assembled genome comprises 26.7 kb in 11 contigs and shows a high degree of synteny to the *P. infestans* mitochondrion Ia [27] and the *Py. ultimum* mitochondrion [10] (Figure S3). Considering the node coverage of the Velvet primary assembly (~150×), 15.6 kb of the mitochondrial genome have >300× node coverage and seem to be duplicated. This might indicate, comparable to the *Py. ultimum*

**Table 1.** Percent of *Ar. thaliana* ecotypes resistant to *A. laibachii* Em1 and Nc14 isolates.

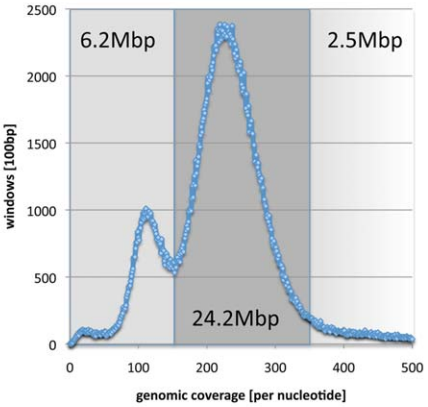
| *A. laibachii* Isolate Tested | Percent Resistant *Ar. thaliana* Accessions | | |
|---|---|---|---|
| | Per Each of the *A. laibachii* Isolates | To Both Isolates | Specifically to Only One of the Isolates |
| Em1 | 14.3 | 7.1 | 7.1 |
| Nc14 | 9.5 | | 2.4 |

Results indicate that the fresh isolate Nc14 is more virulent than Em1, which has been cultivated and propagated in the lab for more than 15 y.
doi:10.1371/journal.pbio.1001094.t001

A

| Nc14 ID / lane | pair | insert | GA | read length | all reads | aligned reads | % aligned |
|---|---|---|---|---|---|---|---|
| 49 / 1 | 1 | 800 | GA1 | 36 | 2592936 | 2423971 | 93.5 |
| 49 / 1 | 2 | 800 | GA1 | 36 | 2604589 | 2013293 | 77.3 |
| 49 / 2 | 1 | 400 | GA1 | 36 | 4452796 | 4353397 | 97.8 |
| 49 / 2 | 2 | 400 | GA1 | 36 | 4451946 | 4265672 | 95.8 |
| 49 / 3 | 1 | 400 | GA1 | 36 | 4639124 | 4530895 | 97.7 |
| 49 / 3 | 2 | 400 | GA1 | 36 | 4638463 | 4428583 | 95.5 |
| 49 / 8 | 1 | 800 | GA1 | 36 | 3219683 | 3008718 | 93.4 |
| 49 / 8 | 2 | 800 | GA1 | 36 | 3234913 | 2506759 | 77.5 |
| 58 / 3 | 1 | 800 | GA2 | 36 | 5266521 | 4889778 | 92.8 |
| 58 / 3 | 2 | 800 | GA2 | 36 | 5282233 | 4177837 | 79.1 |
| 58 / 5 | 1 | 800 | GA2 | 36 | 5938133 | 5525515 | 93.1 |
| 58 /5 | 2 | 800 | GA2 | 36 | 5954389 | 4760042 | 79.9 |
| 58 / 6 | 1 | 800 | GA2 | 36 | 5957734 | 5545776 | 93.1 |
| 58 / 6 | 2 | 800 | GA2 | 36 | 5972827 | 4767272 | 79.8 |
| 58 / 7 | 1 | 400 | GA2 | 36 | 8141242 | 7944121 | 97.6 |
| 58 / 7 | 2 | 400 | GA2 | 36 | 8137936 | 7824385 | 96.1 |
| 58 / 8 | 1 | 400 | GA2 | 36 | 7531910 | 7345696 | 97.5 |
| 58 / 8 | 2 | 400 | GA2 | 36 | 7529511 | 7234670 | 96.1 |
| 95 / 5 | 1 | 400 | GA2 | 76 | 12536014 | 12128361 | 96.7 |
| 95 / 5 | 2 | 400 | GA2 | 76 | 12537817 | 11975089 | 95.5 |
| 95 / 6 | 1 | 400 | GA2 | 76 | 12142358 | 11751747 | 96.8 |
| 95 / 6 | 2 | 400 | GA2 | 76 | 12143720 | 11610040 | 95.6 |
| 95 / 7 | 1 | 800 | GA2 | 76 | 8139569 | 7424231 | 91.2 |
| 95 / 7 | 2 | 800 | GA2 | 76 | 8159867 | 6320200 | 77.5 |
| 95 / 8 | 1 | 800 | GA2 | 76 | 7296923 | 6636963 | 91.0 |
| 95 / 8 | 2 | 800 | GA2 | 76 | 7314825 | 5603981 | 76.6 |
| all | | | | | 175817979 | 160996992 | 91.6 |

B




C

| Em1 ID / lane | pair | insert | GA | read length | all reads | aligned reads | % aligned |
|---|---|---|---|---|---|---|---|
| 82 / 1 | 1 | 400 | GA2 | 36 | 9092140 | 7034977 | 77.4 |
| 82 / 1 | 2 | 400 | GA2 | 36 | 9092140 | 6217935 | 68.4 |
| 82 / 2 | 1 | 800 | GA2 | 36 | 7523758 | 5742848 | 76.3 |
| 82 / 2 | 2 | 800 | GA2 | 36 | 7523758 | 5685920 | 75.6 |
| 83 / 1 | 1 | 400 | GA2 | 36 | 7776094 | 6074451 | 78.1 |
| 83 / 1 | 2 | 400 | GA2 | 36 | 7776094 | 5110984 | 65.7 |
| 83 / 2 | 1 | 800 | GA2 | 36 | 5065239 | 3899228 | 77.0 |
| 83 / 2 | 2 | 800 | GA2 | 36 | 5065239 | 3645968 | 72.0 |
| all | | | | | 58914462 | 43412311 | 73.7 |

D

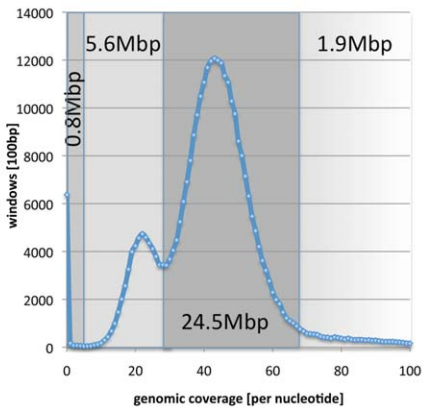



**Figure 1. Genomic sequencing data and coverage of *A. laibachii* Nc14 and Em1 assemblies.** (A) Reads generated for *A. laibachii* Nc14 using Illumina genome analyzer version 1 (GA1) or version 2 (GA2). (B) Distribution of genomic coverage. Grey fields indicate the total amount of sequence represented by the 100-bp windows with corresponding coverage. (C) Reads generated for *A. laibachii* Em1. (D) Distribution of genomic coverage showing Em1 reads aligned to the Nc14 genome using MAQ aligner. Nc14 and Em1 show a major peak at 226× and 43× coverage, respectively. A second peak is detected at 112× or 22×, showing half the coverage of the main peak, indicating highly heterozygous regions that were not merged in the assembly or hemizygous regions.
doi:10.1371/journal.pbio.1001094.g001

mitochondrion genome [10], that ~50% of the genome is duplicated, leading to an estimated genome size of ~43 kb. While the highly repetitive tRNAs are not resolved within the *A. laibachii* mitochondrial genome, regions of high synteny between the *Py. ultimum* and the *P. infestans* mitochondrial genome are found in ribosomal proteins and subunits of the NADH dehydrogenase as well as cytochrome C oxidase.

## Features of the *A. laibachii* Nuclear Genome

Approximately 22% of the *A. laibachii* Nc14 genome assembly consists of repetitive regions (Figure 3; Tables S4 and S5). The majority of repeats are represented by transposable elements (96%), while 4% of all repeats are *A. laibachii*-specific (Table S5). Compared to other obligate biotrophs, the number of repeats is low. *H. arabidopsidis*, for example, with an estimated genome size of 100 Mb, contains ~43.3% repeats [4], while transposable elements account for 64% of the ~120-Mb *Bl. graminis* (powdery mildew) genome [3]. We identified 45 contigs carrying telomeric repeats; amongst these, 25 contigs have telomeric repeats located at one end of a contig. We therefore postulate that the *A. laibachii* Nc14 genome is distributed over 12 or 13 chromosomes (Table S6). tRNA genes are difficult to resolve because of their high copy number [28]. Within our Illumina assembly, 153 tRNA genes were detected with 48 distinct anticodons (Figure S4; Table S7). Our ability to resolve all these repeats within the Illumina short read assembly illustrates its quality.

Based on read depth, both Nc14 and Em1 isolates possess ~6 Mbp of hemizygous or highly heterozygous regions (6.2 and 5.6 Mbp for Nc14 and Em1, respectively) (Figure 1B and 1D) as well as ~13,000 heterozygous loci (13,116 and 13,523 for Nc14 and Em1, respectively) (Figure 2C). Remarkably, most of the hemizygous/highly heterozygous regions are shared between Nc14 and Em1.

Compared to other sequenced oomycetes like *P. infestans* (240 Mbp), *H. arabidopsidis* (100 Mbp), or even *Py. ultimum* (42.8 Mbp), *A. laibachii* has a highly compact genome structure (Figure 4A). Approximately 50% of the *A. laibachii* genome assembly matched cDNA reads, and transcriptionally active regions are further clustered, resulting in transcriptional hot spots and silent genomic regions (Figure 4B).

## Annotation and Validation of Protein Coding Genes

A reference set of 13,032 gene models was generated incorporating cDNA reads from different stages of infection (Figure S5A). From extensive cDNA sequencing of infected *Arabidopsis* leaves, approximately 20 M (~1.5 Gbp) unique Illumina reads match the Nc14 genome assembly but not *Ar. thaliana* TAIR 9.0, and these were used to generate training sets for *ab initio* gene predictions and as evidence sets for consensus gene

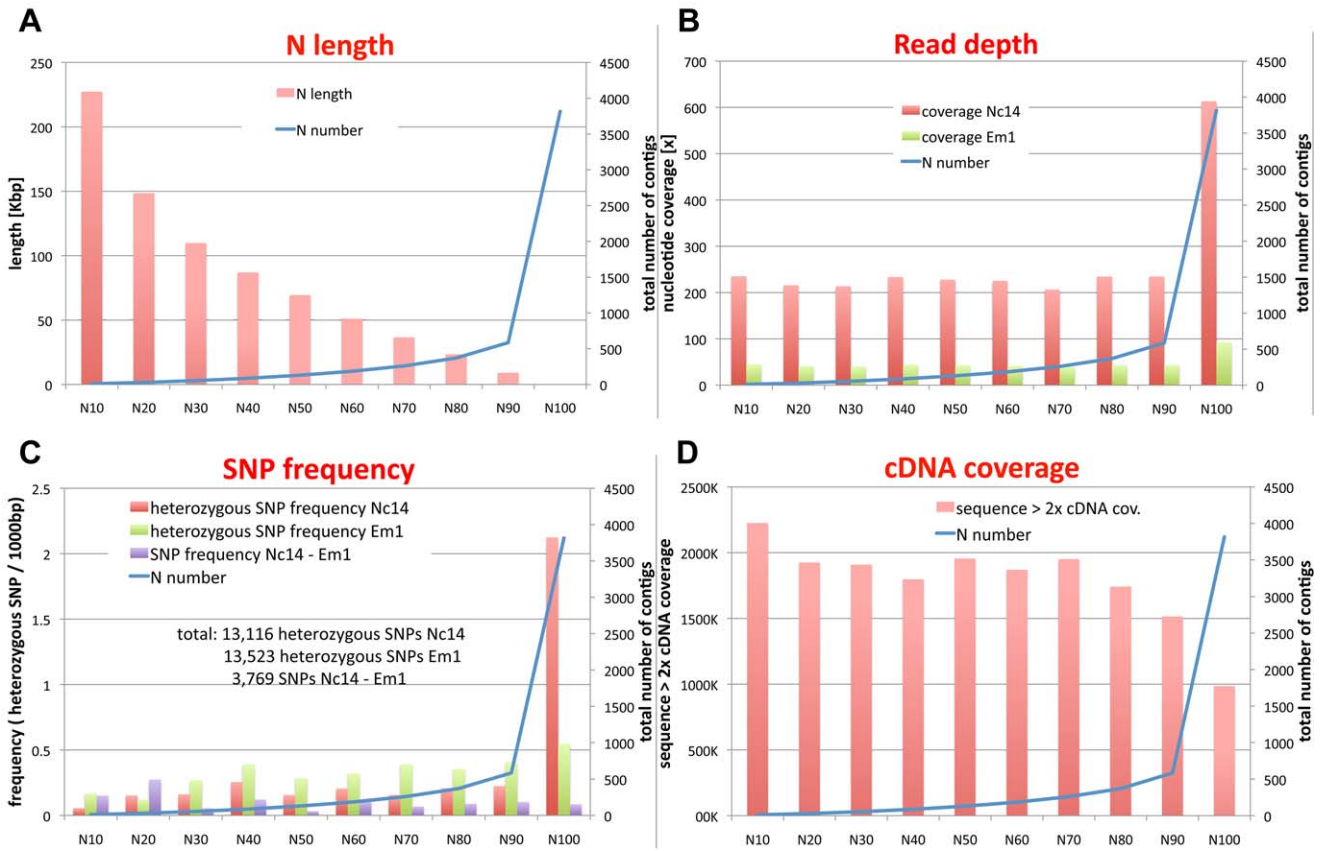



**Figure 2. Distribution of contig length, nucleotide coverage, SNP frequency, and cDNA coverage in the *A. laibachii* assembly.** (A) Genomic distribution of contig length (N length) versus contig number (N number). N lengths were calculated by ordering all sequences according to their length and then adding the length from longest to shortest until the summed length exceeded 10% (N10), 20% (N20), etc., up to 100% (N100) of the assembled contigs (32.7 Mbp). Blotting the N length versus the N number (number of contigs in each N category) indicates that 90% of the assembled genome show high continuity, while the last 10% are highly fragmented. (B) Average coverage for each category for Nc14 (red) and Em1 (green). In all, 90% of the genome shows low variation, consistent with 210–240× coverage for Nc14 and 40–50× for Em1. The last 10% show highly elevated coverage, indicating unresolved highly repetitive regions present in Nc14 and Em1. (C) Distribution of heterozygosity in each N category shows elevated levels in the set of short contigs. Heterozygous positions were accepted only if coverage was >180× and <350× for Nc14 (red) or >27× and <80× for Em1 (green). SNPs between Nc14 and Em1 were calculated ignoring heterozygous positions (lilac). (D) Alignment of Nc14 cDNA and summing up all regions showing >2× coverage indicate that the more continuous part of the genome contains more transcribed regions than the highly repetitive regions of the genome (in the histogram, N length and N number are cumulative while read depth, SNP frequency, and cDNA coverage are presented as binned data).
doi:10.1371/journal.pbio.1001094.g002

prediction. In all, 88.3% of all gene models are supported by at least three cDNA hits.

For validation of these gene models, a set of 860 annotated core eukaryotic orthologous groups (KOGs) [29] was compiled and tested. In all, 75% of these groups are present in the current annotation. For comparison, 78% of KOGs were present in *P. infestans*, 73% in *H. arabidopsidis*, 42% in *Pl. falciparum*, and 85% in *Ar. thaliana* (Figure S5B). In addition, 49.9% of all gene models show Pfam support, resulting in 2,505 Pfam domains, and 803 genes were functionally assigned to pathways using ASGARD [30] and manual annotation. Transcriptional units show an even more compact, clustered occurrence than *P. sojae* or *P. ramorum* and an occurrence pattern clearly different from that of *P. infestans* [9] (Figure 4C).

From our annotations using ASGARD we identified major enzymes of the lipopolysaccharide biosynthesis pathway, as have been described for *P. infestans* [31]. These analyses revealed, in addition, the possibility that *A. laibachii* is able to synthesize brassinosteroids. We identified potential homologues to the *Ar. thaliana* brassinosteroid biosynthesis genes *Dwf4* and *DET2* (Table S8). Although ASGARD identified homologues of *Br6ox*, *D2*, and *CPD*, manual annotation revealed that assigning function to members of the superfamily of cytochrome P450 enzymes in *A. laibachii* is difficult based on homology alone (Table S8). It has been hypothesized that the frequency of functionally redundant genes is reduced in obligate biotrophs, as reported for *Bl. graminis* [3]. Combining ASGARD and manual annotation we identified the absence of the whole steroid biosynthesis pathway, and, like other oomycetes, *A. laibachii* probably relies on the host as a source of sterols. We hypothesize that *A. laibachii* would need to take up campesterol from the plant as a precursor for brassinosteroid synthesis.

## Ancestral Red and Green Algae Genes in the *A. laibachii* Genome

During evolution, plastids of both red algae and green algae were transferred to other lineages by secondary endosymbiosis. How often and when secondary endosymbiosis occurred is difficult to address but of importance to clarify the origin of chromalveolates and their gain and loss of endosymbionts. There are two distinct hypotheses for what took place. The monophyletic hypothesis posits that a red alga was taken up only once, followed

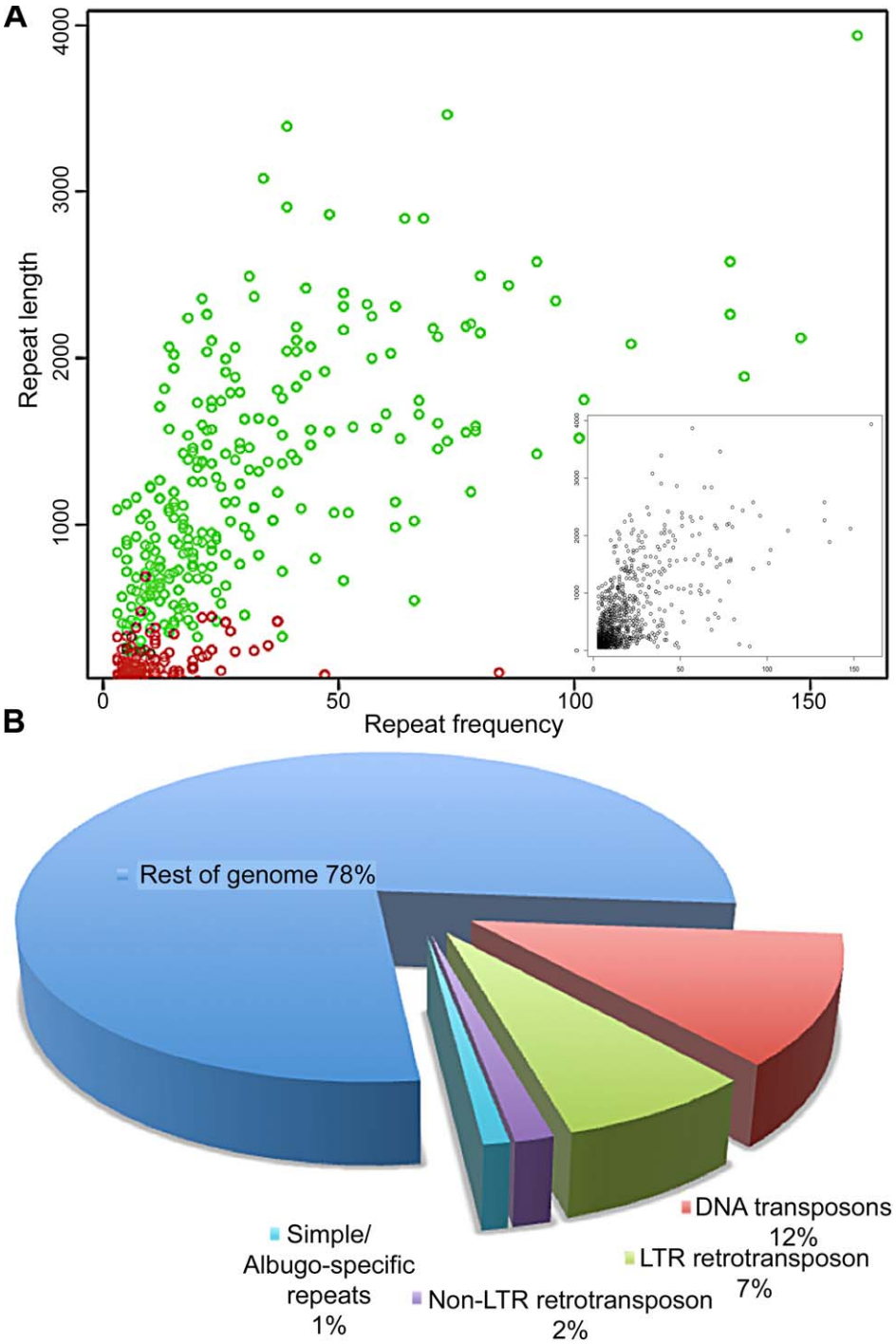



**Figure 3. Repeats identified in the *A. laibachii* Nc14 contigs.** Initial run of RepeatScout produced a library of 1,252 consensus repetitive sequences that include transposable elements, recently duplicated paralogous genes, and other dispersed duplicated regions. (A) Inset: The distribution of lengths of the identified repeats versus their frequency in the genome is shown. The majority of repeats fell into the category of short and rare in the assembly. The primary plot in (A) shows that the majority of the longest and most frequent repeats in the genome are transposon elements (shown in green and Table 1), while *Albugo*-specific repetitive sequences are mostly short (shown in red). (B) Summary of the proportion of the repetitive sequences (percent) in the *A. laibachii* Nc14 genome.
doi:10.1371/journal.pbio.1001094.g003

by repeated losses of this algal genome, giving rise to the highly divergent group of chromalveolates [32]. An alternative and more common view hypothesizes polyphyletic origins of the Chromalveolata, with in some cases multiple events of secondary endosymbiosis [33–35].

Molecular divergence of *A. laibachii* from other species within the Chromalveolata was assessed by examining the percentage of amino acid identity between orthologous gene pairs (Figure 5). These analyses demonstrate that the green alga *Chlamydomonas reinhardtii*, the brown alga *Ectocarpus siliculosus*, and the diatom

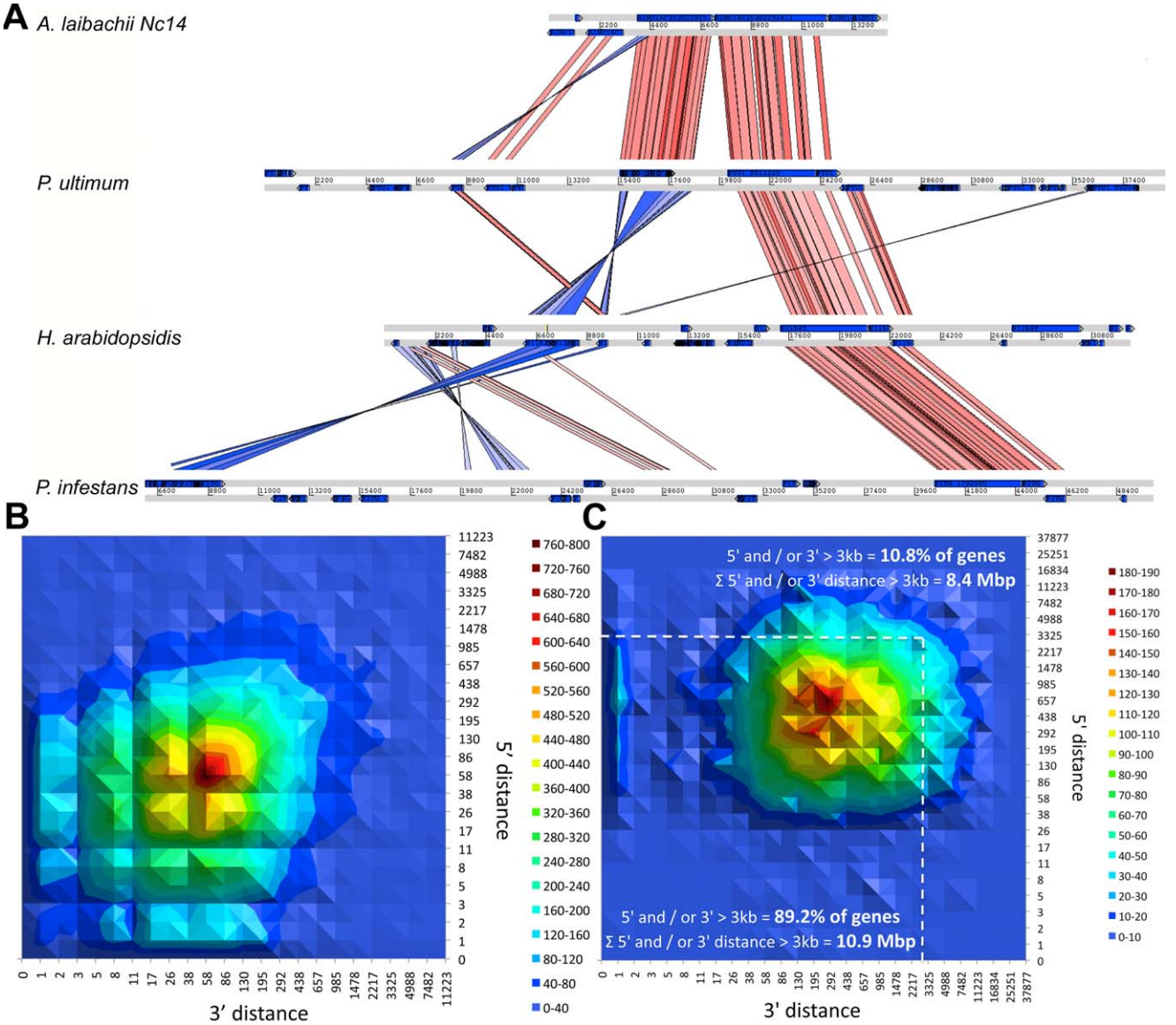



**Figure 4. *A. laibachii* has a compact genome with expression clusters.** (A) Synteny between *A. laibachii*, *Py. ultimum*, *H. arabidopsidis*, and *P. infestans*. The region shown is an example of the dense clustering of genes in the pentafunctional AROM polypeptide and a P-type ATPase. The AROM polypeptide comprises five enzymes of the shikimate pathway in one enzyme. With increasing genome size the distance between both genes increases and re-organisations occur (red, synteny without inversion; blue, inverted regions). (B) Plotting the distance between transcriptional islands based on the 5′ to 3′ orientation of the forward strand reveals that transcriptional regions are clustered close together. The maximum peak reflects the average intron size. Regions with no 3′ but with 5′ distance and vice versa reflect overlapping 3′ and 5′ non-coding regions of genes. Analysing the distance distribution between transcriptional units reveals a median distance between genes of 45 bp, showing that within transcribed regions, nearly all the DNA sequence corresponds to genes. (C) Plotting the 5′–3′ distance for all genes from ATG to stop to the next gene confirms the gene clustering. Only 10.8% of all genes have a distance to the next gene or the end of the contig greater than 3 kb. Summing the distance between these genes contributes to only 10.9 Mbp of the genome because of the close packaging, while summing the distance of the few genes that are not in clusters contributes to 8.4 Mbp of the genome.
doi:10.1371/journal.pbio.1001094.g004

*Phaeodactylum tricornutum* show the same distribution of percentage amino acid identity to *A. laibachii* Nc14 regarding the cumulative frequency of orthologous pairs. In contrast, previous systematic analyses suggested that brown algae and diatoms are the closest relatives of oomycetes and that secondary endosymbiosis occurred with a red alga [32], although there are suggestions that oomycetes diverged before this event [36]. Using a set of >1,700 genes that are of "green" origin (from green algae) or "red" origin (from red algae) and that have been integrated into the diatom nuclear genome [37], we found more oomycete genes that show significant BLAST hits to green algae than to red algae (34 "green" compared to five "red") (Figure S6; Table S9). These findings are consistent with the results published by Moustafa et al. [37] for diatoms. In a separate approach we identified genes showing high similarity between oomycetes, green algae, and red algae that are absent from diatoms (32 "green"; 11 "red") (Tables S10 and S11). This result might indicate the presence of all these genes in a common ancestor, followed by loss or expansion of the gene family

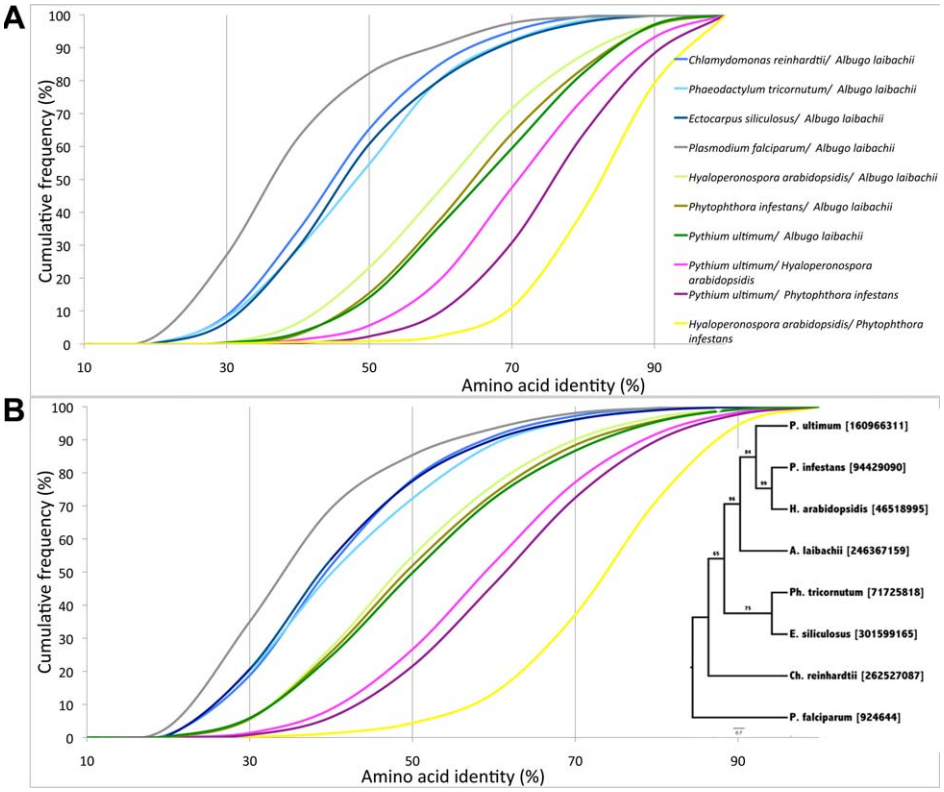



**Figure 5. Molecular divergence between *A. laibachii* and other species based on pairwise comparisons.** (A) Molecular divergence based on all pairwise comparisons of the one-to-one orthologues. In the figure, the cumulative frequencies of amino acid identity across each set of potential orthologous pairs is presented, indicating that although *H. arabidopsidis* and *A. laibachii* are both biotrophs, *H. arabidopsidis* is less diverged from *P. infestans* than it is from *A. laibachii* (e.g., in the *H. arabidopsidis*–*A. laibachii* comparison, ~22% of all orthologues show an amino acid identity of <50%, while only ~14% in a *Py. ultimum*–*A. laibachii* comparison show an amino acid identity of <50%). *A. laibachii* shows the highest amino acid identity to *Py. ultimum*. (B) Molecular divergence between *A. laibachii* and other species based on the subset of core eukaryotic genes to show stability of the test. Results are consistent with the one-to-one orthologue analyses although differences between *A. laibachii*, *P. infestans*, *H. arabidopsidis*, and *Py. ultimum* are less obvious, indicating the lack of selection pressure on the core eukaryotic genes [37]. For comparative reasons, a tree using ITS2 sequences is added. The represented tree is a maximum likelihood tree produced with PhyML.
doi:10.1371/journal.pbio.1001094.g005

depending on adopted live style. To address this question, we further analysed genes absent from *A. laibachii* Nc14 and studied their presence/absence in three other oomycetes, *Pl. falciparum*, and the brown alga *E. siliculosus* (Table S12). The majority of genes absent from *A. laibachii* Nc14 are absent from other oomycetes and from *Pl. falciparum* but are present in the brown alga. These genes are involved in the photoautotrophic, aquatic life style of diatoms and algae, such as a sodium/bile acid cotransporter, a haloacid dehalogenase-like hydrolase, fatty acid biosynthesis genes, a zeaxanthin epoxidase and a fucoxanthin chlorophyll a/c binding protein. In contrast to the genes lost, we found that certain gene families like aspartic proteases or proteases containing MORN (membrane occupation and recognition nexus) repeats [38] show expansion in *A. laibachii* Nc14 compared to in diatoms. Although our results fit the hypothesis of a common ancestor, we cannot exclude horizontal gene transfer and uptake of an endosymbiont after the divergence between a brown algal ancestor and an oomycete ancestor, given the low number of diagnosed genes that we could analyse.

Potentially green-algae-derived proteins carrying MORN repeat domains (Figure S7) are involved in the complex process of internal budding in apicomplexans [39], which may be similar to the zoospore formation of oomycetes within oospores or zoosporangia or gamete formation in diatoms [40]. While oomycetes with a motile zoospore stage like *A. laibachii* and *P. infestans* carry the MORN repeat proteins, these proteins are absent in the non-motile *H. arabidopsidis* and absent in the non-motile red alga *Cyanidioschyzon merolae* [41]. We therefore hypothesize that loss of this gene of hypothetical green algal origin could have led to the evolutionary loss of the whole flagellum apparatus in *H. arabidopsidis* [4]. However, we cannot rule out that depletion of any major flagellar protein could have caused evolutionary loss of the whole flagellum apparatus. Inspection of the flagellar inner arm dynein 1 heavy chain alpha, which is absolutely necessary for flagellum function, reveals that genomic regions carrying flagellar inner arm dynein 1 heavy chain alpha genes show a high degree of synteny between oomycetes like *Py. ultimum* and *A. laibachii*. In contrast, a syntenic region in *H. arabidopsidis* shows replacement of the flagellar dynein by Mariner- or Gypsy-like transposable elements (Figure S8).

## Comparative Genomics to Identify Genes Implicated in Biotrophy

Since within the peronosporalean lineage, biotrophy evolved twice independently [11], we compared *A. laibachii* with the other obligate biotroph *H. arabidopsidis* [4], hemibiotroph *P. infestans* [9], and necrotroph *Py. ultimum* [10] (Figure 5; Tables S13 and S14). We found that *H. arabidopsidis* is the most diverged from *A. laibachii*.

*H. arabidopsidis* shares the fewest (4,826) orthologous genes with *A. laibachii*, versus the average of 5,722 in *A. laibachii*/*P. infestans* and *A. laibachii*/*Py. ultimum* comparisons. Meanwhile, *H. arabidopsidis* genes show the highest amino acid identity with the genes of *P. infestans*, on average 73% of amino acid identity between all single copy orthologous pairs.

*Py. ultimum* shares the highest number of orthologous genes with *A. laibachii* (5,910 pairs). *P. ultimum* proteins also have a slightly higher percentage of amino acid identity with *A. laibachii* proteins than with other oomycetes (Figure 5). Yet, *Py. ultimum* itself is closer to *H. arabidopsidis* and *P. infestans* than to *A. laibachii*, sharing with them more orthologous genes with higher mean amino acid identity.

These analyses support the hypothesis that *A. laibachii* and *H. arabidopsidis* evolved biotrophy independently; genes missing in one or the other genome compared to the necrotroph *Py. ultimum* or hemibiotroph *P. infestans* may be correlated with biotrophy (Table S15). One of these genes is that for molybdenum-cofactor-dependent nitrate reductase. Nitrate reductase catalyzes pyridine-nucleotide-dependent nitrate reduction for nitrogen acquisition [42]. Both biotroph pathogens have a set of transporters showing homology to amino acid transporters, but other uptake mechanisms or sources could also enable nitrogen acquisition from their hosts [43]. While *H. arabidopsidis* lost only the nitrate reductase, *A. laibachii* also lost the sulphite oxidase and the whole molybdopterin (a cofactor required for nitrate reductase and sulphite oxidase function) biosynthesis pathway. In *Pl. falciparum*, which shows a high degree of adaptation to parasitism, nitrate reductase, sulphite oxidase, and the whole molybdopterin biosynthesis pathway are also missing. Most likely the loss of the two Mo-containing enzymes and the Mo-cofactor biosynthesis is the outcome of biotrophy and not the reason for biotrophy, though conceivably there may have been selection against this pathway if other nitrogen or sulphate sources are less energy-consuming and therefore enhance fitness during parasitism. Molybdenum has been reported to interfere with function of chaperones like Hsp90 [44,45]. Avoiding the uptake of molybdenum might prevent this Hsp90 inhibition and increase fitness on *Ar. thaliana* accessions with high molybdenum levels like Col-0 [46]. *H. arabidopsidis* therefore could be in a less advanced stage of host adaptation compared to *A. laibachii* and *Pl. falciparum*.

Besides biotrophy, the formation of haustoria and haustorium-like structures evolved several times in peronosporalean biotroph and hemibiotroph pathogens. Haustoria in fungi are sites of enhanced nutrient uptake [47] and metabolism, such as thiamine biosynthesis [48]. In the oomycetes, all haustorium-forming species have lost the thiamine biosynthetic pathway. We infer that haustorial oomycetes obtain thiamine from the host.

We therefore hypothesize that evolution to biotrophy is initiated not by gene loss, but rather from the ability to build a haustorium and therefore differentiate a sophisticated interface with a host. The critical step to adopting biotrophy is likely to be efficient defence suppression to enable persistence of functioning haustoria; subsequent loss of biosynthetic pathways is likely to be secondary.

## The *A. laibachii* Secretome

Well-adapted human pathogens like *Pl. falciparum* and plant pathogenic fungi like *Ustilago maydis* have small secretomes (320 [49] and 426 [50] proteins, respectively) compared to necrotrophic fungi like *Aspergillus fumigatus* (up to 881 proteins [51]). We found that the same is true for oomycetes. Using SignalP [52] to predict potential secretion signal peptides and MEMSAT [53] to predict transmembrane (TM) domains, we identified 2,473 (2,136 without TM domains) potentially secreted proteins in the hemibiotroph *P. infestans* and 1,636 (1,222 without TM domains) in the necrotroph *Py. ultimum*. For *H. arabidopsidis* only 1,350 (1,054 without TM domains) and for *A. laibachii* 949 (672 without TM domains) were identified. Analysing the secretome for pathogenicity-related proteins like proteases, glucosyl hydrolases, and potential elicitins or lectins reveals a significant reduction in the *H. arabidopsidis* and *A. laibachii* secretome (Tables 2 and S16). We postulate that biotrophs reduce their activation of host defence by reducing their inventory of secreted proteins, particularly cell wall hydrolyzing enzymes.

## The *A. laibachii* Effector Complement

The ability to establish a sophisticated zone of interaction like the parasitophorous vacuole in *Pl. falciparum* or the haustorium in oomycetes and fungi requires sophisticated host defence suppression [54], which is predominantly achieved via secreted proteins delivered into the host cell [55,56]. The *A. laibachii* secretome comprises 672 secreted proteins without TM domains. Genetically identified oomycete avirulence (Avr) proteins are secreted proteins that have signal peptide and RXLR motifs [57,58]. In many oomycete genomes the RXLR motif is over-represented and positionally constrained within the secreted protein [59]. We identified 25 RXLR and 24 RXLQ effector candidates in the *A. laibachii* secretome. To determine the likelihood that RXLR or RXLQ motifs occur merely by chance in the *A. laibachii* secretome based on amino acid content, we performed *in silico* permutation of the motifs (Figure 6A and 6B). We concluded that the RXLR and RXLQ motifs were not likely to occur merely by chance, and that the likelihood of occurrence by chance is higher in the proteome as a whole than among secreted proteins. It was shown for *P. infestans* that effectors are often located in gene-depleted repetitive regions of the genome [9]. We therefore investigated RXLR candidate proteins in highly repetitive regions of the genome. We identified two RXLRs, one in a highly conserved repeat region with $\sim$10 repeats in Nc14 and one in a more diverged repeat region with $>$80 repeats within the genome. The first region also exists in *A. laibachii* isolate Em1; the diverged repeat of the second identified region exists but without the RXLR gene-containing region (Figure S9). There are 563 RXLR effector candidates identified in *P. infestans* [9], so RXLR effectors are less likely to be relevant for *A. laibachii* virulence.

Similar conclusions can be drawn for the CRN protein family, which shows expansion in *P. infestans* [9,60] but not *A. laibachii*, where only three members of the CRN family could be identified with signal peptides. Eight additional CRN-like proteins were identified where no signal peptide has been predicted.

To identify new classes of effectors in the Albuginales clade, the secretome of *A. laibachii* was computationally screened for genes either showing heterozygosity or showing nucleotide polymorphisms between Nc14 and Em1. We identified a new class carrying a "CHXC" motif by inspection of the first 80 amino acids after the signal peptide cleavage site. CHXC candidates are significantly enriched within the secretome (Figure 6C). Comparisons of the N-terminal part of the CHXC proteins revealed additional conserved amino acids, particularly a glycin at +6 to the CHXC motif (Figure 6D).

## Intraspecies Comparison between *A. laibachii* Nc14 and *A. laibachii* Em1

In host–pathogen interactions, intraspecies comparisons enable the search for virulence alleles that undergo positive selection and fixation within the population [61,62]. Secreted proteins with close contact to the host cell, such as effector proteins, often show enhanced levels of positive selection [63,64]. By comparing the

**Table 2.** Quantitative comparison of pathogenicity-related proteins.

| Protein | *A. laibachii* (Secreted Only) | *A. laibachii* (All) | *H. arabidopsidis* (All) | *P. infestans* | *Py. ultimum* | *T. pseudonana* | *Pl. falciparum* |
|---|---|---|---|---|---|---|---|
| Aspartyl proteases | 1 | 10 | 9 | 12* | 22** | 5 | 11 |
| Serine carboxypeptidases | 6 | 32** | 6 | 24* | 1 | 6 | 0 |
| Cysteine proteases | 16 | 16 | 7 | 33** | 32* | 16 | 4 |
| Glycosyl hydrolases | 15 | 44 | 66 | 157** | 85* | 31 | 1 |
| Pectin esterases | 0 | 0 | 4* | 11** | 0 | 0 | 0 |
| Pectate lyases | 0 | 1 | 8 | 30** | 15* | 0 | 0 |
| Cutinase | 2* | 2* | 2* | 4** | 0 | 0 | 0 |
| Lipases | 3 | 12 | 10 | 19* | 19* | 22** | 9 |
| Phospholipases | 3 | 13 | 13 | 36** | 6 | 18* | 15 |
| Protease inhibitors, all | 0 | 0 | 3 | 38** | 1 | 11* | 1 |
| Cytochrome P450s | 1 | 3 | 16 | 19* | 39** | 7 | 17 |
| ABC transporters | 3 | 41 | 53 | 156* | 173** | 50 | 10 |
| NPP1-like proteins (necrosis-inducing proteins) | 0 | 0 | 24* | 27** | 7 | 0 | 0 |
| Elicitin-like proteins | 1 | 3 | 1 | 40** | 7* | 0 | 0 |
| Lectin-like proteins | 5 | 6 | 6 | 10* | 20** | 0 | 0 |
| Crinklers (CRN family) candidates | 2 | 3 | 20 | 196** | 26* | 0 | 0 |
| RXLR/Q effector candidates | 49 | 49 | 115* | 505** | 57 | 62 | 4 |
| CHXC effector candidates | 29** | 29** | 3 | 5* | 4 | 2 | 1 |

Genes were predicted for all datasets using Pfam prediction and BLASTP against NCBI data or specific datasets of selected protein groups. Results were further compared to data published by Haas et al. for *P. infestans* [9] or Levesque et al. for *Py. ultimum* [10], or Baxter et al. for *H. arabidopsidis* [4]. The data indicate that the *P. infestans* secretome of pathogenicity-related proteins is bigger than that of all other compared and annotated genomes (**, highest number; *, second highest number).
doi:10.1371/journal.pbio.1001094.t002

two *A. laibachii* isolates Nc14 and Em1, we identified a significantly higher frequency of non-synonymous to synonymous mutations within the predicted secretome compared to the rest of the proteome. Our analyses showed that this was particularly true for heterozygous positions and less convincing for homozygous SNPs (Table S17). Genes that are highly conserved between species, like KOGs, showed comparable non-synonymous and synonymous substitution rates, with a slight excess of synonymous mutations. There are significantly more genes within the KOGs showing a non-synonymous/synonymous ratio less than 1 than genes with values greater than 1. Comparing this to candidate effector classes like RXLRs, RXLQs, and CHXCs reveals that in particular the CHXCs show significantly higher frequencies of non-synonymous to synonymous mutations. This supports the idea that the CHXC sub-class of secreted proteins is under positive selection, similar to other described oomycete effectors like ATR1 or ATR13 from *H. arabidopsidis* [57,65].

Further to this we identified Nc14 genes absent or highly diverged from the Em1 complement. We defined a gene as absent or highly diverged if >10 bp showed 0 coverage in the Em1 alignment. Out of the 672 secreted proteins without TM domains, we identified seven as absent from Em1 (1.04%). We also detected two with a predicted TM domain (0.73%) that are absent from Em1. Regarding all gene models, 96 were absent (0.74%). This finding is a further indication for a greater selection pressure on secreted than on non-secreted proteins, as has been found in species or interspecies comparisons in *Phytophthora* sp. [66] and *Ustilago/Sporisorium* [67].

## Validation of Effector Delivery

We tested *A. laibachii* effector candidates (one CHXC, one RXLR, and one CRN effector candidate) for their host delivery efficiency using a *P. capsici–Nicotiana benthamiana* translocation assay [68]. Briefly, N-terminal domains of candidate effectors were fused to the *P. infestans* Avr3a effector domain, transformed into *P. capsici*, and tested for whether they confer translocation of Avr3a into *N. benthamiana* carrying *R3a*, resulting in avirulence. Statistical analyses of the delivery efficiency (Figure 7) clearly indicate that the *A. laibachii* CRN3 N-terminus and CHXC9 N-terminus are as efficient as the Avr3a N-terminus in Avr3a translocation, while the RXLR1 N-terminal domain is less efficient. An alanine replacement construct of the CHXC motif supports the importance of this motif for delivery efficiency. The Avr3a C-terminus alone confers a low basal delivery level without the need for the N-terminal enhancer. These findings reveal the potential of the CHXC proteins to be delivered into the host cell, similar to RXLRs and CRNs, though the delivery mechanism for all these effector classes requires further investigation.

## Validation of Virulence-Conferring Function of *A. laibachii* Effector Candidates

To assay the effectors for virulence function, we used *Pseudomonas syringae* pv. tomato (Pst) DC3000 luciferase [69] carrying "effector detector vector" (EDV) constructs to deliver effectors into the plant cytoplasm via type III secretion [70] (Figure 8). Tests on *Ar. thaliana* Nd-0 plants revealed that several selected *A. laibachii* RXLRs, CRNs, and CHXCs enhance virulence compared to a non-functional AvrRps4 (AvrRps4[AAAA]). On *Ar. thaliana* Col-0, in contrast, the CRN and one RXLR (RXLR1) do not enhance virulence while RXLR2 and CHXCs still do. These tests indicate that CHXCs carry the capacity to enhance virulence in phytopathogenic bacteria, perhaps by suppression of host resistance mechanisms [54,70].

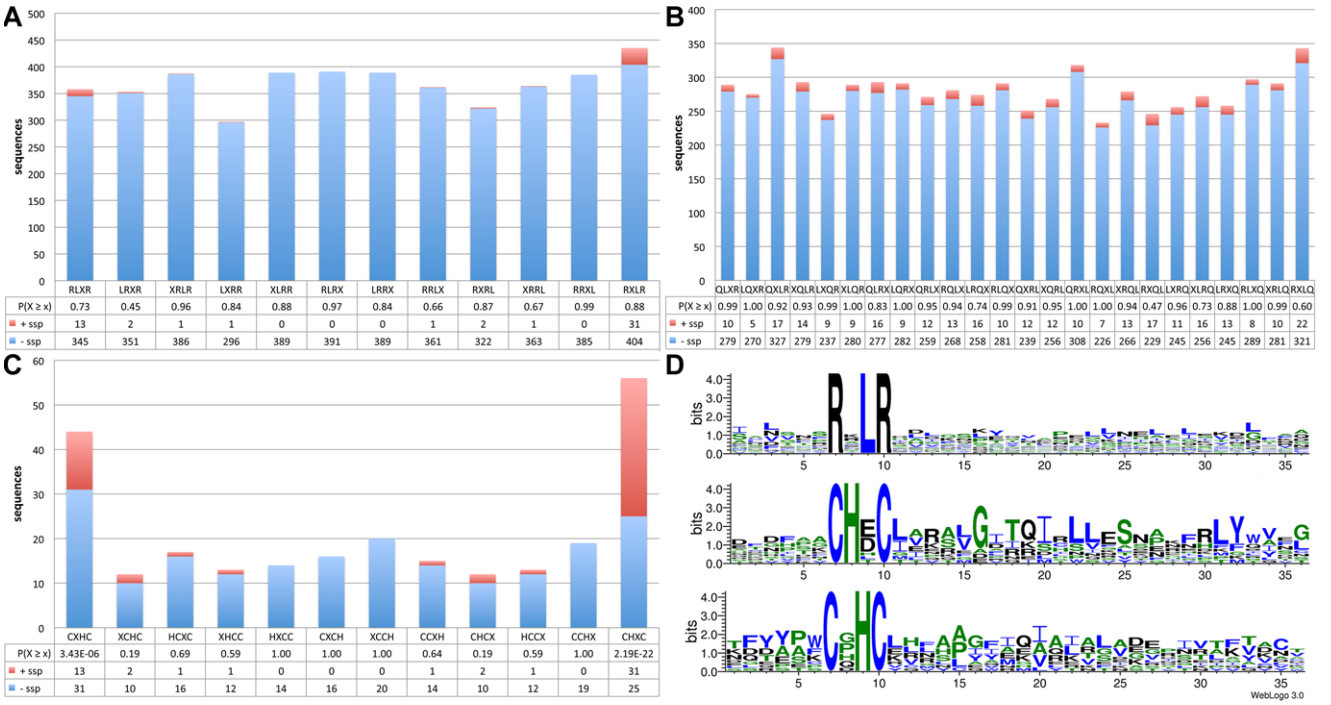



**Figure 6. Validation and identification of potential delivery motifs.** To identify potential effector delivery motifs we analysed RXLR (A), RXLQ (B), and the new CHXC (C) effector candidates for enrichment in the secretome. Motif shuffling was used to identify background levels. Calculating the cumulative hypergeometric probability to analyse the enrichment of secreted proteins (red) over non-secreted proteins (blue) for each of the permutated motifs reveals a significant enrichment of CHXCs in the secretome ($p[X \geq x] = 2.19 \times 10^{-22}$). None of the RXLR or RXLQ motifs or permutations shows significant enrichment. There is also enrichment for CXHC proteins in the secretome. Except for one CHHC protein, there is no overlap between the two motif classes. The logo blot (D) clearly indicates that RXLR-containing proteins are conserved only within the selected amino acids, while for CHXCs it is not only the motif but also sequences C-terminal to it are conserved, including conserved glycine, leucines, and a tyrosine.
doi:10.1371/journal.pbio.1001094.g006

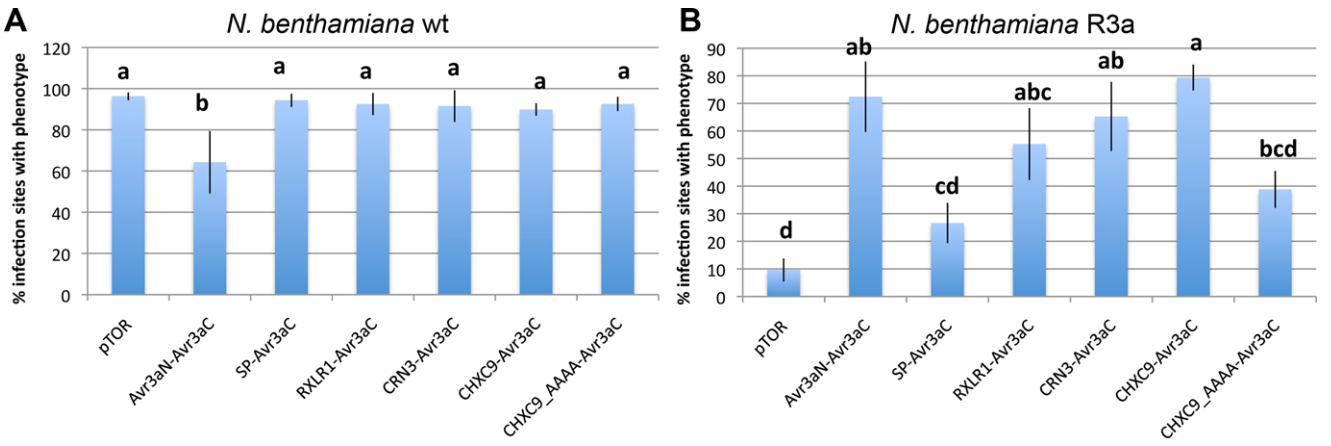



**Figure 7. *P. capsici* test for delivery motif.** To identify known and new classes of transfer motifs in *A. laibachii*, the *P. capsici–N. benthamiana* translocation assay was used. This test system is based on Avr3a-mediated avirulence in plants carrying R3a [68]. (A) virulence assay to show that transgenic *P. capsici* is not impaired in growth; (B) Hypersensitive response (HR) assay on R3a-carrying plants for delivery assay. The Avr3a RxLR translocation domain is replaced by the N-terminus of different *A. laibachii* effector candidates. For the assay, RXLR1, CRN3, and CHXC9 carrying the newly identified CHXC motif were used. Our results validate that known motifs like the CRN motif are functional while the selected RXLR shows low delivery efficiency. CHXC9 shows the same efficiency as Avr3a does, and dependency of the CHXC motif could be identified (statistical analyses using the Tukey test; means with the same letter are not significantly different; error bars denote standard error of the mean). wt, wild type.
doi:10.1371/journal.pbio.1001094.g007

These virulence assays together suggest that *A. laibachii* uses at least three different major effector classes.

### Evolutionary Origin of CHXC Effectors

To try to identify the evolutionary source of CHXCs, we investigated enrichment of CHXC-motif-containing proteins in the secretomes of *P. infestans*, *Py. ultimum*, *H. arabidopsidis*, *Saprolegnia parasitica*, *Thalassiosira pseudonana* (diatom), *Pl. falciparum* (Apicomplexa), *E. siliculosus* (brown alga), *C. merolae* (red alga), *Ch. reinhardtii* (green alga), *Volvox carteri* (green alga), and *Ar. thaliana*. Only *A. laibachii* contained a significant enrichment of CHXCs in its secretome. Although not significantly enriched, both the fish pathogen *S. parasitica* and the land plant *Ar. thaliana* contained more than ten CHXC proteins carrying potential secretion signals (14 and 11, respectively) (Figure S10). In contrast to CHXC-containing proteins, almost all inspected organisms show a high number of CXHC-containing potentially secreted proteins; a common CXHC protein is protein disulphide isomerase (Table S18).

Given that *A. laibachii* CHXCs show the closest clustering with *S. parasitica*, *V. carteri*, *Ch. reinhardtii*, and *Ar. thaliana* CHXCs (Figure 9), conceivably this candidate effector class evolved from an ancestral green-alga-derived gene.

Whatever their origin, we conclude that CHXC proteins are present in all organisms analysed but evolved effector function only in Albuginales and possibly Saprolegniales. In Albuginales, one N-terminal sub-class of CHXCs (CHxCLx(4)Gx(5–6)L) shows significant expansion, with 23 members, while other CHXCs are distinct from this clade. *S. parasitica* CHXCs are distinct from this major *A. laibachii* clade and therefore remain to be tested in future experiments.

### Conclusions

The *A. laibachii* genome assembly sheds light on the evolution of biotrophy since it allows the first comparison, to our knowledge, of two oomycete obligate biotroph pathogens (*A. laibachii* and *H. arabidopsidis*) that evolved biotrophy independently. In addition, *A. laibachii* shows the highest overall amino acid identity to the necrotroph pathogen *Py. ultimum* and the hemibiotroph *P. infestans*. One of the striking results of this comparison is that all organisms able to build haustoria have lost their thiamine biosynthesis pathway, presumably because thiamine is easily obtained from hosts via the haustorial interface. A closer interface requires effective host defence suppression. We therefore hypothesize that the evolution of biotrophy involves a series of steps: step 1, involving progressively more effective effectors to suppress defence, step 2, attenuated activation of defence by reduction in the inventory of cell wall hydrolyzing enzymes, resulting in, step 3, weak selection to maintain certain biosynthetic pathways if the products of the pathways can be directly obtained from the host. This results in progressively more comprehensive auxotrophy and culminates in irreversible biotrophy (Figure 10).

## Materials and Methods

### Field Isolate

An infected leaf was harvested from an *Ar. thaliana* plant grown in a heavy infected field plot in Norwich (UK; 52.6236,1.2182) [21] in December 2007. Zoosporangia were washed off the leaf surface and used to infect *Ar. thaliana* Ws-0-eds1 plants. After 1 wk one pustule was punched out, and spores were placed on ice for 30 min to release zoospores. Unhatched zoosporangia were removed by filtration, and zoospores were diluted to ~10 zoospores/ml and sprayed on *Ar. thaliana* Ws-0 plants (~100 µl/plant). This procedure was repeated 4× until spores were bulked up on *Ar. thaliana* Ws-0 plants. Zoosporangia were harvested using a home-made cyclone spore collector [71].

### Plant Inoculation

Zoospores were suspended in water ($10^5$ spores/ml) and incubated on ice for 30 min. The spore suspension was then sprayed on plants using a spray gun (~700 µl/plant), and plants were incubated in a cold room in the dark over night. Infected plants were kept under 10-h light and 14-h dark cycles with a 20°C day and 16°C night temperature.

### DNA Extraction and Sequencing

High molecular weight DNA was extracted from zoosporangia using a phenol/chloroform-based purification method after grinding in liquid nitrogen, adapted from [72]. Library prepara-

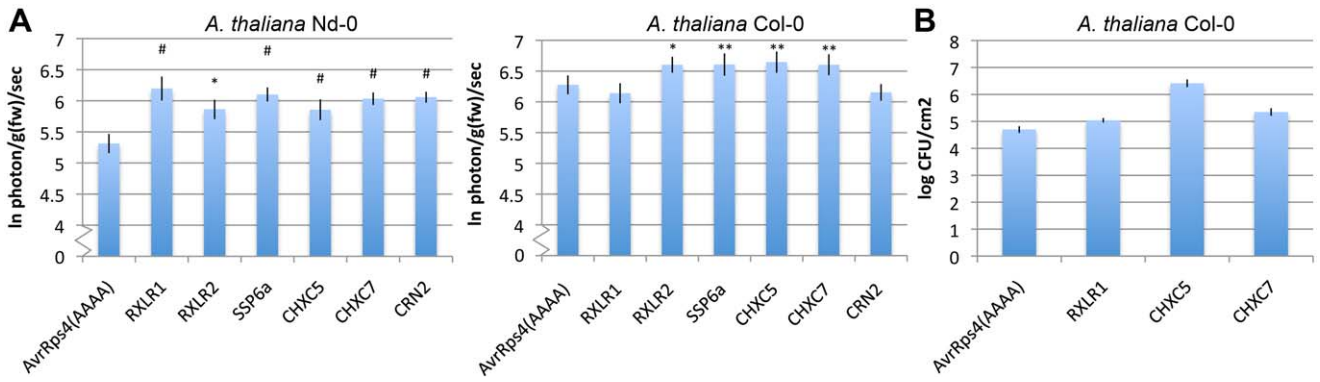



**Figure 8. Candidate *A. laibachii* Nc14 effectors contribute to Pst DC3000 virulence.** (A) *Arabidopsis* plants (4- to 5-wk-old) were spray inoculated with $5\times10^8$ CFU Pst DC3000 lux harbouring candidate effectors cloned in pEDV6. Bacterial growth was measured as an increase in luciferase photon emission per gram fresh weight per second (photon/g[fw]/sec). The histogram represents the log median of photon emission of three independent experiments, each with four technical replicates. Error bars denote standard error of the mean. Two-way ANOVA: #, $p<0.001$; **, $p<0.01$; *, $p<0.05$ from AvrRps4(AAAA). (B) Plants 4- to 5-wk-old were infected with $5\times10^8$ CFU of Pst DC3000 ΔAvrPto/ΔAvrPtoB harbouring candidate effector cloned in pEDV6. Bacterial populations were sampled 4 d post-inoculation. The histogram represents the median colony count of two independent experiments, each with more than four technical replicates. Error bars denote standard error of the mean.
doi:10.1371/journal.pbio.1001094.g008

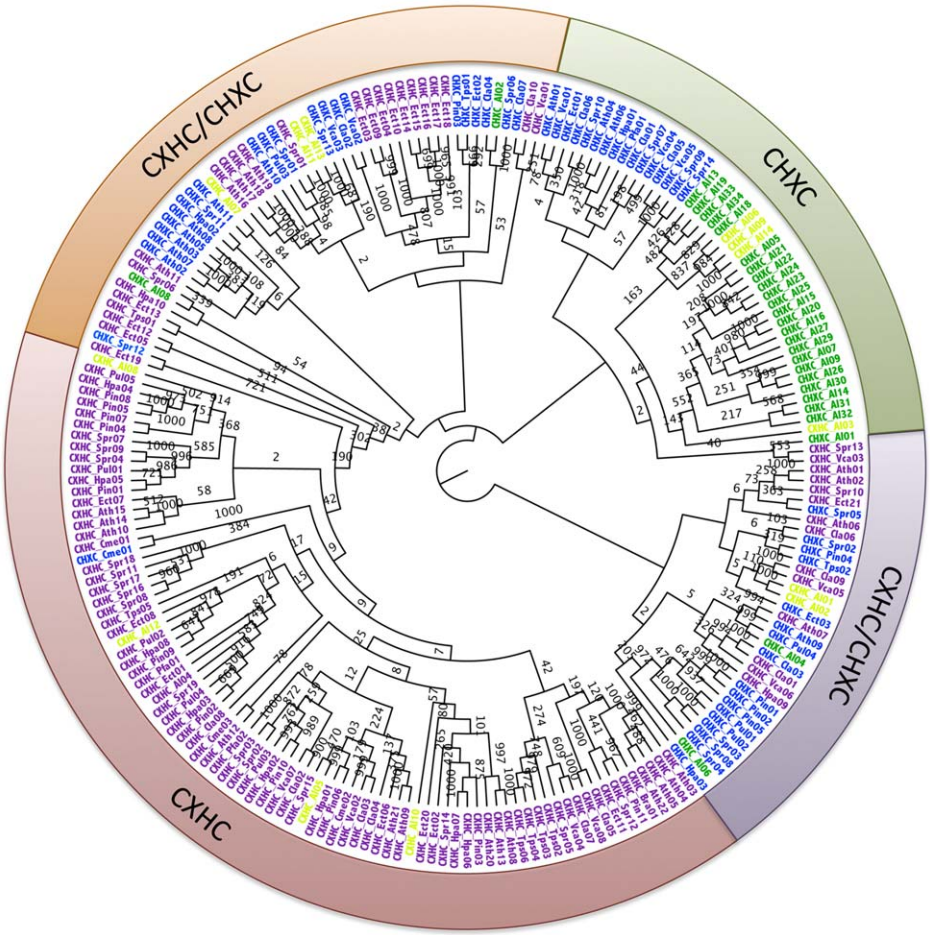



**Figure 9. Result of neighbour-joining analyses using N-termini of all predicted CHXCs or CXHCs from the genomes of *P. infestans*, *Py. ultimum*, *H. arabidopsidis*, *T. pseudonana*, *Pl. falciparum*, *E. siliculosus*, *C. merolae*, *Ch. reinhardtii*, *V. carteri*, *S. parasitica*, as well as *Ar. thaliana*.** The outer ring summarises clades with N-termini predominantly carrying CHXC or CXHC motif or mixed clades (CXHC/CHXC) into classes. *A. laibachii* CHXCs are mainly clustered in the CHXC class (green), containing besides *A. laibachii* distantly related CHXCs from *S. parasitica*, *V. carteri*, *Ch. reinhardtii*, and *Ar. thaliana*. CHXCs are distant from endoplasmic reticulum proteins like disulphide isomerases that predominantly carry the CXHC motif and are grouped within the CXHC class (red). Between the CHXC class and the CXHC class, mixed clades contain protease and defensin homologues (orange) or *Ar. thaliana* cystein-rich proteins (violet). (Names in green indicate *A. laibachii* CHXCs and in yellow, *A. laibachii* CXHCs. Blue indicates CHXCs from other species; magenta indicates CXHCs from other species; 16 amino acids before and 45 amino acids after the CHXC or CXHC motif in the N-terminus were used. The tree is midpoint rooted. All bootstrap counts refer to 1,000 replications.). Ath, *Ar. thaliana*; Cla, *Ch. reinhardtii*; Cme, *C. merolae*; Ect, *E. siliculosus*; Hpa, *H. arabidopsidis*; Pfa, *Pl. falciparum*; Pin, *P. infestans*; Pul, *Py. ultimum*; Spr, *S. parasitica* (Spr); Tps, *T. pseudonana*; Vca, *V. carteri*.
doi:10.1371/journal.pbio.1001094.g009

tion for Illumina sequencing was performed as described [28]. All data were generated using paired-end reads. 800 bp and 400 bp paired-end sequencing libraries were constructed, and 8.8 Gbp of usable data were generated (for read and insert length, see Figure 1A).

## Calculation of Expected Nc14 Genome Size

Figure 1A lists all reads after purification from plant and bacterial contamination as well as all reads aligned to the assembly. In summary, 91.6% of all reads can be aligned to the contigs, suggesting 2.8 Mbp missing from the assembly. Since 32.7 Mbp are in the assembly, the genome can be estimated to 35.5 Mbp. In another approach considering all reads and their read length, 8.8 Gbp (~7% correction for lower quality of second read pair) were generated, which would lead to an expected coverage of the 32.7 Mbp genome of ~270×. The mean coverage using single copy genes (glycolysis and TCA) is 240×. Considering the 2.5 Mbp of repeats (Figure 1B, right side, coverage underestimated) with an average coverage of 1,086×, which is ~4.4 times more than the mean coverage of the contigs, this repeat region corresponds to 10.9 Mbp. In contrast to this, the genome contains ~6.2 Mbp of hemizygous regions (Figure 1B, left side, coverage overestimated). These calculations suggest a genome size of ~43 Mbp, given all repeats resolved, or an effective genome size of ~37 Mbp.

## cDNA Preparation and Sequencing

*A. laibachii*–infected *Ar. thaliana* Ws-0 plants were harvested 0 (after cold room, see plant inoculation), 2, 4, 6, 8 and 10 d after infection. Total RNA was extracted using TRI Reagent RNA Isolation Reagent (Sigma), and Dynabeads (Invitrogen) were used to enrich for mRNA. First and second strand cDNA synthesis was performed according to manufacturer's instructions using the SMART cDNA Library Construction Kit (Clontech), and cDNA was normalized using the Trimmer kit from Evrogen. cDNA samples were mixed in equal amounts and fragmented using a

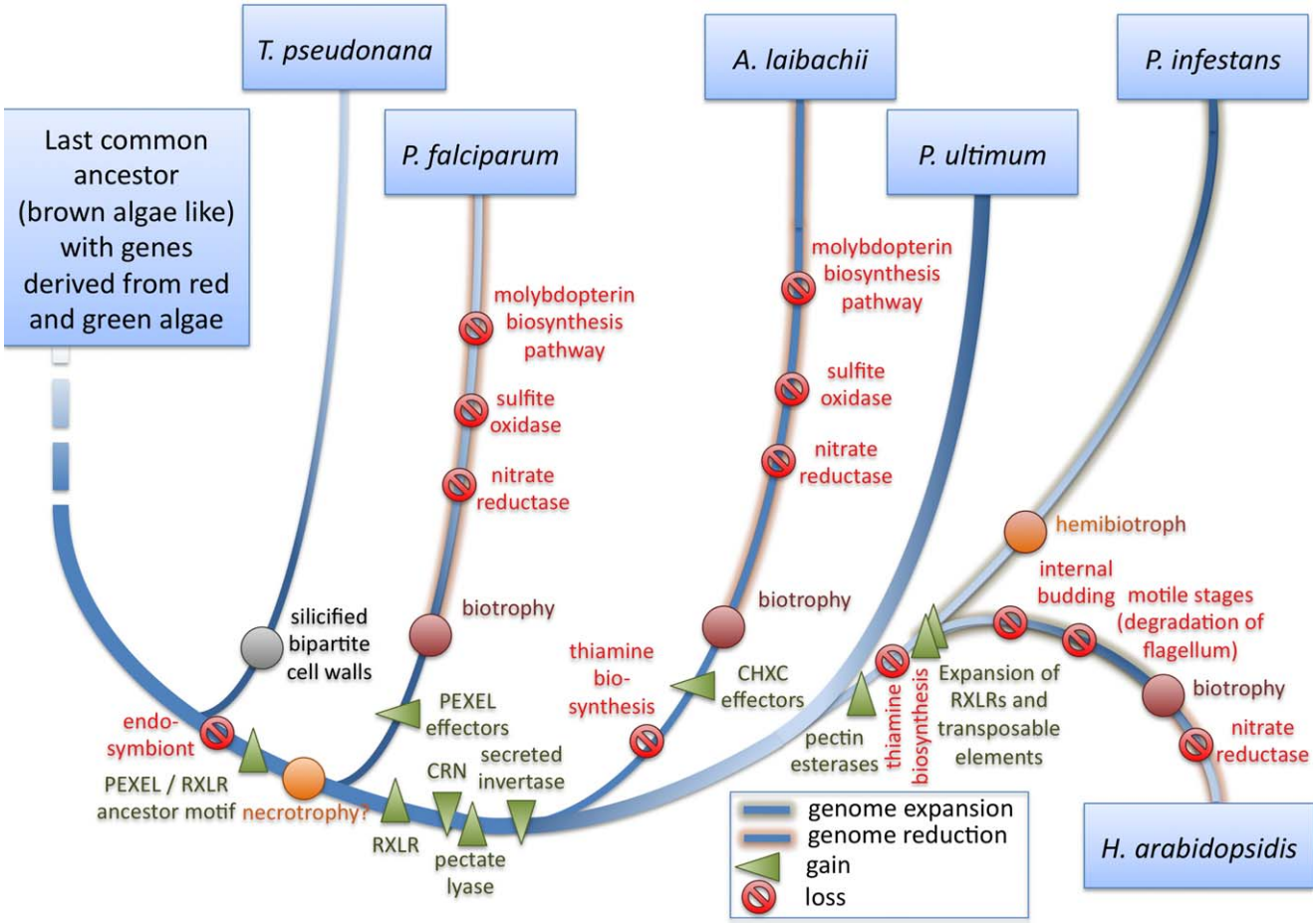



**Figure 10. Gain and loss of genes and pathways for selected Chromalveolata in comparison to *A. laibachii*.** It was hypothesized that the last common ancestor of Chromalveolata was a brown-alga-like organism with genes from green and red algae integrated into the nuclear genome after primary and secondary endosymbiosis [32,37]. While some heterokonts kept their secondary endosymbiont and, in the case of diatoms, acquired a silicated bipartite cell wall [97], others lost their secondary endosymbiont. We postulate that after the loss of the endosymbiont, convergent evolution led to effector proteins like PEXEL [49,98] and RXLR precursors. PEXEL effectors might have enabled *Pl. falciparum* to achieve more complex interactions with its host and establish intercellular growth. In addition to the RXLR effector proteins, oomycetes acquired or evolved another class of effectors, the CRNs [60] and a secreted invertase that allows use of sucrose from host plants [99]. Oomycetes that are biotrophs or hemibiotrophs today lost their thiamine biosynthesis pathway and, in the case of *A. laibachii*, evolved a new "CHXC" effector class. After taking up the biotroph lifestyle, the genomes of *Pl. falciparum*, *H. arabidopsidis*, and *A. laibachii* started a gene reduction that is exemplified by looking at enzymes that require molybdenum cofactors and the molybdopterin biosynthesis pathway. Hemibiotroph *P. infestans* instead shows a strong genome expansion [9]. In this context, *H. arabidopsidis* showed a genome expansion and acquired biotrophy late, based on the loss of only one molybdenum-dependent enzyme.
doi:10.1371/journal.pbio.1001094.g010

Covaris sonicator (Covaris). Illumina libraries were prepared as described for fragmented genomic DNA [28].

## Data Acquisition

Data for comparative genomics were downloaded from the sources listed in Table 3.

## Genome Assembly

First Velvet [24] was used, running different kmer-lengths and different sequencing library subsets (kmer-length: 23, 31, 41, 45, 49, 55, 61, 67, and 73; subsets: 400-bp insert only, 800-bp insert only). N50 number and length were determined for each of the assemblies, and the best assembly was selected as the matrix to be used with the Minimus2 genome merge pipeline [25]. For the current assembly the 400-bp only subset with kmer-length 61 was used as matrix, and for kmer-lengths 49, 55, 61, 67 and 73, all 400- and 800-bp assemblies were added (Minimus parameters: consensus error $<0.001$; minimum identity $>99\%$; 20-bp maximum trimming). A set of genes showing high heterozygosity was used to ensure that contigs were properly joined. Parameters were changed through several rounds, and minimum overlap, in particular, was lowered from 100 bp to 15 bp. An overlap of 15 bp was found to be the optimum for difficult heterozygous regions. After each Minimus assembly, all reads were back aligned to the contigs using MAQ aligner [73]. Regions showing less than $3\times$ average coverage were removed, and redundant fragments were removed using BLASTN with an e-value cut-off of $1e^{-20}$ and 99.9% identity. After this step a next round of Minimus was started, with changing minimum overlap in steps of 20 bp down from 100 bp. Below 20 bp steps were changed by 5 bp (See Figure S1 for work flow).

Since it is impossible to cultivate obligate biotrophs under sterile conditions, plant and bacterial contaminations were removed by using BLAST against genome sequences of the host plant *Ar.*

**Table 3.** Sequence sources for comparative genomics analyses.

| Organism | Download Site | Reference | Genome Version | Annotation/ Proteins | Download Site Host |
|---|---|---|---|---|---|
| *Chlamydomonas reinhardtii* | http://genome.jgi-psf.org/Chlre4/Chlre4.home.html | [100] | 4.0 | 3.1 | DOE Joint Genome Institute |
| *Ectocarpus siliculosus* | http://www.ebi.ac.uk/ena/ | [101] | 2.0 | 20100616100000 | European Bioinformatics Institute |
| *Phaeodactylum tricornutum* | http://genome.jgi-psf.org/Phatr2/Phatr2.download.ftp.html | [102] | 2.0 | 20070523 | DOE Joint Genome Institute |
| *Thalassiosira pseudonana* | http://genome.jgi-psf.org/Thaps3/Thaps3.info.html | [103] | 3.0 | 2.0 | DOE Joint Genome Institute |
| *Saccharomyces cerevisiae* | ftp://genome-ftp.stanford.edu/pub/yeast/sequence/genomic_sequence/ | [104] | Nov 30 2006 | Jan 06 2010 | Saccharomyces Genome Database |
| *Debaryomyces hansenii* | http://www.ebi.ac.uk/2can/genomes/eukaryotes/Debaryomyces_hansenii.html | [105] | CR382133.2 | CR382133 | European Bioinformatics Institute |
| *Toxoplasma gondii* | http://toxodb.org/common/downloads/release-6.0/Tgondii/ | [106] | 6.0 | 6.0 | ToxoDB |
| *Plasmodium falciparum* | http://plasmodb.org/common/downloads/release-6.3/Pfalciparum/ | [107] | 6.3 | 6.3 | PlasmoDB |
| *Homo sapiens* | ftp://iubio.bio.indiana.edu/eugenes/2003/man/ | [108] | Jun 23 2002 | Jun 24 2002 | IUBio Archive |
| *Takifugu rubripes* | http://www.fugu-sg.org/downloads/downloads3.htm | [109] | 5 | 5 | Institute of Molecular and Cell Biology |
| *Phytophthora sojae* | http://genome.jgi-psf.org/Physo1_1/Physo1_1.home.html | [110] | 1.1 | 1.1 | DOE Joint Genome Institute |
| *Phytophthora infestans* | http://www.broadinstitute.org/annotation/genome/phytophthora_infestans/MultiDownloads.html | [9] | 4.1 | 4.1 | Broad Institute |
| *Saprolegnia parasitica* | http://www.broadinstitute.org/annotation/genome/Saprolegnia_parasitica/Downloads.html | [111] | 1 | 1 | Broad Institute |
| *Pythium ultimum* | http://pythium.plantbiology.msu.edu/download.html | [10] | Release 1 | Release 1 | Michigan State University |
| *Hyaloperonospora arabidopsidis* | http://vmd.vbi.vt.edu/download/index.php | [4] | 8.3.2 | 8.3.2 | Virginia Bioinformatics Institute |
| *Ustilago maydis* | http://www.broadinstitute.org/annotation/genome/ustilago_maydis.2/MultiDownloads.html | [50] | Release 2 | 1 | Broad Institute |
| *Fusarium oxysporum* | http://www.broadinstitute.org/annotation/genome/fusarium_group/MultiDownloads.html | [112] | 2 | 2 | Broad Institute |
| *Volvox carteri* | http://genome.jgi-psf.org/Volca1/Volca1.download.ftp.html | [113] | 2 | 2 | DOE Joint Genome Institute |
| *Cyanidioschyzon merolae* | http://merolae.biol.s.u-tokyo.ac.jp/download/ | [114] | Jul 03 2007 | Jan 18 2008 | University of Tokyo |

doi:10.1371/journal.pbio.1001094.t003

*thaliana* (TAIR 9.0), fungal genomes (*Neurospora crassa*), oomycetes (*H. arabidopsidis*), and diverse bacterial genomes (*Xanthomonas* sp. and *Pseudomonas* sp.).

## Prediction of Heterozygous Loci

To identify heterozygous loci, Illumina reads were aligned using MAQ, and the SNP detection pipeline was used according to the manual, with default parameters and minimum coverage greater than 180× for the Nc14 alignment and greater than 20× for the Em1 alignment. From the MAQ SNP file, positions were selected where two bases are possible and maximum coverage was less than 350×.

## Repetitive Elements

Assembled repetitive elements were identified using the RepeatScout program (http://bix.ucsd.edu/repeatscout/) with a seed size of 14. The frequency of elements and their location in the assembly were estimated with RepeatMasker using a library of repetitive elements built up by RepeatScout. A sequence was considered to be repetitive if it occurred in the genome assembly on at least three different contigs. The resulting library was searched for the sequences homologous to the known transposon elements using TBLASTX (e-value cut-off of $1e^{-5}$) and a database of transposons, RepBase [74]. Consensus repeats that matched predicted Nc14 protein coding genes were filtered out. The remaining consensus repeats that do not match any sequences deposited in the NCBI database or any known transposon element and that do not overlap with Nc14 protein coding genes represent either *Albugo*-specific repeats or simple repeats.

tRNA genes were predicted with the program ARAGORN [75] using first default parameters and second options allowing introns in the gene sequences.

## Genome Quality Using CEGMA

CEGMA was used according to the manual [26] with a local installation.

## cDNA Assembly

For the combined ABySS [76] and Oases [77] assembly, adaptor sequences from the SMART kit cDNA synthesis were

removed for the ABySS assembly, and the ABySS program was used according to the manual. Different kmer-lengths were tested, and a length of 61 used for the final assembly.

Untrimmed cDNA sequences were assembled using Velvet and a kmer-length of 51, 57, 61, and 71. Oases was used for the final assembly of the contigs according to the manual, using default parameters.

MUMmer in maxmatch mode was used to combine all ABySS and Velvet assemblies. Redundant contigs were removed using BLAST.

Since the assembled cDNA is not strand specific but orientation is needed for gene prediction, cDNA 5′ tags were generated by Illumina sequencing (E. Kemen, A. Balmuth, J. D. Jones, unpublished data). Using Bowtie aligner [78], cDNA 5′ tags were aligned onto the assembled cDNA and, based on tag counts, orientated in the 5′ to 3′ direction.

## cDNA Alignments

To map assembled cDNA against the genome, either BLAT [79] in trimT and fine mode or PASA [80] with default settings was used.

Illumina reads were directly mapped to the genome using the Bowtie aligner, in "best" mode and with strand correction (strandfix mode). Pileup files were generated using bowtie-maqconvert and maq pileup allowing four mismatches per 76-bp read. To incorporate this data as hints files for gene prediction, regions with greater than 3× coverage were extracted.

## Gene Prediction and Annotation

To generate a reliable gene set to train further programs, GeneMark [81] was used for *ab initio* gene prediction. ORFs plus 50 bp on the 3′ end and 50 bp on the 5′ end were extracted, and Illumina-sequenced cDNA was aligned to the ORFs using Bowtie. Gene models were selected if the coverage within the ORF didn't drop below three. This dataset with more than 2,000 genes was used as "traingenes" for the automated training program provided with the Augustus package (autoAug.pl). The trained Augustus program was then used for gene prediction including the combined Oases/ABySS-assembled cDNA (mapped using BLAT) as evidence. Default parameters (extrinsic.ME.cfg) were used for all predictions.

For consensus gene predictions with *P. infestans*, SGP2 was used according to the manual [82].

ASGARD [30] alignments were converted into GFF files to be used for consensus predictions.

Consensus gene models were generated using Evigan [83]. cDNA from assemblies and alignments was converted into GFF files and combined with Augustus, GeneMark, SGP2, and ASGARD predictions. The genome was than screened for gene-free regions, and Augustus gene predictions were added if available. In a third round, regions that did not contain consensus gene models or Augustus gene models were extracted, and GeneMark annotations were added if available.

A set of genes was further tested by 5′ and 3′ RACE to validate start and stop sites.

## Orthologous Genes and Divergence Level

Molecular divergence of *A. laibachii* from other species was assessed by examining the percentage of amino acid identity between orthologous gene pairs [75].

Orthologous pairs were identified using the OrthoMCL program with an e-value cut-off of $1e^{-5}$ [84]. Alignments of protein pairs were performed with MUSCLE [85].

Amino acid identity was calculated only for the single copy genes by either excluding alignment gaps from calculations or taking gaps into account. The results show similar trends, so we present only results for the calculations when alignment gaps were excluded.

The total number of orthologous groups identified between species and the number of one-to-one orthologous pairs, as well as a mean amino acid identity, are shown in Table S7. In the comparison of *T. gondii* and *A. laibachii*, we found few orthologous pairs represented by the single copy genes (23 pairs); therefore, we excluded this pair of species from the analyses of sequence divergence.

We also estimated the levels of amino acid identity for the core eukaryotic genes (orthologous genes shared by all examined species); these data are presented in Table S8.

## Green- and Red-Alga-Derived Genes

To identify *A. laibachii* genes with sequence similarity to green- or red-algal-derived diatom genes, a set published by Moustafa et al. [37] was used. All *A. laibachii* proteins showing homology to genes identified by Moustafa et al. [37] were further blasted (BLASTP) against the *Ch. reinhardtii* gene set, the *E. siliculosus* gene set, the *U. maydis* gene set, and the *Fusarium oxysporum* gene set with an e-value cut-off of $1e^{-20}$. Genes were considered to be green-alga-derived only if the protein was absent from *U. maydis* and *F. oxysporum* but present in *Ch. reinhardtii*, and was considered red-alga-derived if not in *U. maydis* or *F. oxysporum* but in *E. siliculosus*. The same analyses were performed on the *Saccharomyces cerevisiae*, *Pl. falciparum*, *H. arabidopsidis*, *P. infestans*, *Py. ultimum*, *V. carteri*, *Ch. reinhardtii*, *C. merolae*, *C. merolae*, *Th. pseudonana*, and *Ph. tricornutum* gene sets.

*A. laibachii* candidate genes with significant sequence similarity to green or red algae and other oomycetes (e-value cut-off of $1e^{-20}$) but not to fungi, brown algae, or diatoms were identified using the criteria in Table 4. Representative organisms for each group are as follows: green algae: *V. carteri*, *Ch. reinhardtii*; red algae: *C. merolae*, *Galdieria sulphuraria*; fungi: *F. oxysporum*; brown algae: *E. siliculosus*; diatoms: *Ph. tricornutum*, *Th. pseudonana*; oomycetes: *P. sojae*, *Py. ultimum*, *H. arabidopsidis*.

Homologues between oomycetes, fungi, brown algae, and diatoms were identified using OrthoMCL (e-value cut-off of $1e^{-20}$ or $1e^{-5}$) [37].

## Synteny

Synteny between multiple species was analysed using the Artemis Comparison Tool [86]. Alignments between genomic sequences were performed using TBLASTX with a score cut-off of 210. Annotations of *P. infestans*, *Py. ultimum*, and *H. arabidopsidis* were transferred using TBLASTN with an e-value cut-off of $1e^{-30}$. LTR_FINDER [87] was used to annotate long terminal repeats (LTRs) within the genomic sequences, and coordinates were manually added. Regions between LTRs were blasted against RepBase [74] to identify the presence and/or type of transposon.

## Defining the Secretome

Secreted proteins were predicted using a local installation of SignalP 3.0 [88]. Proteins were considered to be secreted if both the neural networks and hidden Markov model methods predicted the protein to have a signal peptide. Predictions of TM domains were performed after removing the predicted secretion signal. TM domains were identified using MEMSAT3 [89]. Proteins were considered to be without a TM domain with $p_{\text{non-TM}}>0.0004$ or, for high stringency, $p_{\text{non-TM}}>0.01$.

## Motif Discovery

To identify new motifs, subsets of secreted proteins were selected and analysed using MEME [90] with default parameters.

**Table 4.** Criteria for identification of red- and green-alga-derived genes.

| Category | Presence/Absence | | | | | |
|---|---|---|---|---|---|---|
| | Green Algae | Red Algae | Fungi | Brown Algae | Diatoms | Oomycetes |
| Genes of Nc14 with significant sequence similarity with green algae | + | − | − | − | − | + |
| Genes of Nc14 with significant sequence similarity with red algae | − | + | − | − | − | + |

doi:10.1371/journal.pbio.1001094.t004

Identified motifs were tested against the whole gene set and the Swiss-Prot database using MOTIF Search. In a second step, motifs were selected only if they were positioned within 50 amino acids after the secretion signal.

Tests for over-representation of an identified motif were done using motif and sequence shuffling. Secreted proteins were predicted [88] as described in the previous section, and the signal peptide was removed prior to further analyses.

Each of the sequences without secretion signal was randomly shuffled 30 times. After each shuffling the sequences were screened for the motif in question. If the motif was identified after shuffling, the sequence was excluded from the next round. If the motif was never identified within the 30 times shuffling, the motif in the original protein was counted as "unique empirical". All possible combinations of the amino acid sequence within the motif were calculated. For each of these permutations, the "unique empirical" proteins were calculated.

The 30 times shuffling was repeated 1,000 times to calculate background levels. Background levels were defined as how often a sequence was found again having the motif or the permutated motif. This was called "background (mean)". Motifs that were above this background were considered for further analyses.

The second criterion was if a motif was significantly enriched in the secretome compared to all non-secreted proteins. For statistical validations we calculated the cumulative hypergeometric probability.

## Selection of Candidates for Further Experiments

Candidates for further experiments were evaluated according to a ranking list. Maximum possible score was nine points, and the following scores were given: one point for being on a shorter, repetitive contig (≤3,000 bp) or end of contig, since we assumed that effector candidates might be in repetitive regions as shown for *P. infestans* effectors [9]; one point for having cDNA support; two points for being a short protein (≤400 amino acids); two points for carrying one of the identified motifs (RXLR, RXLQ, CHXC, CRN); one point for being expressed before day 10 after infection; one point for being expressed before day 4 after infection; and one point for showing SNPs in the Em1 comparison.

## *P. capsici* Tests

**Plant and bacterial growth procedures and *P. capsici* culturing.** *N. benthamiana* plant genotypes and *P. capsici* strain LT1534 were grown and cultured as described by Schornack et al. [68]. *P. capsici* transformation was performed as described by Schornack et al. [68].

**Plasmid construction and preparation.** *Phytophthora* transformation constructs SP_AVR3aC, RXLR1_AVR3aC, CRN3_AVR3aC, CHXC9_AVR3aC, CHXC9AAAA_AVR3aC, CHXC7_AVR3aC, and CHXC7AAAA_AVR3aC were synthesized and cloned into pTOR by Genscript. Fusion genes were flanked by ClaI (5′) and SacII (3′), and internal AscI sites were inserted between the N-terminal effector domain and AVR3aC coding domain. N-terminal domains used are listed in Table 5.

***Phytophthora* infection assays.** *Phytophthora* infection assays were performed according to Schornack et al. [68] with slight modifications.

**Plasmid constructs.** Vector pTOR::Avr3a and pTOR::Avr3a (AAAA-AAA) were obtained from Dr. Steve Whisson [91]. The control construct SP_Avr3aP was synthesized using the signal peptide of *P. infestans* Avr3a, fusing the signal peptide directly to the Avr3a C-terminus (GenBank accession number ACX46530.1). Translocation fusion constructs were synthesized using N-terminal coding sequences of *Albugo* RXLR1 (Gene name: AlNc14C278G10072, GI: 325190660), CRN3 (Gene name: AlNc14C196G8578, GI: 325188975), CHXC9 (Gene name: AlNc14C832G12555, GI: 325193652), and CHXC7 (Gene name: AlNc14C191G8449, GI: 325188831), and the AVR3a C-terminus (GenBank accession number ACX46530.1, GI: 260594559). CHXC9_AAAA_AVR3aC and CHXC7_AAAA_AVR3aC fusion constructs were synthesized by replacing the CHXC motif with a quadA motif. All plasmid suspensions used for *P. capsici* transformation were prepared using the Qiagen Midi Prep kit (Qiagen). For a summary of constructs see Table 5.

**Table 5.** Summary of constructs generated for the *Phytophthora* infection assays.

| Construct | N-Terminal Effector Domain (aa) | C-Terminal Avr3ac Domain (aa) |
|---|---|---|
| SP_Avr3aC | *P. infestans* Avr3a | *P. infestans* Avr3a $KI_{67-147}$ |
| RXLR1_Avr3aC | *A. laibachii* NC14 $RXLR1_{1-52}$ | *P. infestans* Avr3a $KI_{67-147}$ |
| CRN3_Avr3aC | *A. laibachii* NC14 $CRN3_{1-90}$ | *P. infestans* Avr3a $KI_{67-147}$ |
| CHXC9_Avr3aC | *A. laibachii* NC14 $CHXC9_{1-112}$ | *P. infestans* Avr3a $KI_{67-147}$ |
| CHXC9_AAAA_Avr3aC | *A. laibachii* NC14 $CHXC9_{1-112}$, $AAAA_{41-44}$ | *P. infestans* Avr3a $KI_{67-147}$ |
| CHXC7_Avr3aC | *A. laibachii* NC14 $CHXC7_{1-107}$ | *P. infestans* Avr3a $KI_{67-147}$ |
| CHXC7_AAAA_Avr3aC | *A. laibachii* NC14 $CHXC7_{1-107}$, $AAAA_{53-57}$ | *P. infestans* Avr3a $KI_{67-147}$ |

doi:10.1371/journal.pbio.1001094.t005

### Effector Detector Vector Assays

Candidate RXLR effectors were cloned from RXLR to stop; all other candidate effectors were cloned from SP cleavage site to stop into pENTR D-TOPO (Invitrogen) and mobilized into pEDV6 [70]. The resulting effector:pEDV6 constructs were conjugated into Pst DC3000 luxCDABE [69] and Pst DC3000 ΔAvrPto/ΔAvrPtoB [92]. The contribution of an individual effector was assessed by spray inoculating 4- to 5-wk-old short day grown plants as previously described [93].

Growth of Pst DC3000 luxCDABE effector:pEDV6 was calculated by measuring whole plant luminescence using a Photek camera system and normalizing this to plant fresh weight [69].

To assess the virulence of Pst DC3000 ΔAvrPto/ΔAvrPtoB effector:pEDV6, bacterial colony counts were performed as previously described [94].

### Accession Numbers

All Illumina sequence reads generated during this study have been submitted to the Sequence Read Archive at EBI and are accessible under the accession number ERA015557. Individual studies are available with accession numbers ERP000440 (Alias: albugo_laibachii_nc14_dna_sequencing, http://www.ebi.ac.uk/ena/data/view/ERP000440), ERP000441 (Alias: albugo_laibachii_nc14_cdna_sequencing, http://www.ebi.ac.uk/ena/data/view/ERP000441), and ERP000442 (Alias: albugo_laibachii_em1_dna_resequencing, http://www.ebi.ac.uk/ena/data/view/ERP000442).

All contigs and annotations are available through EBI or NCBI. The accession range is from FR824046 to FR827861 (3,816 contigs including annotations) and can be accessed through the ENA browser (http://www.ebi.ac.uk/ena/).

## Supporting Information

**Figure S1 Assembly pipeline using Velvet and Minimus.** Blue boxes with white filling indicate the different Velvet assemblies used. For the Minimus assembler the best contig was used as a seed leading to supercontigs v1. Mis-assemblies in this version were identified and corrected by back aligning all reads (Figure 1A) using MAQ [73] and Bowtie [78]. A self-BLAST was used to avoid redundancy in the contigs. This pipeline was retrained using RACE data of highly heterozygous regions using contig-spanning genes.
(TIF)

**Figure S2 The continuity and quality of the assembled contigs were assessed using CEGMA.** In terms of core eukaryotic genes, 93.6% of a selected set of 248 genes could be detected. While 98.4% and 100%, respectively, of the highly conserved classes 3 and 4 were detected, 86.4% and 89.3%, respectively, of the more divergent classes 1 and 2 were found. Since CEGMA distinguishes between partial and full-length predicted genes, it allows studying the continuity of the genome as well. For the *A. laibachii* Nc14 genome only poorly conserved proteins show an elevated number in partial compared to full-length genes. For groups 2, 3, and 4, all genes predicted were present in full length, indicating that none of the genes was split over contigs. The Illumina-assembled Panda genome and the Sanger/Illumina combined genome of *H. arabidopsidis* were compared (dotted lines). The Panda genome shows high fragmentation of genes, indicated by the distance between partial and complete annotations. The *H. arabidopsidis* genome shows high continuity and a high detection level, although some genes are fragmented in the highly conserved class 4.
(TIF)

**Figure S3 Synteny between the *A. laibachii* Nc14 draft, the *P. infestans* Ia, and the *Py. ultimum* mitochondrial sequence.** The much bigger size of the *Py. ultimum* mitochondrial genome is due to a ~22-kb inverted repeat [10]. Several regions within the *A. laibachii* mitochondrion show direct synteny (red) and inverted synteny (blue), reflecting regions within the *Py. ultimum* inverted repeats. The same region is not inverted in comparison to the *P. infestans* mitochondrion (far left and far right contigs of the *A. laibachii* assembly). Gene annotation in the *P. infestans* genome (annotated by BLAST from the protein sequences) shows that some genes don't show synteny in the *A. laibachii* Nc14 sequence, which is due to unresolved tRNA sequences. Genes in regions with synteny are in particular genes coding for ribosomal proteins, NADH dehydrogenase, and cytochrome C oxidase.
(TIF)

**Figure S4 Annotation of tRNA genes.** The trend shows that copy number correlates with possible codons and amino acid usage in the proteome. Exceptions are the tRNA for the start codon that encodes Met and for the codons that encode Val and Pro.
(TIF)

**Figure S5 Gene prediction pipeline and quality control.** (A) To ensure the best possible gene calls, we combined trained (Augustus), *ab initio* (GeneMark), and consensus (SGP2) gene predictions. Consensus gene calls were made using Evigan based on cDNA evidence. Evidence was generated either by direct alignment of cDNA reads from different stages of infection using Bowtie or by assembling the cDNA using Velvet in combination with Oases or/and using ABySS. (B) For validation of these gene models, a set of 860 annotated KOGs was compiled and tested. Results indicate that 75% of these groups are present in the current annotation. For comparison, 78% of KOGs were present in *P. infestans*, 73% in *H. arabidopsidis*, 42% in *Pl. falciparum*, and 85% in *Ar. thaliana*.
(TIF)

**Figure S6 Genes of "green" or "red" origin present in diatoms and a set of other chromalveolates.** Diagram showing the fraction of genes that are in common between the diatom *Ph. tricornutum* and the tested species that are integrated into the nuclear genome but are of green alga or red alga origin [37]. Bars show the percent of genes present in *Ph. tricornutum*; lines show absolute numbers. Coloured bar below the diagram indicates systematic groups (yellow: fungi; light blue: Apicomplexa; blue: Oomycota; green: green algae; red: red algae; brown: brown algae; lilac: diatoms). The diagram shows that oomycetes still carry about 20% of the green-alga-derived genes that diatoms do. The brown alga *E. siliculosus* carries ~60% of the green alga genes the diatoms do. This might indicate that the ancestral brown algae contained far more green alga genes but these genes were replaced by red alga genes.
(TIF)

**Figure S7 Maximum likelihood trees inferred from comparisons of ITS2 (A) or MORN repeat proteins (B).** A comparison between both trees indicates incongruence between the ITS2 tree and the MORN repeat tree. The ITS2 tree reflects current systematics and indicates that brown algae and diatoms are closer to oomycetes than green algae are. Green algae build an isolated clade from brown algae and chromalveolates. The MORN repeat analyses indicate closer clustering of green algae to brown algae and oomycetes than to diatoms and apicomplexans. These analyses might support a hypothesis that brown-alga-like ancestors accumulated green alga genes. (All bootstrap counts were calculated from 100 replications. Both trees are midpoint rooted.)
(TIF)

**Figure S8 Synteny of a region in *A. laibachii* containing the flagellar inner arm dynein 1 heavy chain alpha (essential for flagellar function) to *Py. ultimum* and *H. arabidopsidis*.** *Py. ultimum* is able to form mobile zoospores while *H. arabidopsidis* isn't. Compared to *A. laibachii* the region is expanded in *Py. ultimum* and *H. arabidopsidis*, but while *Py. ultimum* maintains the flagellar dynein, *H. arabidopsidis* shows a region with synteny but an insertion with homology to transposable elements. LTR sites were annotated using LTR_Finder (labelled in red).
(TIF)

**Figure S9 Gbrowse view of two repetitive regions in the *A. laibachii* Nc14 and *A. laibachii* Em1 genome.** Both regions contain RXLR effector candidates. (A) A highly conserved repeat region with ~10 repeats in Nc14 and ~6 repeats in Em1. (B) A more diverged repeat region with >80 repeats within the Nc14 genome but deletion of the gene-containing region within the Em1 repeats.
(TIF)

**Figure S10 Representation analyses of permutated CHXC motifs in the proteome of selected chromalveolates, red and green algae, and *Ar. thaliana*.** The analyses reveal that only *A. laibachii* and *S. parasitica* contain a significant number of CHXC-motif-containing proteins in the secretome. Only for *A. laibachii* is there a significant enrichment of secreted CHXCs over non-secreted CHXCs. All organisms show a significant number of CXHC proteins, with a high proportion of secreted proteins. CXHC proteins are conserved between genomes and are predominantly enzymes like disulphide isomerase.
(TIF)

**Table S1 Host range of *A. laibachii* Nc14 and *A. laibachii* Em1 tested on 126 *Ar. thaliana* ecotypes.** Twelve ecotypes could be identified that show resistance to only one of the *A. laibachii* isolates, indicating a difference in host range (red: *Ar. thaliana* ecotypes resistant to both *A. laibachii* isolates; orange: ecotypes resistant to one; green: ecotypes susceptible to both).
(DOC)

**Table S2 Genes missing from the CEGMA prediction.** Genes not detected by CEGMA in the *A. laibachii* Nc14 assembly were further analysed and compared to the *P. infestans* genome and *H. arabidopsidis* Emoy2 genome. In all, 12 out of 28 core eukaryotic genes not predicted in *A. laibachii* Nc14 were not predicted in the other two oomycete genomes as well (light grey shading). In addition, three were present in only one of the tested genomes. To rule out the possibility that genes were not predicted because of unusual gene models that cannot be predicted by CEGMA, a BLAST and manual curation was performed on all missing candidates. Eleven could not be identified in the genome as well, while some genes gave multiple results (e.g., ATB binding domains) and were therefore ignored (labelled with "?"). The blast cut-off value was $1e^{-20}$. (Asterisk indicates partial genes.)
(DOC)

**Table S3 Primer pairs used to validate genome continuity and accuracy.** Genomic regions were selected and PCR amplified. The first column gives the primer name and orientation, the second column, primer sequence, the third column, expected length of the PCR product, and the last column indicates if the region could be amplified or not.
(DOC)

**Table S4 Repetitive elements in the *A. laibachii* assembly.** After a search of the library generated with RepeatScout for sequences homologous to transposons, we identified 270 consensus elements showing significant similarity to known transposons. The most abundant in the genome were *mariner* (DNA transposon) and *copia* (LTR retrotransposon) elements. Consensus repeats that do not match any deposited in the NCBI database and do not overlap with Nc14 protein coding genes are either *Albugo*-specific repeats (light grey background) or simple repeats. We identified 191 such consensus sequences that compose about 1% of the assembly.
(DOC)

**Table S5 Distribution of repetitive elements relative to contig length.** Out of the total 3,816 contigs in the assembly, 2,211 contigs have regions with similarity to transposons or other repetitive sequences. Most of these contigs (1,528 contigs) are less than 5,000 bp long.
(DOC)

**Table S6 Distribution of repeats matching telomeric consensus sequences.** Forward and reverse telomeric consensus sequences were identified with RepeatScout. A total of 45 contigs have repeats matching telomeric consensus sequences; amongst these, 25 contigs have telomeric repeats located either at the beginning or at the end of a contig. In all, 5,925 bp of telomeric repeats was assembled.
(DOC)

**Table S7 Characterisation of tRNA genes in the assembled *A. laibachii* contigs.** Type of tRNA gene, number of genes (without and with introns), number of anticodons, type of anticodon, and frequency of usage as a number of stars; 15 tRNA genes were predicted with introns.
(DOC)

**Table S8 Annotations for the brassinosteroid biosynthesis pathway.** The first column gives the enzyme commission numbers (EC numbers) of possible genes. The second column indicates gene names in *Ar. thaliana*. Question marks indicate genes that are difficult to annotate for a certain function (genes that belong to the superfamily of cytochrome P450s). The third column indicates genes identified using the ASGARD annotation pipeline, and the fourth column indicates manual annotation. (GI numbers in brackets.)
(DOC)

**Table S9 Potentially green-alga-derived genes that were identified based on results of a set of green- and red-alga-derived genes present in the diatom *Ph. tricornutum*.** Genes listed here had to be present in the green alga *Ch. reinhardtii* (chloroplast or nuclear genome) but had to be absent from the red alga *C. merolae* and from the fungi *F. oxysporum* and *U. maydis*. (Orange: in *A. laibachii*, *P. infestans*, *Py. ultimum*, *Ph. tricornutum*, *Th. pseudonana*, *Ch. reinhardtii*, and *Pl. falciparum* but not in *H. arabidopsidis* and *E. siliculosus*. Brown: as before but in *E. siliculosus*. Green: shared at least between *Pl. falciparum* and oomycetes. Annotations for identified genes were taken from the list published by Moustafa et al. [37].)
(DOC)

**Table S10 Green alga genes showing homology to *A. laibachii* genes but not to diatome, red alga, brown alga, or fungal genes.** Genes listed here had to be present in the green algae *Ch. reinhardtii* (chloroplast or nuclear genome) and *V. carteri* but had to be absent from the red alga *C. merolae*, the fungi *F. oxysporum* and *U. maydis*, and the brown alga *E. siliculosus* (for the BLAST analyses, an e-value cut-off of $1e^{-20}$ was used; proteins retained by repeating the analyses using an e-value cut-off of $1e^{-5}$ are indicated in blue).
(DOC)

**Table S11 Red alga genes showing homology to *A. laibachii* genes but not to diatom, green algae, brown**

**alga, or fungal genes.** Genes listed here had to be present in the red algae *C. merolae* and *G. sulphuraria* but had to be absent from the green algae *Ch. reinhardtii* (chloroplast or nuclear genome) and *V. carteri*, the fungi *F. oxysporum* and *U. maydis*, and the brown alga *E. siliculosus* (for the BLAST analyses, an e-value cut-off of $1e^{-20}$ was used; proteins retained by repeating the analyses using an e-value cut-off of $1e^{-5}$ are indicated in blue).
(DOC)

**Table S12 Potentially green-alga-derived genes that are present in the diatoms *Ph. tricornutum* and *Th. pseudonana* but not in *A. laibachii* Nc14.** Genes listed here had to be present in the green alga *Ch. reinhardtii* (chloroplast or nuclear genome) but had to be absent from *A. laibachii* Nc14, the red alga *C. merolae*, and the fungi *F. oxysporum* and *U. maydis*. Columns 3–7 show presence/absence in *Py. ultimum*, *P. infestans*, *H. arabidopsidis*, *Pl. falciparum*, and *E. siliculosus* using the same criteria. (a, absent; p, present. Annotations for identified genes were taken from the list published by Moustafa et al. [33].)
(DOC)

**Table S13 Molecular divergence of *A. laibachii* based on all orthologous genes.**
(DOC)

**Table S14 Molecular divergence of *A. laibachii* based on core eukaryotic gene pairs.**
(DOC)

**Table S15 Presence and absence of important metabolic enzymes.** Red indicates absence and green indicates presence of genes. Genes present were annotated or validated in each organism. Remarkable is the absence of all molybdopterin biosynthesis genes, and enzymes using the cofactor, in *A. laibachii* and *Pl. falciparum*. *P. infestans* and *H. arabidopsidis* each lack one of the molybdopterin biosynthesis enzymes but contain molybdopterin-dependent enzymes, which might indicate that other enzymes can compensate for the missing step; in case of B73, the missing enzyme might be replaced by a multifunctional Cnx1 or by high concentrations of Mo inside the cell [95].
(DOC)

**Table S16 List of all annotated proteins of *A. laibachii* that might be associated with pathogenicity.** Annotation and identification were done using Pfam and BLASTP against the NCBI database. Localisation was predicted using a local installation of WoLF PSORT [96]. SignalP 3.0 was used for secretion prediction.
(DOC)

**Table S17 Intraspecies comparison between Nc14 and Em1.** All genes, genes with a predicted secretion signal peptide and without a TM domain, genes representing KOGs, or genes carrying a CHXC, RXLR, or RXLQ motif were compared. The second column in the table indicates heterozygosity (het) within Nc14; the third column indicates heterozygous positions within Em1 (green) or homozygous (hom) SNPs between Nc14 and Em1 (blue). The fourth column shows Em1-specific heterozygous positions or SNPs corrected against Nc14 heterozygous positions carrying the same nucleotide in one of the haplotypes. Frequencies of non-synonymous and synonymous mutations (darker coloured fields, mutations per 100 bp) are almost balanced in the all-gene and KOG gene comparisons, while a comparison of all secreted proteins indicates a 3:1 ratio (non-synonymous:synonymous). RXLRs and, particularly, RXLQs show an imbalance (~2:1), with high variation due to the small sample size. CHXCs, with a ratio of ~5:1, show a significant imbalance in the comparison between Nc14 and Em1. Considering total number and percentage of genes with a ratio of non-synonymous/synonymous $<1$ or $>1$ (light-coloured fields), only KOG genes show a significantly higher number of genes with a value $<1$, while all other classes show more genes with a value $>1$.
(DOC)

**Table S18 CHXC and CXHC candidate genes.** This table gives an overview of all predicted CHXC (white background) and CXHC (grey background) candidates from various species. The first column of the table indicates name of CHXC or CXHC candidates used for the phylogenetic analyses (Figure 9). The second column indicates the species name, and the third column indicates the locus tag within the corresponding genome. The fourth column shows the best BLAST hit against the NCBI nr database with an e-value$<10^{-50}$. The last column indicates the accession number of the best hit.
(DOC)

## Acknowledgments

The authors thank S. Whisson for the pTOR and pTOR::Avr3a constructs; Caroline Dean and colleagues for access to their *Arabidopsis* field plots; Jonathan Urbach for his critical comments and suggestions on the manuscript; Jodie Pike and Matthew Smoker for technical assistance; Michael Burrell for computational assistance; Richard Gibson (European Bioinformatics Institute) for his help in making our data publicly available; and Brett Tyler, Sophien Kamoun, and Marco Thines for discussions.

## Author Contributions

The author(s) have made the following declarations about their contributions: Conceived and designed the experiments: EK JDGJ. Performed the experiments: EK TSL ACK ALB ARS KB. Analyzed the data: EK AG. Contributed reagents/materials/analysis tools: EH DJS DM. Wrote the paper: EK JDGJ. Prepared data for submission to databases and did the submission: DM.

## References

1. Yarwood CE (1956) Obligate parasitism. Ann Rev Plant Physiol 7: 115–142.
2. Goker M, Voglmayr H, Riethmuller A, Oberwinkler F (2007) How do obligate parasites evolve? A multi-gene phylogenetic analysis of downy mildews. Fungal Genet Biol 44: 105–122.
3. Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, et al. (2010) Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. Science 330: 1543–1546.
4. Baxter L, Tripathy S, Ishaque N, Boot N, Cabral A, et al. (2010) Signatures of adaptation to obligate biotrophy in the Hyaloperonospora arabidopsidis genome. Science 330: 1549–1551.
5. Friml J, Wisniewska J, Benkova E, Mendgen K, Palme K (2002) Lateral relocation of auxin efflux regulator PIN3 mediates tropism in *Arabidopsis*. Nature 415: 806–809.
6. Spanu P, Kamper J (2010) Genomics of biotrophy in fungi and oomycetes—emerging patterns. Curr Opin Plant Biol 13: 409–414.
7. Parfrey LW, Barbero E, Lasser E, Dunthorn M, Bhattacharya D, et al. (2006) Evaluating support for the current classification of eukaryotic diversity. PLoS Genet 2: e220. doi:10.1371/journal.pgen.0020220.
8. Cavalier-Smith T (1999) Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. J Eukaryot Microbiol 46: 347–366.
9. Haas BJ, Kamoun S, Zody MC, Jiang RH, Handsaker RE, et al. (2009) Genome sequence and analysis of the Irish potato famine pathogen Phytophthora infestans. Nature 461: 393–398.
10. Levesque CA, Brouwer H, Cano L, Hamilton JP, Holt C, et al. (2010) Genome sequence of the necrotrophic plant pathogen, Pythium ultimum, reveals

original pathogenicity mechanisms and effector repertoire. Genome Biol 11: R73.
11. Thines M, Kamoun S (2010) Oomycete-plant coevolution: recent advances and future prospects. Curr Opin Plant Biol 13: 427–433.
12. Holub EB, Beynon JL (1997) Symbiology of mouse-ear cress (Arabidopsis thaliana) and oomycetes. Adv Bot Res24: 227–273.
13. Soylu S (2004) Ultrastructural characterisation of the host-pathogen interface in white blister-infected *Arabidopsis* leaves. Mycopathologia 158: 457–464.
14. Mims CW, Richardson EA, Holt BF, 3rd, Dangl JL (2004) Ultrastructure of the host-pathogen interface in Arabidopsis thaliana leaves infected by the downy mildew Hyaloperonospora parasitica. Can J Bot 82: 1545–1545.
15. Coates ME, Beynon JL (2010) Hyaloperonospora arabidopsidis as a pathogen model. Annu Rev Phytopathol 48: 329–345.
16. Cooper AJ, Latunde-Dada AO, Woods-Tor A, Lynn J, Lucas JA, et al. (2008) Basic compatibility of Albugo candida in Arabidopsis thaliana and Brassica juncea causes broad-spectrum suppression of innate immunity. Mol Plant Microbe Interact 21: 745–756.
17. Saharan GS, Verma PR (1992) White rusts: a review of economically important species. Rugby (United Kingdom): ITDG Publishing. 380 p.
18. Petkowskia JE, Cunningtona JH, Minchintona EJ, Cahillb DM (2010) Molecular phylogenetic relationships between Albugo candida collections on the Brassicaceae in Australas. Plant Pathol 59: 6.
19. Thines M, Choi YJ, Kemen E, Ploch S, Holub EB, et al. (2009) A new species of Albugo parasitic to Arabidopsis thaliana reveals new evolutionary patterns in white blister rusts (Albuginaceae). Persoonia 22: 123–128.
20. Elena SF, Lenski RE (2003) Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. Nat Rev Genet 4: 457–469.
21. Wilczek AM, Roe JL, Knapp MC, Cooper MD, Lopez-Gallego C, et al. (2009) Effects of genetic perturbation on seasonal life history plasticity. Science 323: 930–934.
22. Holub EB, Brose E, Tor M, Clay C, Crute IR, et al. (1995) Phenotypic and genotypic variation in the interaction between Arabidopsis thaliana and Albugo candida. Mol Plant Microbe Interact 8: 916–928.
23. Borhan MH, Holub EB, Beynon JL, Rozwadowski K, Rimmer SR (2004) The arabidopsis TIR-NB-LRR gene RAC1 confers resistance to Albugo candida (white rust) and is dependent on EDS1 but not PAD4. Mol Plant Microbe Interact 17: 711–719.
24. Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18: 821–829.
25. Sommer DD, Delcher AL, Salzberg SL, Pop M (2007) Minimus: a fast, lightweight genome assembler. BMC Bioinformatics 8: 64.
26. Parra G, Bradnam K, Korf I (2007) CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics 23: 1061–1067.
27. Avila-Adame C, Gomez-Alpizar L, Zismann V, Jones KM, Buell CR, et al. (2006) Mitochondrial genome sequences and molecular evolution of the Irish potato famine pathogen, Phytophthora infestans. Curr Genet 49: 39–46.
28. Farrer RA, Kemen E, Jones JD, Studholme DJ (2009) De novo assembly of the Pseudomonas syringae pv. syringae B728a genome using Illumina/Solexa short sequence reads. FEMS Microbiol Lett 291: 103–111.
29. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, et al. (2003) The COG database: an updated version includes eukaryotes. BMC Bioinformatics 4: 41.
30. Alves JM, Buck GA (2007) Automated system for gene annotation and metabolic pathway reconstruction using general sequence databases. Chem Biodivers 4: 2593–2602.
31. Whitaker JW, McConkey GA, Westhead DR (2009) The transferome of metabolic genes explored: analysis of the horizontal transfer of enzyme encoding genes in unicellular eukaryotes. Genome Biol 10: R36.
32. Keeling PJ (2010) The endosymbiotic origin, diversification and fate of plastids. Philos Trans R Soc Lond B Biol Sci 365: 729–748.
33. Bodyl A, Stiller JW, Mackiewicz P (2009) Chromalveolate plastids: direct descent or multiple endosymbioses? Trends Ecol Evol 24: 119–121; author reply 121-112.
34. Parfrey LW, Grant J, Tekle YI, Lasek-Nesselquist E, Morrison HG, et al. (2010) Broadly sampled multigene analyses yield a well-resolved eukaryotic tree of life. Syst Biol 59: 518–533.
35. Riisberg I, Orr RJ, Kluge R, Shalchian-Tabrizi K, Bowers HA, et al. (2009) Seven gene phylogeny of heterokonts. Protist 160: 191–204.
36. Stiller JW, Huang J, Ding Q, Tian J, Goodwillie C (2009) Are algal genes in nonphotosynthetic protists evidence of historical plastid endosymbioses? BMC Genomics 10: 484.
37. Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, et al. (2009) Genomic footprints of a cryptic plastid endosymbiosis in diatoms. Science 324: 1724–1726.
38. Takeshima H, Komazaki S, Nishi M, Iino M, Kangawa K (2000) Junctophilins: a novel family of junctional membrane complex proteins. Mol Cell 6: 11–22.
39. Gubbels MJ, Vaishnava S, Boot N, Dubremetz JF, Striepen B (2006) A MORN-repeat protein is a dynamic component of the Toxoplasma gondii cell division apparatus. J Cell Sci 119: 2236–2245.
40. Round FE, Crawford RM (1990) The diatoms. Biology and morphology of the genera. Cambridge: Cambridge University Press.
41. Misumi O, Matsuzaki M, Nozaki H, Miyagishima SY, Mori T, et al. (2005) Cyanidioschyzon merolae genome. A tool for facilitating comparable studies on organelle biogenesis in photosynthetic eukaryotes. Plant Physiol 137: 567–585.
42. Campbell WH (1999) Nitrate reductase structure, function and regulation: bridging the gap between biochemistry and physiology. Annu Rev Plant Physiol Plant Mol Biol 50: 277–303.
43. Divon HH, Fluhr R (2007) Nutrition acquisition strategies during fungal infection of plants. FEMS Microbiol Lett 266: 65–74.
44. Hartson SD, Thulasiraman V, Huang W, Whitesell L, Matts RL (1999) Molybdate inhibits hsp90, induces structural changes in its C-terminal domain, and alters its interactions with substrates. Biochemistry 38: 3837–3849.
45. Millson SH, Nuttall JM, Mollapour M, Piper PW (2009) The Hsp90/Cdc37p chaperone system is a determinant of molybdate resistance in Saccharomyces cerevisiae. Yeast 26: 339–347.
46. Baxter I, Muthukumar B, Park HC, Buchner P, Lahner B, et al. (2008) Variation in molybdenum content across broadly distributed populations of *Arabidopsis thaliana* is controlled by a mitochondrial molybdenum transporter (*MOT1*). PLoS Genet 4: e1000004. doi:10.1371/journal.pgen.1000004.
47. Latijnhouwers M, de Wit PJ, Govers F (2003) Oomycetes and fungi: similar weaponry to attack plants. Trends Microbiol 11: 462–469.
48. Sohn J, Voegele RT, Mendgen K, Hahn M (2000) High level activation of vitamin B1 biosynthesis genes in haustoria of the rust fungus *Uromyces fabae*. Mol Plant Microbe Interact 13: 629–636.
49. Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, Harrison T, et al. (2004) A host-targeting signal in virulence proteins reveals a secretome in malarial infection. Science 306: 1934–1937.
50. Kämper J, Kahmann R, Bolker M, Ma LJ, Brefort T, et al. (2006) Insights from the genome of the biotrophic fungal plant pathogen Ustilago maydis. Nature 444: 97–101.
51. Tsang A, Butler G, Powlowski J, Panisko EA, Baker SE (2009) Analytical and computational approaches to define the Aspergillus niger secretome. Fungal Genet Biol 46Suppl 1): S153–S160.
52. Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10: 1–6.
53. Jones DT, Taylor WR, Thornton JM (1994) A model recognition approach to the prediction of all-helical membrane protein structure and topology. Biochemistry 33: 3038–3049.
54. Dodds PN, Rathjen JP (2010) Plant immunity: towards an integrated view of plant-pathogen interactions. Nat Rev Genet 11: 539–548.
55. Bos JI, Armstrong MR, Gilroy EM, Boevink PC, Hein I, et al. (2010) Phytophthora infestans effector AVR3a is essential for virulence and manipulates plant immunity by stabilizing host E3 ligase CMPG1. Proc Natl Acad Sci U S A 107: 9909–9914.
56. Kelley BS, Lee SJ, Damasceno CM, Chakravarthy S, Kim BD, et al. (2010) A secreted effector protein (SNE1) from Phytophthora infestans is a broadly acting suppressor of programmed cell death. Plant J 62: 357–366.
57. Rehmany AP, Gordon A, Rose LE, Allen RL, Armstrong MR, et al. (2005) Differential recognition of highly divergent downy mildew avirulence gene alleles by RPP1 resistance genes from two *Arabidopsis* lines. Plant Cell 17: 1839–1850.
58. Armstrong MR, Whisson SC, Pritchard L, Bos JI, Venter E, et al. (2005) An ancestral oomycete locus contains late blight avirulence gene *Avr3a*, encoding a protein that is recognized in the host cytoplasm. Proc Natl Acad Sci U S A 102: 7766–7771.
59. Win J, Kanneganti TD, Torto-Alalibo T, Kamoun S (2006) Computational and comparative analyses of 150 full-length cDNA sequences from the oomycete plant pathogen Phytophthora infestans. Fungal Genet Biol 43: 20–33.
60. Torto TA, Li S, Styer A, Huitema E, Testa A, et al. (2003) EST mining and functional expression assays identify extracellular effector proteins from the plant pathogen Phytophthora. Genome Res 13: 1675–1685.
61. Oleksyk TK, Smith MW, O'Brien SJ (2010) Genome-wide scans for footprints of natural selection. Philos Trans R Soc Lond B Biol Sci 365: 185–205.
62. Aguileta G, Refregier G, Yockteng R, Fournier E, Giraud T (2009) Rapidly evolving genes in pathogens: methods for detecting positive selection and examples among fungi, bacteria, viruses and protists. Infect Genet Evol 9: 656–670.
63. Kamoun S (2006) A catalogue of the effector secretome of plant pathogenic Oomycetes. Annu Rev Phytopathol 44: 41–60.
64. Soanes DM, Talbot NJ (2008) Moving targets: rapid evolution of oomycete effectors. Trends Microbiol 16: 507–510.
65. Allen RL, Bittner-Eddy PD, Grenville-Briggs LJ, Meitz JC, Rehmany AP, et al. (2004) Host-parasite coevolutionary conflict between Arabidopsis and downy mildew. Science 306: 1957–1960.
66. Raffaele S, Farrer RA, Cano LM, Studholme DJ, MacLean D, et al. (2010) Genome evolution following host jumps in the Irish potato famine pathogen lineage. Science 330: 1540–1543.
67. Schirawski J, Mannhaupt G, Munch K, Brefort T, Schipper K, et al. (2010) Pathogenicity determinants in smut fungi revealed by genome comparison. Science 330: 1546–1548.
68. Schornack S, van Damme M, Bozkurt TO, Cano LM, Smoker M, et al. (2010) Ancient class of translocated oomycete effectors targets the host nucleus. Proc Natl Acad Sci U S A 107: 17421–17426.

69. Fan J, Crooks C, Lamb C (2008) High-throughput quantitative luminescence assay of the growth in planta of Pseudomonas syringae chromosomally tagged with Photorhabdus luminescens luxCDABE. Plant J 53: 393–399.
70. Sohn KH, Lei R, Nemri A, Jones JD (2007) The downy mildew effector proteins ATR1 and ATR13 promote disease susceptibility in Arabidopsis thaliana. Plant Cell 19: 4077–4090.
71. Mehta YR, Zadoks JC (1971) Note on the efficiency of a miniaturized cyclone spore collector. Eur J Plant Pathol 77: 60–63.
72. McKinney EC, Ali N, Traut A, Feldmann KA, Belostotsky DA, et al. (1995) Sequence-based identification of T-DNA insertion mutations in Arabidopsis: actin mutants act2-1 and act4-1. Plant J 8: 613–622.
73. Li H, Ruan J, Durbin R (2008) Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res 18: 1851–1858.
74. Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, et al. (2005) Repbase Update, a database of eukaryotic repetitive elements. Cytogenet Genome Res 110: 462–467.
75. Laslett D, Canback B (2004) ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res 32: 11–16.
76. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, et al. (2009) ABySS: a parallel assembler for short read sequence data. Genome Res 19: 1117–1123.
77. Schulz M, Zerbino DR (2010) Oases: De novo transcriptome assembler for very short reads. Published online: http://www.ebi.ac.uk/~zerbino/oases/.
78. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10: R25.
79. Kent WJ (2002) BLAT—the BLAST-like alignment tool. Genome Res 12: 656–664.
80. Haas BJ, Delcher AL, Mount SM, Wortman JR, Smith RK, Jr., et al. (2003) Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. Nucleic Acids Res 31: 5654–5666.
81. Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M (2008) Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. Genome Res 18: 1979–1990.
82. Parra G, Agarwal P, Abril JF, Wiehe T, Fickett JW, et al. (2003) Comparative gene prediction in human and mouse. Genome Res 13: 108–117.
83. Liu Q, Mackey AJ, Roos DS, Pereira FC (2008) Evigan: a hidden variable model for integrating gene evidence for eukaryotic gene prediction. Bioinformatics 24: 597–605.
84. Li L, Stoeckert CJ, Jr., Roos DS (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res 13: 2178–2189.
85. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32: 1792–1797.
86. Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, et al. (2005) ACT: the Artemis Comparison Tool. Bioinformatics 21: 3422–3423.
87. Xu Z, Wang H (2007) LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. Nucleic Acids Res 35: W265–W268.
88. Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc 2: 953–971.
89. Jones DT (2007) Improving the accuracy of transmembrane protein topology prediction using evolutionary information. Bioinformatics 23: 538–544.
90. Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol 2: 28–36.
91. Whisson SC, Boevink PC, Moleleki L, Avrova AO, Morales JG, et al. (2007) A translocation signal for delivery of oomycete effector proteins into host plant cells. Nature 450: 115–118.
92. Lin NC, Martin GB (2005) An avrPto/avrPtoB mutant of Pseudomonas syringae pv. tomato DC3000 does not elicit Pto-mediated resistance and is less virulent on tomato. Mol Plant Microbe Interact 18: 43–51.
93. Zipfel C, Robatzek S, Navarro L, Oakeley EJ, Jones JD, et al. (2004) Bacterial disease resistance in Arabidopsis through flagellin perception. Nature 428: 764–767.
94. Whalen MC, Innes RW, Bent AF, Staskawicz BJ (1991) Identification of Pseudomonas syringae pathogens of Arabidopsis and a bacterial locus determining avirulence on both Arabidopsis and soybean. Plant Cell 3: 49–59.
95. Schwarz G, Schulze J, Bittner F, Eilers T, Kuper J, et al. (2000) The molybdenum cofactor biosynthetic protein Cnx1 complements molybdate-repairable mutants, transfers molybdenum to the metal binding pterin, and is associated with the cytoskeleton. Plant Cell 12: 2455–2472.
96. Horton P, Park KJ, Obayashi T, Fujita N, Harada H, et al. (2007) WoLF PSORT: protein localization predictor. Nucleic Acids Res 35: W585–W587.
97. Raven JA, Waite AM (2004) The evolution of silicification in diatoms: inescapable sinking and sinking as escape? New Phytol 162: 45–61.
98. Marti M, Good RT, Rug M, Knuepfer E, Cowman AF (2004) Targeting malaria virulence and remodeling proteins to the host erythrocyte. Science 306: 1930–1933.
99. Voegele RT, Wirsel S, Möll U, Lechner M, Mendgen K (2006) Cloning and characterization of a novel invertase from the obligate biotroph *Uromyces fabae* and analysis of expression patterns of host and pathogen invertases in the course of infection. Mol Plant Microbe Interact 19: 625–634.
100. Merchant SS, Prochnik SE, Vallon O, Harris EH, Karpowicz SJ, et al. (2007) The Chlamydomonas genome reveals the evolution of key animal and plant functions. Science 318: 245–250.
101. Cock JM, Sterck L, Rouze P, Scornet D, Allen AE, et al. (2010) The Ectocarpus genome and the independent evolution of multicellularity in brown algae. Nature 465: 617–621.
102. Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, et al. (2008) The Phaeodactylum genome reveals the evolutionary history of diatom genomes. Nature 456: 239–244.
103. Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, et al. (2004) The genome of the diatom Thalassiosira pseudonana: ecology, evolution, and metabolism. Science 306: 79–86.
104. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, et al. (1996) Life with 6000 genes. Science 274: 546563–547.
105. Dujon B, Sherman D, Fischer G, Durrens P, Casaregola S, et al. (2004) Genome evolution in yeasts. Nature 430: 35–44.
106. Kissinger JC, Gajria B, Li L, Paulsen IT, Roos DS (2003) ToxoDB: accessing the Toxoplasma gondii genome. Nucleic Acids Res 31: 234–236.
107. Gardner MJ, Hall N, Fung E, White O, Berriman M, et al. (2002) Genome sequence of the human malaria parasite Plasmodium falciparum. Nature 419: 498–511.
108. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. (2001) Initial sequencing and analysis of the human genome. Nature 409: 860–921.
109. Aparicio S, Chapman J, Stupka E, Putnam N, Chia JM, et al. (2002) Whole-genome shotgun assembly and analysis of the genome of Fugu rubripes. Science 297: 1301–1310.
110. Tyler BM, Tripathy S, Zhang X, Dehal P, Jiang RH, et al. (2006) Phytophthora genome sequences uncover evolutionary origins and mechanisms of pathogenesis. Science 313: 1261–1266.
111. Thoquet P, Gherardi M, Journet EP, Kereszt A, Ane JM, et al. (2002) The molecular genetic linkage map of the model legume *Medicago truncatula*: an essential tool for comparative legume genomics and the isolation of agronomically important genes. BMC Plant Biol 2: 1.
112. Ma LJ, van der Does HC, Borkovich KA, Coleman JJ, Daboussi MJ, et al. (2010) Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. Nature 464: 367–373.
113. Prochnik SE, Umen J, Nedelcu AM, Hallmann A, Miller SM, et al. (2010) Genomic analysis of organismal complexity in the multicellular green alga Volvox carteri. Science 329: 223–226.
114. Matsuzaki M, Misumi O, Shin IT, Maruyama S, Takahara M, et al. (2004) Genome sequence of the ultrasmall unicellular red alga Cyanidioschyzon merolae 10D. Nature 428: 653–657.