Uncovering the molecular mechanism underlying the virulence of *Striga gesnerioides*

Chun Su

Wuhan, China

B.S., Huazhong Agricultural University, 2011

A Dissertation presented to the Graduate Faculty of the University of Virginia in Candidacy for the Degree of Doctor of Philosophy

Department of Biology

University of Virginia

December, 2017 will be conferred

Abstract

Cowpea (*Vigna unguiculata L.*) is the most important food and forage legume in the African Sahel providing essential income and protein nutrition to millions of farmers. While most cowpea cultivars are susceptible to the root parasitic weed *Striga gesnerioides*, cultivar B301 is resistant to all known parasite races except SG4z. When challenged by races SG4 and SG3, the roots of B301 display a hypersensitive response (HR) at the site of parasite attachment followed by death of the invading parasite. In contrast, no visible response occurs in B301 roots parasitized by SG4z and the parasite successfully penetrates the host root cortex, forms vascular connections, and grows to maturity. This study seeks to clarify the molecular mechanism underlying the virulence of *S. gesnerioides* overcoming host resistance.

In order to identify potential parasite components that specifically suppress or elicit host defense responses, I have carried out transcriptome profiling using high throughput RNA-seq analysis on the tissues isolated from parasite haustoria during compatible and incompatible response of resistant (B301) and susceptible (Blackeye) cowpea cultivars with three races of *S. gesnerioides* (SG3, SG4 and SG4z). Comparative transcriptomics and *in silico* analysis revealed the first *S. gesnerioides* transcriptome and several candidate effectors that display differential expression between individual races, that may contribute to race-specific virulence.

To clarify the molecular mechanism underlying the hypervirulence of SG4z, I molecularly characterized a leucine-rich receptor (LRR)-protein kinase (PK) homolog dubbed *SHR4z* (Suppressor of Host Resistance 4z) that is highly expressed in SG4z haustoria and secreted into the host root. I found that overexpression of *SHR4z* in transgenic B301 roots leads to suppression of HR elicitation and loss of host innate immunity by targeting a host BTB-BACK domain containing ubiquitin E3 ligase homolog (VuPOB1). Subsequent silencing *VuPOB1* expression in transgenic B301 roots lowers the frequency of hypersensitive response (HR) while overexpression of *VuPOB1* results in decreased parasitism by SG4z suggesting VuPOB1 functions as a positive regulator of HR and plant innate immunity.

Finally, to explore host effect on *Striga* virulence, I examined global expression changes of the parasite during compatible and incompatible interactions with cowpea. I found that resistant hosts trend to induce the expression of disease resistant genes in SG3 and SG4 relative to susceptible hosts but suppress the expression of genes responsible for development. Comparative analysis on the expression profile of SG4z revealed that, even though SG4z triggers the same response from

B301 and Blackeye, several candidate effectors still display differential expression at the interactions with different hosts. These candidates may contribute to the suppression of host resistance by SG4z.

Overall, this thesis study provides new insight into the role of secreted effectors as part of the strategy used by parasitic weeds to overcome host immunity and complete their life-cycles. It could potentially contribute to the development of novel strategies for controlling *Striga* and other parasitic weeds thereby enhancing crop productivity and food security globally.

Acknowledgements

I am indebted to many people who assisted me over the course of this thesis work. First of all, I would like to thank my advisor, Michael Timko, for his patient guidance, constant inspiration, and continuous support over the years. I am very grateful that Mike took me in on my third year of Ph.D and provide valuable resources and training opportunities for thesis work. Accomplishment of this work is impossible without his critical advice and continuing encouragement. It is his encouraging mentorship that shaped me as a scientist and helped me develop into an independent thinker.

I want to thank current and former members of Timko Lab, who not only gave me a valuable guidance in technical aspects of research, but also created a friendly and joyful working environment in the lab. My special thanks goes to Tatyana Kotova who is like our 'lab mom' taking care of me, Hai Liu whose expertise and friendship helped me overcome the difficult time during my PhD work, Erik Ohlson who always triggered interesting and thoughtful discussion on the research, and Shante Kidd who was the most reliable undergraduate researcher ever.

I also want to express my sincere thanks to Tony Spano, who is always supportive from every aspect of research and life. Tony was always there to help me troubleshoot experiments and made me think through all the possible solutions. He also shared his wisdom on life with me, my many conversations with him always enlightening and drove me to press forward even during the stressful moments at graduate school. His hardworking devotion to science set a good example for me of how a true scientist should conduct themselves.

I would like to thank my former advisor, Lei Li, who taught me useful presentation skills and helped me wrap up first two-year work of Ph.D to my first publication. I also want to thank my current and former committee members: Martin Wu, Christin Danna, Stephen Rich and Stephen Turner, Ben Blackman, and Aaron Mackey. Their critical advice and individual expertise contributed a lot to this work. I would also like to thank Christopher Deppmann for agreeing to sit in on my dissertation defense.

I am thankful for our lab's collaborators: Claude dePamphilis, Jim Westwood, John Yoder, Julie Scholes and Steven Runo. Their intellectual input inspired me a lot in this thesis work. I am thankful for them sharing resources and suggestions valuable to the completion of this work. I

also want to thank Sarah Siegist and Sarah Kucenas for allowing me use their Axiozoom and confocal microscope.

Finally, I want to thank my friends and family for all their support and patience. I thank Che-Hseun, Chao, Susie, Olga, and Sumanth for sharing in the good and bad times over these years. I really appreciate the understanding from my parents and grandma who encouraged me to pursue my dream and gave me unconditional love through my whole life. I also want to thank my wonderful husband, Matthew Pahl, who was there with me every step of the way and supported me intellectually and emotionally. Their presence in my life helped make this work a reality.

Table of Contents

Abbreviation

BAK1: brassinosteroid-associated kinase 1

CDPK: Ca2+-dependent protein kinases

CERK1: chitin elicitor receptor kinase 1

CWDEs: cell wall degrading enzymes

ET pathway: ethylene pathway

ETI: effector-triggered immunity

HR: hypersensitive response

JA pathway: jasmonic acid pathway

LysM: lysine motifs

M/PAMPs: microbial- or pathogen-associated molecular patterns

MAPK: mitogen-activated protein kinase

NBS-LRR receptors: nucleotide-binding-site-leucine-rich-repeat receptors

PTI: PAMP-triggered immunity

RK: receptor kinase

RLCKs: receptor-like cytoplasmic regulatory kinases

RLK: receptor-like proteins

ROS: reactive oxygen species

RRRs: pattern recognition receptors

SA pathway: salicylic acid pathway

SERKs: somatic embryogenesis receptor kinases

T3SS: bacterial type III secretion system

General Introduction

1.1 Plant Immunity

Being sessile, plants are inevitably exposed to various biotic challenges throughout their growth and development. In order to survive, they evolved multidimensional defense system to detect the presence of potential pathogenic threats and respond to specific stressors accordingly. In addition to physical barriers, such as cell walls and constitutively produced phytochemicals, recent molecular research have revealed a sophisticated innate immune system that depends on two interconnected tier of immune signaling to detect and respond to invading parasites (Dangl et al. 2013) (Figure 1.1).

The first tier of plant immune signaling initiates with the perception of evolutionarily conserved pathogenic proteins or structures called microbe or pathogen-associated molecular patterns (MAMPs or PAMPs) (Monaghan & Zipfel 2012). Host transmembrane pattern recognition receptors (PRRs) detect slowly evolving MAMPs/PAMPs, such as bacterial flagellin and fungal chitin, and trigger profound physiological changes in plant cells to minimize pathogen virulence (Monaghan & Zipfel 2012). These changes are usually referred as PAMP-triggered immunity (PTI), which includes dramatic transcriptional reprogramming, bust of reactive oxygen species (ROS), and induction of several hormonal signaling pathways (Andrew F. Bent & Mackey 2007). These changes contribute to the plant mounting an effective innate immune response to constrain pathogen growth.

While PTI is sufficient to stop the invasion of opportunistic parasites, adapted parasites secrete numerous effector proteins to interfere and subvert PTI. However, some effectors, usually called antivirulence (*Avr*) genes, are specifically detected by a group of host intracellular nucleotide-binding leucine-rich repeat (NB-LRR) receptors, which are generally referred as R genes. This recognition will trigger plant second tier of immunity, called effector-triggered immunity (ETI). Different from PTI, ETI usually involves physiologically striking changes, collectively called the hypersensitive response (HR), which is characterized by browning and localized death of the infected cells of the host to constrain pathogen proliferation.

Due to the rapid and highly efficient suppression of pathogen invasion, *R* genes have frequently been the targets of breeding programs for improving crop resistance to parasites (Andrew F. Bent & Mackey 2007). However, the resistance conferred by *R* genes is quite short-lived in the field due to fast evolution of parasites. Effectors can be lost or mutated to avoid the recognition by *R*

gene. Additional effectors can also be generated to interfere interaction between R protein and Avr. This natural selection pressure from parasites, in turn, results in diversifying selection on *R* genes to improve the host fitness. The high rate of turnover of susceptibility alleles and regenerating new resistance alleles drives the evolution of plant immune system in response to highly diverse pathogenic environment.

In this session, I will describe in detail about plant immunity with a subset of the recent discoveries
that solidify the two-tier zig-zag model (Figure 1.1) and expand our understanding of the plant immune system.

1.1.1 Extracellular recognition of pathogen patterns

Plants are subject to attack from various parasites including bacteria, fungi, nematodes, aphids, and weeds. However, the pathogenic proteins recognized by plants are converged to a group of evolutionarily conserved molecular patterns called PAMPs/MAMPs. The classic examples of PAMPs includes bacterial flagellin, elongation factor Tu (EF-Tu), fungal cell wall-derived chitin fragments, peptidoglycans and lipopolysaccharides (Zipfel et al. 2004; Kunze et al. 2004; Kaku et al. 2006; Erbs et al. 2008). Since PAMPs are often molecules that are essential for survival, they tend to be slowly evolving and under strong negative selection (Andrew F Bent & Mackey 2007).

Corresponding to conserved PAMPs, the host plant receptors that recognize them, referred as pattern recognition receptors (PRRs), are usually highly conserved as well (Figure 1.2). PRRs, which are homologous to animal *Toll* receptors, tend to be either receptor kinase (RK) or receptor-like kinase (RLK) that localize to the plasma membrane (Couto & Zipfel 2016). Based on the nature of their ligands, the extracellular domain of PRRs can be subdivided to several groups. Leucine-rich repeat (LRR)-containing PRRs preferentially bind proteins or peptides, such as bacterial flagellin or EF-Tu. FLS2 (flagellin receptor) and EFRs (EF-Tu receptors) in Arabidopsis are typical examples of this subgroup (Zipfel et al. 2004; Zipfel et al. 2006). PRRs with lysine motifs (LysM) bind carbohydrate ligands such as fungal chitin and bacterial peptidoglygan (Gimenez-Ibanez et al. 2009; Wan et al. 2008). The first chitin-binding PRR was identified in rice as the LysM-RLP CEBiP, which homodimerized upon the reception of chitin (Shimizu et al. 2010). The known peptidoglycans receptors include Arabidopsis AtLYM1 or AtLYM3 and rice ortholog OsLYP4 or OsLYP6 (Lannoo & Van Damme 2014).

PAMPs binding to PRRs induces a rapid downstream response mediated by a series of membrane regulatory receptor kinases (Figure 1.2). Co-receptor brassinosteroid-associated kinase 1 (BAK1, also known as SERK3) is associated with LRR-containing PRRs (FLS2 and EFRs) in a PAMP-perception manner in Arabidopsis (Chinchilla et al. 2007; Schulze et al. 2010; Roux et al. 2011). Similarly, the rice BAK1 ortholog SERK2 constitutively combines to LRR-receptor XA21 and links XA21-mediated immune response (Chen et al. 2014). Those LRR-containing PRRs usually contains intracellular kinase domain which auto-phosphorylate or phosphorylate co-receptor kinase like BAK1 upon recognition of PAMPs. However, some PRRs which lack signaling kinase domains require additional RLK like SUPPRESSOR OF BIR1-1 (SOBIR1) or SOBIR1-like LRR-receptor kinases to ensure binding with BAK1 (Gust & Felix 2014; Liebrand et al. 2014). Similar to the role of BAK1 with LRR-type PRRs, LysM-containing PRRs associate with a different regulatory co-receptor CHITIN ELICITOR RECEPTOR KINASE 1 (CERK1) to activate downstream PTI. Rice LysM-RLP CEBiP homodimers binds CERK1 to form hetero-complex following chitin binding. Arabidopsis AtLYM1 and AtLYM3 also binds to CERK1 during peptidoglycan recognition to mediate antibacterial immune response (Gimenez-Ibanez et al. 2009).

PTI signaling is transmitted from the plasma-membrane to the cytoplasm via a group of receptor-like cytoplasmic regulatory kinases (RLCKs) (Figure 1.2). Arabidopsis BOTRYTIS-INDUCED KINASE 1 (BIK1) is the best-studied example of RLCKs. BIK1 associates with FLS2 at the plasma membrane during resting conditions. Upon flg22 elicitation, BAK1 forms an elicitor-dependent complex with FLAGELLIN SENSING 2 (FLS2) and immediately phosphorylates BIK1 to dissociate BIK1 from the PRR complex (Lu et al. 2010). BIK1 is also required for elf18, AtPep1 and chitin activated immune response and functions as a hub to integrate multiple PRR mediated signaling (Zhang et al. 2010). There are other RLCKs, like Pattern-Triggered Immunity Compromised Receptor-like Cytoplasmic Kinases (PCRKs) and PBS1-like kinase (PBL) proteins, also mediated PTI signaling (Tang et al. 2017). Most of them specifically mediate immune responses triggered by certain groups of PAMPs and vary in their ability to activate distinct branches of PTI signaling. The large repertoire of RLCKs enable plants to have more flexibility in their immune system in face of various biotic challenges.

The immediate downstream signaling cascades of the PRR complex include a dramatic rise of cytosolic Ca^{2+} levels and rapid production of reactive oxygen species (ROS) (Boller & Felix 2009; Seybold et al. 2014) (Figure 1.2). Although identity of the channel(s) responsible for this

Ca^{2+} burst is still unclear, it has been found that the elevation of cytosolic Ca^{2+} transcriptionally regulates the genes in salicylic acid (SA) biosynthesis and trigger SA signaling during the PTI through a group of Ca2+ sensor proteins, such as calmodulin (CaM) and Ca^{2+}-dependent protein kinases (CDPK). ROS production was regulated by NADPH oxidase RESPIRATORY BURST OXIDASE HOMOLOGUE PROTEIN D (RBOHD) which is directly phosphorylated by several RLCKs, including BIK1 and related PBLs, upon PAMP elicitation (Kadota et al. 2014; Li et al. 2014). The activity of RBOHD was also regulated by CDPK-mediated phosphorylation (Kobayashi et al. 2007; Ogasawara et al. 2008).

Another main PTI signal following activation of the PRR complex is activation of mitogen-activated protein kinase (MAPK) pathway which conveys immune signaling to the nucleus, by phosphorylating multiple transcription factors and resulting in transcriptional reprogramming to establish PTI (Meng & Zhang 2013) (Figure 1.2). The MITOGEN-ACTIVATED PROTEIN KINASEs (MAPK) involved in PTI include MPK3 and MPK6 which are activated by the MAPK kinases (MAPKKs) MKK4 and MKK5 (Asai et al. 2002), but the particular MAPK kinases kinases (MAPKKKs) acting upstream of MKK4 and MKK5 and RLCKs that link PRR elicitation to the MAPK pathway are still unknown(Couto & Zipfel 2016). The targets acting downstream of the MAPK cascade in PAMP triggered immunity response are transcriptionally induced by MAPK-activated transcription factors. Several downstream transcriptional factors have been identified: including the bZIP transcription factor VIP1 which activates *PR1* gene expression (Djamei et al. 2007), WRKY33 which activate the expression of camalexin biosynthetic genes (Mao et al. 2011) and EFR104 which regulates ET signaling (Bethke et al. 2009).

Although the induction of PTI is essential for plant defense, it must be strictly regulated to maintain immune homeostasis and plant fitness. There are several different layers to control PTI responses, including disrupting the recruit of co-receptor kinases to form PRR complex, deactivating or degrading the PRR complex, monitoring cytoplasmic kinase signal transducing pathways, and control of transcriptional reprogramming (Couto & Zipfel 2016). Those were generally achieved either by various kinases/phosphatases regulating the phosphorylation status of the PRR complex and RLCKs or by E3 ubiquitin ligases mediated ubiquitination and degradation of PTI components. Several kinases/phosphatases and E3 ubiquitin ligases have been identified to be involved in the negative regulation of PTI. For example, the rice protein phosphatases type 2C (PP2C/XB15) directly binds to and dephosphorylates the pattern recognition receptor XA21. As a result, PP2C/XB15 negatively regulates the XA21-mediated

innate immune response (Park et al. 2008). Similarly, several members of the Plant U-box (PUB) family of ubiquitin E3 ligases are known to negatively regulate PTI responses. PUB12 and PUB13 have been implicated in the ubiquitination and degradation of FLS2 upon flg22 treatment (Lu et al. 2011). The *pub22/pub23/pub24* triple mutant displayed impaired downregulation of responses triggered by PAMPs. It results in prolonged ROS production, increased MPK3 activity, and transcriptional upregulation of defense-marker genes *PR1* and *PDF1.2* in triple mutant (Trujillo et al. 2008).

In summary, PTI is a highly complex and tightly regulated response. PRRs dynamically associate with a high variety of transmembrane co-receptors and RLCKs to initiate signaling cascades to promote immunity against opportunistic parasites.

1.1.2 R gene mediated plant defense

Throughout the history of agriculture, crop diseases caused devastating economic damage and severely antagonized human civilization. In addition to pesticide applications and field rotation, over centuries of breeding practice and genetic research in nearly all crops discovered a group of ancient intracellular receptors with NBS-LRR domain, referred as the resistant genes (*R* gene).

1.1.2.1 Structure of R protein

R genes typically encode proteins containing nucleotide-binding site domain (NBS) and the leucine-rich repeat domain (LRR). While NBS is suggested to bind ATP and initiate protein conformation change for signal activation, C-terminal LRR domain is predicted to form an intra-molecular protein interaction surface with NBS for auto-inhibition of signaling or/and an inter-molecular interaction for effector recognition (Głowacki et al. 2011).

Besides these two essential domains, plant R proteins are further grouped to two subfamilies based on N-terminal domain which usually exhibits either a toll/interleukin-1 receptor (TIR) domain or a coiled coil domain (CC). These two subgroups differ in the downstream signaling components that they recruit. TIR-NB-LRRs require Enhanced Disease Susceptibility 1 (EDS1), while most CC-NB-LRRs require Non-race specific Disease Resistance 1 (NDR1) for activation of immune responses. LRR domain provides a protein binding site to recognize pathogenic elicitors (DeYong & Innes 2006).

1.1.2.2 Recognition of effector by R protein

Direct interactions between R proteins and effectors have been assumed from gene-for-gene specificity and coevolution of the *R* genes and *Avr* genes. The first direct interaction between NBS-LRR proteins and pathogen effectors was shown in the rice blast fungus *Magnaporthe grisea* (Jia et al. 2000). Both yeast hybrid experiment and in vitro immunoprecipitation detected that functional portion of effector AVR-Pita directly interacts with the LRR-like domain of *R* gene (Pi-ta) specific resistant to *M. grisea* (Jia et al. 2000). In this way, the ligand-receptor model has been proposed based on direct physical interactions between LRR domain of R protein and pathogenic effectors.

However, direct physical interaction between effector and R protein is not always the case. R proteins can also recognize effectors indirectly by guarding host proteins that physically interact with effectors. In the guard model of plant immunity, the interaction between effector and guarded effector target (also called the guardee) is surveilled by the R protein (Dangl & Jones 2001) A classic example of a host 'guardee' mediator is RIN4, which was directly manipulated by *Pseudomonas syringae* type III effector molecules AvrRpm1 and AvrB. The modification of RIN4 is recognized by NBS-LRR protein RPM1 and triggered plant defense in Arabidopsis (Mackey et al. 2003; Mackey et al. 2002). Another example of guardee-mediated detection of parasites was described in the regulation of RPS5, a CC-NBS-LRR protein, which was inhibited by a host protein kinase, PBS1, at resting stage. The activity of RPS5 is induced upon *P. syringae* inoculation since a bacterial protease AvrPphB cleaves PBS1 and releases the inhibition of RPS5 from LRR domains (Zhang et al. 2010). Thus, the effector target (guardee) acts as a bridge between effectors and R protein and is an essential regulator in R protein-mediated plant resistance (ETI). In the absence of R protein, guardee protein is an essential part of PTI. Being targeted by effectors leads to the comprising of PTI and proliferation of parasites. Therefore, due to their dual functions (involvement in both PTI and ETI), the guardee protein is under conflicting selection pressures from R protein and effector. On one hand, guardee protein is prone to bind effectors so that R protein can recognize the association between guardee proteins and effectors to trigger ETI. From the other hand, guardee protein is selected against binding to effectors due to the fact that being targeted by effector would compromise the PTI when R protein is not present.

However, this conflicting selection pressures on the effector interaction surface of the guardee/host targets can be released by a new model --- decoy model. In the decoy model, host targets act as decoys to trap the effectors and are dispensable for the virulence activities of

effectors in plants (van der Hoorn & Kamoun 2008). Thus, in contrast to guardee model, the decoys are only involved in the ETI and selected towards binding effectors for the recognition by R protein. The interaction between decoys and effectors at absence of R protein will not affect PTI. The decoy model was first proposed based on targets of the *P. syringae* pv tomato DC3000 effector protein AvrPto (Zipfel & Rathjen 2008). AvrPto is a kinase inhibitor that suppresses PTI by directly targeting pattern recognition receptors FLS2 and EFR (Xiang et al. 2008). AvrPto also interacts with a mimic of the FLS2, Pto, which surveilled by R protein, Prf, and elicits strong defenses through ETI. Pto is a decoy subjected to the regulation of Prf but does not function independently to promote plant immunity when Prf is absent (Zipfel & Rathjen 2008).

1.1.2.3 Signaling following recognition of effector

It is well known PTI and ETI share many signaling components after recognition of PAMPs or effectors. Both pathways result in increased ROS, plant hormonal signaling (SA, JA, ET), and transcriptional reprogramming. Rapid ROS production is regulated by NADPH oxidase AtRbohD upon perception of PAMPs in PTI (Zhang et al. 2007). The recognition of effectors by R proteins elicits much higher magnitude of ROS production which is also largely AtRbohD-dependent (Torres et al. 2006). Similarly, both PTI and ETI trigger increased hormone signaling to alter the physiology of the plant and mount an effective immune response. Delayed-dehiscence 2 (*DDE2*) encodes one of the key enzymes in the JA biosynthesis pathway. SA induction-deficient 2 (*SID2*) and phytoalexin deficient 4 (*PAD4*) are essential components of SA signaling. Ethylene Insensitive 2 (*EIN2*) is implicated in cross-talk between ET, JA and SA. Quadruple mutants (*dde2/ein2/pad4/sid2*) for these pathways have impaired signaling in response to both flg22-triggered immunity (flg22-PTI) and AvrRpt2-triggered immunity (AvrRpt2-ETI), which suggests SA, ET, and SA signaling all contribute positively to PTI and ETI (Tsuda et al. 2009). Furthermore, a transcriptome-wide study revealed that there is an significant overlap between genes induced by PTI and genes induced by ETI (Navarro et al. 2004). All these evidence implied that ETI and PTI shared induced downstream signaling.

However, immune responses in ETI are more robust, prolonged and rapid than those provoked by PTI (Tsuda & Katagiri 2010). Although common signaling machinery is employed differently in PTI and ETI, some specific features of ETI following effector perception are worthy of mention.

R genes coordinate many transcriptional changes during the immune response. Following the effector recognition, NB-LRR R proteins are activated by conformational changes. This

conformational change is regulated by intramolecular interaction in which the negative regulator domain LRR dissociates NB-ARC domain to allow its binding to ATP (Rairdan & Moffett 2006). Immediately following activation, most R proteins translocate to nucleus and regulate immune related transcriptional changes directly. The NB-LRR protein SNC1 associates with the transcriptional corepressor TPR1 (Topless Related 1) to inhibit the expression of negative immune regulators and trigger defense responses (Zhu et al. 2010). Similarly, the family of intracellular mildew A (MLA) R proteins from barley interferes with the WRKY transcriptional repressor function upon recognition of the fungal Avr10 effector, linking the effector-specific ETI to the basal resistance responses (Shen et al. 2007). More interestingly, the Arabidopsis RRS1-R NB-LRR protein carries a C-terminal WRKY DNA binding domain itself and forms a R protein complex with another NB-LRR protein RPS4. While RRS1-R behaves as a decoy in the nucleus to detect the effectors (PopP2 and AvrRps4) that originally target other WRKY proteins, RPS4 activates defense upon such perception (Sarris et al. 2015). The direct involvement of R protein in transcriptional re-programming after effector recognition may partially contribute to rapid response of ETI.

One specific phenotype associated with ETI is the hypersensitive response (HR) which is a form of programmed cell death localized at the site of attempted pathogen invasion. Although the signaling events leading to HR in plants after effector recognition is not fully elucidated, two separate signaling modules, EDS1 and NDR1, have been found to play a central role in *R* gene-mediated pathways by integrating redox signals downstream of NADPH oxidase AtRbohD to SA accumulation. In turn, ROS and SA act synergistically to drive HR. It has been reported that EDS1, homologous to eukaryotic lipases, is recruited by TIR-NB-LRR proteins in multiple plant species and acted upstream of oxidative burst and the accumulation of salicylate (Falk et al. 1999; Mateo et al. 2004; G. Hu et al. 2005). It also regulates expression of pathogenesis-related (*PR*) genes which are the signatures of plant response to biotic stress (G. Hu et al. 2005). In comparison to EDS1, NDR1 mediates the CC-NB-LRR signaling pathway. As a plasma membrane-localized protein, it interacted with effector target RIN4, which also localizes to plasma membrane. Since RIN4 has been found to interact with at least two R proteins (RIN4-RPS2 and RIN4-RPM1), in addition to the interaction between RIN4 and NDR1 suggests as an essential form of switch for activating HR (Day et al. 2006).

1.1.3 Gene-for-gene theory and beyond

Due to specific one-for-one relationship between *R* gene and *Avr* gene defined in early years, a gene-for-gene model has been proposed, suggesting that each parasite *Avr* may only be identified by its counterpart *R* gene in host (Flor 1971). A small number of *R-Avr* gene pairs have been isolated in parasite-plant interactions. The comprehensive research in *Cladosporium fulvum* - tomato system has provided support for the gene-for-gene hypothesis, where four resistance genes, *Cf-2*, *Cf-4*, *Cf-4E* and *Cf-9* recognize a specific counterpart *Avr* gene, *Avr2*, *Avr4*, *Avr4E* and *Avr9*, respectively (van Esse et al. 2008). Similarly, *avrRpt2-RPS2* gene-for-gene interaction was also demonstrated in *P. syringae*-Arabidopsis pathosystem. However, the majority of gene-for-gene interactions are not molecularly defined, but implied by observation that a given host genotype is specifically resistant to certain population or isolates of parasites. Our lab has previously shown that the Cowpea cultivar B301 is resistant to seven races of *S. gesnerioides* in the field, but resistant gene *RSG3-301* can trigger HR only when host is infected with a specific population (SG3) derived from Nigeria and Niger. Silencing *RSG3-301* in B301 was not able to compromise the host resistance to other races. It is proposed that each race of *Striga* has their own Avr that interacts with specific counterpart R protein (Li & Timko 2009). Therefore, isolation of more *Avr-R* gene pairs is required to evaluate gene-for-gene hypothesis.

Effector-triggered immunity was not as simple as direct one-to-one interaction implied by the gene-for-gene hypothesis. Arabidopsis R protein RPM1 recognized two completely unrelated Avr proteins, AvrRpm1 and AvrB, from *P. syringae* (Grant et al. 1995). Later Dangl et al (2006) discovered that RPM1 does not directly recognize the molecular structure of AvrB and AvrRpm1 (Dangl & Jones 2001). Instead, it monitors an 'guardee' protein RIN4 which is modified by AvrB and AvrRpm1 during the infection (Mackey et al. 2002). Finally, phosphorylation of RIN4 activates RPM1 resistance. Thus, new progress leads us to re-interpret the gene-for-gene hypothesis of plant resistance and expands this old concept to include indirect multi-specific resistance.

In summary, plant defense is complex and multi-level response. It relies on membrane-anchored PRRs or/and cytoplasmic NBS-LRR receptors to recognize pathogenic molecules and trigger synergistic signaling pathways.

1.2. Parasite Virulence

1.2.1 Classification of Parasitism

Parasitism is a relationship between two different species where the parasite benefits at expense of the host. Plants, as a general host, are attacked by various parasites, including bacteria, fungus, oomycetes, nematodes and angiospermous parasitic plants. Based on the nutrition mode, parasites are categorized into necrotrophs and biotrophs. While necrotrophic parasites actively damage their host and utilize dead tissues, biotrophic parasites maintain the viability of their host and obtain nutrients from living cells.

Necrotrophic parasites can be bacteria or fungi. They actively damage the host, resulting in extensive necrosis, tissue maceration, and plant rot (van Kan 2006). The pathogenesis process begins with the secretion of lytic enzymes into extracellular region to decompose the plant tissue. Subsequently necrotrophs eject necrosis-inducing factors and toxins into the plant and kill the cell rapidly. Finally, they feed on dead cell contents. The well-studied necrotrophs include the fungal parasites *Botrytis cinerea* and *Sclerotinia sclerotiorum*, and the bacterial pathogen *Erwinia carotovora* (van Kan 2006; Friesen et al. 2008).

In contrast to necrotrophs, biotrophs use sophisticated methods during infection to balance their pathological virulence and likelihood of detection by their host. The plants infected with biotroph usually survive although they often fail to thrive and may abort reproductive phases. When resistant hosts are inoculated with biotroph, effector-triggered immunity (ETI) usually results in hypersensitive response, which features controlled local cell death at feeding or infected site. Since biotrophs depend on the living cells to support its lifecycle, the local cell death not only prevents the further invasion into the host tissue, but also cuts the parasite off the further nutrient supply from living host cells. *Pseudomonas syringae* is among mostly studied biotrophic bacterial parasites (Jin et al. 2003). It can infect a wide range of species and has been represented as a model in the molecular study of biotrophic parasites.

In addition to biotroph-necrotroph classification, parasites can also be categorized into either facultative parasites or obligate parasites based on their degree of host dependence. While facultative parasites occasionally parasitize plants when the option is available to them, obligate ones strictly rely on the host to complete their life cycles (Westwood & Yoder 2010). This classification has been widely discussed in nematode and parasitic plants. The phylum Nematoda contains many facultative parasites of plants. One example is *Bursaphelenchus xylophilus*, which feeds on living pine tree epithelial cells and produce large numbers of progeny that causes pine

wilt. After the host pine tree dies and living cells are no longer available, *B. xylophilus* switches to feed on fungus and completes its typical free-living life cycles in dead wood (Fukushige 1991). In contrast to *B. xylophilus*, *B. cocophilus* has been defined as an obligate plant parasite and causes red ring disease in coconuts. This species cannot be cultured on any fungi media (Gerber & Giblin-Davis 1990) . In the Angiospermae, both facultative and obligate parasitic plants were observed in several lineages. The Parasitic Plant Genomic Project (PPGP) initiated a comparative transcriptomic analysis of three root parasitic genera from Orobanchaceae (*Triphysaria, Striga, and Orobanche*). Among them, *Triphysaria is* a facultative hemiparasite which can live as regular plant without host attachment, while *Striga* and *Orobanche* require a host plant to provide nutrient supply for their growth. Parasitic plants are further divided into hemiparasite and holoparasites according to their photosynthetic competency (Westwood & Yoder 2010). While hemiparasites still retain photosynthesis ability, holoparasites are completely obligatory and undergo evolutionary reduction in their chloroplast genome (Wickett et al. 2011; Bungard 2004).

1.2.2 Structural and molecular basis in parasitism

The process of parasitism is facilitated by specific structures and an array of molecules including secreted enzymes, virulence effector proteins and even transferable RNAs. They work cooperatively or independently to circumvent host detection and suppress host defense.

1.2.2.1 Structural basis for parasitism

The most well studied parasitic structure is bacterial type III secretion system (T3SS). This transmembrane nano-machine injects virulence effector proteins directly to host cells and is highly conserved among many bacterial parasites of animals and plants (Mateo et al. 2004). Different from this membrane-embedded protein complex in bacteria, fungi evolved a highly specialized invasion organ called the “haustorium”. The haustorium is an intracellular hypha that projects out of the fungus cell surface and is surrounded by a host-derived extra-haustorial membrane (Bushnell 1972). Fungi used this highly differentiated structure to penetrate their host cell walls and facilitate nutrient transfer (Szabo & Bushnell 2001). Parasitic plants also have a functionally similar organ called the haustoria, but the structure of organs differs in these two groups. Instead of being an unicellular hypha, the haustorium of parasitic plants is a multicellular organ which penetrates between the host cells and finally builds vascular connections between host and parasite (Riopel & Timko 1995; Mayer 2006). Similar to haustoria, nematodes protruded a sclerotized stylet into the cells of their host plant to form a feeding site and take-up nutrients from or release toxins.

1.2.2.2 Cell wall degradation enzyme

In spite of diverse infection structures, all plant parasites first need to breach the cell wall of the host plant. Necrotrophic fungi degrade plant cell wall by producing a large array of cell wall degrading enzymes (CWDEs) like polygalacturonases, hemicellulases and cellulases (ten Have et al. 2002). Those enzymes are assisted by acetyl esterases and pectin methylesterases, which extensively disrupt the integrity of plant cell walls but also are recognized as PAMPs to trigger the initial defenses of the plant (Bellincampi et al. 2014). Similar to fungi, cyst nematodes secrete a mixture of cellulase and pectinase enzymes to dismantle plant cell walls. Pectate lyases PEL-1 from potato-cyst nematode *Globodera rostochiensis* bears high sequence similarity with similar enzymes of bacterial and fungal origin and provides the first evidence of parasitic degradation of pectin in plant cell walls by an animal (Kudla et al. 2007). Recently, genomic research in parasitic plants revealed that differentially upregulated putative parasitism genes in haustorial development were also enriched in cell wall modification process (Yang et al. 2015). Therefore, although plant parasites are diversified ranging from prokaryotes to eukaryotes, the mechanism to invade plant first defense requires breaking down plant cell walls.

1.2.2.3 Effectors

Besides suppressing cell wall-based defense by CWDEs, there are groups of secreted proteins called effectors that directly interact with plant host defense targets and suppress host immunity. According to the two-tier plant immunity model, effectors generally accumulate after the plant basal defense is triggered (Dangl & Jones 2001). Effectors suppress PTI by reprogramming the host cellular response and regulating expression change of immunity genes to enhance the compatible parasitic interaction with host. In the case of host plants that carry the *R* gene, subsets of effectors (*Avr* genes) induce incompatible host-microbe interactions (ETI) and cause HR-like cell death localized near the parasite and ultimately death of parasites. Therefore, current research on pathogenic effectors has generally focused two interconnected functions: (1) the role of effectors in virulence by targeting PTI defense genes and (2) the avirulence activity of effectors on pathogen growth by triggering host resistance.

1.2.2.3.1 Bacterial effectors

Molecular research has revealed a handful of effectors in bacterial parasites. *P. syringae pv. glycinea* type III effector (T3E) AvrB was one of earliest cloned plant pathogen effectors (Staskawicz et al. 1987). It interacts with plasma membrane anchored host protein RIN4 and promotes its phosphorylation to facilitate pathogen growth in genotypes lacking R protein RPM1

(Mackey et al. 2002). However, when *RPM1* is presented, RPM1 will recognize AvrB-mediated phosphorylation of RIN4 and confer resistance against *P. syringae.* Similar to AvrB, non-homologous effector AvrRpm1 can also induce the RIN4 phosphorylation and trigger *RPM1*-mediated resistance response (Mackey et al. 2003). Different from the phosphorylation activity of AvrB and AvrRpm1, AvrRpt2 from *P. syringae pv*. tomato exhibits cysteine protease activity and cleave RIN4 into two major proteolytic products. The degradation of RIN4 was monitored by *R* gene *RPS2* and eventually activates *RPS2*-mediated defense pathway (Lim & Kunkel 2004; Axtell et al. 2003). Thus, even though the molecular sequences and biochemical activity were diverse, effectors are converged to a certain key regulator (RIN4 as an example here, also called 'guardee' protein), which provides a mechanistic link to R protein and triggers specific *R* gene mediated defense pathway.

In addition to targeting one regulator hub, effectors also coincidentally subvert MAPK pathway to block defense signaling, which is a conserved virulence strategy shared by a wide range of parasites. *P. syringae pv. tomato 0288-9* effector HopAI1 directly interact with multiple MAP kinase kinases (MKK) and dephosphorylates MKK3, MKK4 and MKK6 to inactivate PAMP-induced defense (Zhang et al. 2007). Similarly, *P. syringae pv.* tomato DC3000 effector HopF2 suppresses Arabidopsis innate immunity by blocking MKK5 activation in the MEKK1/MEKKs-MKK4/5-MPK3/6 cascade and a upstream plasma membrane-localized receptor-like kinase BAK1 (Y. Wang et al. 2010; Zhou et al. 2014). However, no R genes are currently known to trigger ETI in response to HopAI1 and HopF2.

Furthermore, a special group of effectors called TAL (Transcription Activator-Like) effectors from plant pathogenic *Xanthomonas spp.* specifically target promoters of defense regulatory genes in nucleus. AvrBs3 from *X. campestris pv. Vesicatoria* is one of the best-characterized TAL proteins that induces cellular hypertrophy and promotes bacterial fitness in susceptible *Capsicum annuum* (Kay et al. 2007). It specifically binds to UBA box at promoter region of nodulin MtN3 family protein UPA16 and induces its expression. In resistant C. annuum, AvrBs3 also binds the promoter of the pepper resistance gene Bs3 and triggers HR (den Ackerveken et al. 1996; Schornack et al. 2004; Römer et al. 2007).

1.2.2.3.2 Effectors of fungus and oomycetes

In contrast to bacteria, research on effectors of fungi and oomycetes have been delayed given the difficulties of cloning and manipulating *Avr* genes without well annotated genomes. With recent

advancement of next generation sequencing technology, *in silico* identification and characterization of effector catalogs has recently become a big trend in the research of fungi and oomycetes. By 2016, it has been estimated that more than 3000 effector candidates have been identified from 11 plant fungal and oomycete parasites with available genomes and dozens of effector databases have been established (Sonah et al. 2016). Although the computational tools used for effector identification vary from one to another, one common criterion is prediction of the secretome since effector proteins are mostly secretory proteins that manipulate host cells to suppress defense mechanisms.

Based on the location of their function, fungal and oomycetes effectors were further grouped to two distinguishable categories. Apoplastic effectors are secreted into the plant extracellular space while cytoplasmic effectors are translocated into the plant cell. Apoplastic effectors are usually comparatively small (less than 300 amino acids). They contain an N-terminal signal peptide and functional motif at the C-terminal. Effectors that are secreted into the extracellular matrix mainly protect the parasites from plant hydrolytic enzymes, such as proteases, glucanases, and chitinases. Well-studied examples include Avr2 and Avr4 from *Cladosporium fulvum* and EPIC1 and EPIC2 from *Phytophthora infestans*. *C. fulvum* Avr2 directly interacts and inhibits tomato proteases RCR3 and PIP1 to enhance host susceptibility (van Esse et al. 2008; Shabab et al. 2008). One way that resistant plants detect Avr2 is by expressing a decoy of *RCR3* that is surveyed by R gene *Cf-2*, which upon detecting the interaction between the decoy and Avr2 will trigger HR (Dixon et al. 1996). Similarly, Oomycete *P. infestans* proteases inhibitors *EPIC1* and *EPIC2* are also upregulated during the infection of tomatoes. EPIC2 targets tomato proteases PIP1 and C14 to inhibit plant basal defense, but no *R* genes have been reported to mediate resistance specific to EPIC effectors (Tian et al. 2007). Different from protease inhibitors, Avr4 binds to chitin and protects fungal cell walls against tomato chitinases and β-1,3-glucanases(van den Burg et al. 2006). However, Avr4 can also be recognized by membrane-anchored R protein Cf-4 (Kruijt et al. 2005). It has been recently shown that Avr4-Cf4 interaction recruits cell surface receptor-like kinases BAK1/SERK3 to bind Cf4 and induce the downstream resistance and trigger endocytosis of the fungal cell wall (Postma et al. 2016).

Similar to apoplastic effectors, the N-terminal of cytoplasmic effectors also contains a signal peptide for extracellular secretion. In addition, it also includes highly conserved amino acid motifs to enable translocation from extracellular matrix to host cell (Whisson et al. 2007). The most common motif is RxLR-dEER, which is present in over 700 effector candidates from two

Phytophthora species, *P. sojae and P. ramorum* (Jiang et al. 2008). However, little sequence similarity is shared outside this conserved motif, suggesting that this family of effectors are rapid evolving. The molecular functions of several RxLR effectors have been characterized in last two decades. They can interfere with plant immunity and growth through multiple mechanisms, including inhibiting plant immunity protease (Bozkurt et al. 2011), disrupting cell integrity (Bouwmeester et al. 2011), impeding host auxin physiology (Evangelisti et al. 2013) and so on. One intensively studied RxLR effector is Avr3a which translocates into the host cells and stabilizes a host U-box E3 ligase CMPG1 to suppress elicitor protein INF1-triggered cell death (Bos et al. 2010). At the same time, recognition of Avr3a by the R protein R3a in resistant plant induces strong effector-triggered cell death.

1.2.2.3.3 Effectors of nematode

Nematodes are an emerging research focus of parasitism over the last decade. Effectors of nematode are proposed to be mainly synthesized and stored in elaborate secretory gland cells and injected into host plants cells through a specialized stylet structure (Jasmer et al. 2003). By isolating genes expressed preferentially in the gland cells of parasitic stages, researchers found a dozens of deduced proteins were predicted to be extracellular and comprise the parasitism candidates from soybean cyst nematode *Heterodera glycines* (Gao et al. 2001). Based on those candidates, Hewezi *et al* (2008,2010) was able to clone two effector genes, *CBP* and *10A06*, and heterogeneously express them in *Arabidopsis* to test its virulence. Transient expression of both effectors affected plant morphology and increased susceptibility to *H. schachtii*. While CBP effector directly interacts with *Arabidopsis* Pectinesterase inhibitor 3 (PME3) to modify the plant cell wall and eventually facilitate syncytium formation and development (Hewezi et al. 2008), 10A06 specifically increases the activity of Spermidine Synthase2 (SPDS2) and induces the cellular antioxidant machinery against plant basal defense in syncytia (Hewezi et al. 2010).

Opposite to *CBP* and *10A06*, heterogeneously expression of some effectors induces plant resistance, which fits with an avirulence role of effector activity. For example, constitutive expression of *H. schachtii 4D09* in Arabidopsis led to significant increase of resistance to nematode. However, this resistance activation is not mediated by the typical NB-LRR R protein. Instead, 4D09 specially binds to 14-3-3ε which induces expression of genes with functions in basal defense responses (Hewezi & Baum 2013). A canonical ETI example that involves *R* gene was found in cyst nematode *Globodera rostochiensis. G. rostochiensis* effector VAP1 induced perturbations of the tomato apoplastic protein RCR3 which was guarded by Cf-2 resistance

protein (Lozano-Torres et al. 2012). This RCR3-Cf2 resistance pathway is also against fungus *C. fulvum* mentioned above and demonstrates an example of an evolutionary conserved resistance strategy of monitoring multiple parasites.

Besides directly facilitating host susceptibility or resistance to pathogen like fungal and bacterial effectors, nematode effectors are also engaged in root structural alteration. Roots infected with cyst nematodes are often dwarfed and branched, while roots infected with root-knot nematodes are stubby and galled (Hewezi & Baum 2013). Cyst soybean nematode *H. glycines* secrete CLAVATA3/ESR (CLE)-like peptides HgCLE2 (4G12) to mimic the Arabidopsis root developmental regulator CLE and reprogram root cells to branch (J. Wang et al. 2010; Wang et al. 2011). Similarly, another signaling peptide 16D10 from the root-knot nematode *Meloidogyne incognita* was found to interact with the tomato Scarecrow-like transcription factors and regulate root radial patterning (Huang et al. 2006). The effector changes the host root development and facilitates formation of syncytia and eventually host susceptibility to nematode.

With advancement of sequencing technology, transcriptome profiling of esophageal glands became a promising avenue to discover effector repertoires and understand evolution of parasitism in plant-parasitic nematodes (Maier et al. 2013). By mining transcriptome of esophageal glands, 18 novel effector candidates were found in *M. incognita* (Rutter et al. 2014). They show little to no homology to known proteins from free-living nematode species, suggesting that they are specific to plant parasitic nematodes and support the parasitic lifestyle.

1.2.2.3.4 Convergent targeting of host proteins by pathogen effectors

Although effectors from different parasites vary in their molecular functions and lack the homolog to each other, they seem to target an overlapping subset of host proteins. These host targets are the hubs of the plant basal immunity network and are subverted by effectors to facilitate pathogen fitness.

In 2011, Mukhtar and his colleagues used yeast-2-hybrid screening method and generated an interaction network of plant-pathogen effectors from two different parasites, the bacterium, *P. syringae,* and the obligate oomycete, *Hyaloperonospora arabidopsidis*, with their common host *Arabidopsis* (Mukhtar et al. 2011). It reveals an interactome (PPIN-1) containing 3,148 interactions among 926 proteins, including 83 pathogen effectors and 843 host interactors. They included biotrophic fungal *Golovinomyces orontii* two years later and built a more comprehensive

interactome (PPIN-2) that ranges across three kingdoms (Weßling et al. 2014). 165 putative host direct interactors (PPIN-1) are enriched for GO annotations in regulation of transcription, metabolism, nuclear localization and some immune- and hormone-related terms. 18 out of 165 were targeted by effectors from both parasites, which is significantly higher than the simulation that randomly assigns connections between effectors and host proteins (Mukhtar et al. 2011). Experimental validation of those 17 common targets revealed that 15 were required for plant immune system function, either positively or negatively regulating plant immunity. Parasites can activate or suppress host interactors to manipulate plant defense and facilitate their own growth.

In addition to direct interactors, interactome also revealed that the majority of NBS-LRR proteins are not directly interact with effectors. Instead, they monitor the effectors by directly interacting with effector host targets (Mukhtar et al. 2011). This observation is consistent with previously mentioned examples that NB-LRR proteins mediate plant defense by surveilling 'guardee' proteins like RIN4 and RCR3, and are activated when effectors modify the integrity of 'guardee' protein.

This study reveals that evolutionarily divergent effectors target a limited set of well-connected cellular hubs in massive plant immune networks and provides a systematic view on parasitism mechanism shared by diverse plant parasites.

1.2.2.4 transfer RNA and small RNA

The movement of effector proteins from parasite to host is an important parasitism strategy for disease development. However, increasing evidence suggests that RNA may also be involved in suppressing host immunity.

Pioneer study done by Weiberg and his colleagues (2013) provided the first evidence that necrotrophic fungal pathogen *B. cinerea* delivers small RNAs (sRNA) to plant to suppress host immunity. Those sRNA hijack host RNA interference (RNAi) pathway and specifically silence host immunity genes like MAPK signaling proteins (MPK2 and MPK1), oxidative stress-related protein (PRXIIF) and cell wall-associated kinase (WAK) (Weiberg et al. 2013). Heterogeneously expressing *B. cinerea* sRNA in host plant Arabidopsis and tomato significantly enhanced disease susceptibility to fungal parasites and pheno-copied mutant lines of sRNA target gene.

Besides sRNA, genomic-scale exchange of mRNA is also observed between parasitic plant *Cuscuta pentagona* (dodder) and its hosts (Kim & Westwood 2015). A strikingly large scale of mRNAs move was observed bi-directionally and contribute a big proportion of transcriptome in both *Arabidopsis* (45%) and dodder (24%). Mobile transcripts tend to be abundantly expressed and significantly enriched in GO annotation like hydrolase activity and response to stimulus compared to non-mobile transcripts. Although there was no experimental evidence of the involvement of mobile transcripts in parasite virulence and host defense, it implies that mRNA can be transferred like effector proteins and be a potential strategy of parasitism.

1.2.3 Adaptive evolution on host-parasite interactiion

Parasites and host rapidly evolve to develop adaptations against each other. Effectors and host effector targets are usually in the front line of this arms race and therefore are under continuous selection pressure (Terauchi & Yoshida 2010; Andrew F. Bent & Mackey 2007). Two popular models have been proposed to explain this gene-for-gene co-evolutionary process. The ‘arms race’ model proposed that the alleles that enhance fitness will increase in frequency within a population and eventually become fixed in both the parasite and host populations. The recurring directional selection shapes allele frequency distribution of both effectors and host effector targets or defense proteins. In contrast, the ‘red queen’ model favors the maintenance of multiple alleles in the population. The frequency of the most common allele in parasite is reduced due to the increasing of matching host allele, while the allele with previously low frequency in parasite takes the advantage to increase their frequency. The balancing selection makes the allele frequencies of matching genes from parasite and host oscillate periodically. These two types of co-evolutionary process leaves opposite patterns of DNA signature which have been detected at different effectors from various plant parasites (Terauchi & Yoshida 2010).

Balancing selection seems to be the most common evolutionary pressure on effectors that direct interact with R genes. One typical example is the avirulence factor ATR13, which triggers *RPP13*-mediated resistance and exhibits high levels of amino acid polymorphism in correspondence to high variability of *R* gene *RPP13* (Allen et al. 2004). A similar example was also observed in direct interaction between flax rust avirulence effectors AvrL567 and *Linum usitatissimum* (flax) *L* resistance genes. At least 12 members of the *AvrL567* gene family have been identified, which represents a recognition specificity with corresponding *L* gene. Some rust strains obtains *AvrL567* variants that escape *L* gene recognition by mutating Avr-R recognition region while maintaining stability for biochemical activity (Dodds et al. 2004; Dodds et al. 2006).

Those direct Avr-R protein interaction examples represents gene-for-gene arm race leading to diversification of both *R* and *Avr* genes.

For effectors that are recognized indirectly by *R* genes, it may be rare to evolve variants that escape recognition without perturbing their virulence activities, since virulence of such effectors usually depends on the interaction with their host targets. Although there was one experimental mutation that successfully uncouples virulence and avirulence in AvrRpt2 to decrease efficiency in RIN4 degradation (Lim & Kunkel 2004), indirectly recognized effectors usually tend to be 'deleted' to escape the surveillance of *R* gene resistance (Andrew F Bent & Mackey 2007). Transcriptional silencing of effector genes can be one of those 'deletion' method. Avr1b indirectly triggered RPS1b-mediated host resistance in soybean by interacting with U-box E3 ligase (Li 2010). Polymorphic variants at promoter region of *Avr1b* silenced its transcription in some virulent strains of *P. sojae*, which could survive on soybean cultivar with *Rps1b* (Cui et al. 2012). Thus, transcriptional silencing of effector genes is another mechanism that plant parasites employ to avoid the activation of host R-gene-mediated immunity.

With an increasing number of available genome sequences from parasites and advancement of next generation sequencing technology, intra- and interspecies comparisons reveal more evidence of adaptive evolution shaping the pathogen effector repertoires. Win et al (2007) used the draft genome sequences of three oomycete plant parasites ---- *P. sojae, P. ramorum and H. parasitica* ---- to generate genome-wide catalogs of RXLR effector genes and discovered that the positive selection mainly acts on the C-Terminal of the RXLR Effectors (Win et al. 2007). The inter-population comparison on Avr effectors of *Melampsora lini* reveals that populations were significantly differentiated with respect to allelic representation at the Avr loci. The specific local selection distinguished genetic structures between pathogen populations, suggesting that strong selective sweeps and demographic bottlenecks have limited diversity within populations (Barrett et al. 2009).

1.3. Parasitic plant Striga and host interaction

Parasitism independently evolved in the Angiospermae multiple times. At least twelve lineages have been recognized to contain parasitic genera at various degrees of dependence on host (Westwood et al. 2010). Plants with parasitic lifestyle have been named parasitic plants. They all share a specialized organ called the haustorium which penetrates the host roots or shoots tissue and derives water and nutrient from host plants via vascular connection.

With the parasitic lifestyle, some of parasitic plants are economically destructive. Witchweeds (*Striga spp.*) is among the most notable noxious species. *Striga* affect millions of hectares of arable farmland in over 25 countries in sub-Saharan Africa and cause over 7 billion USD loss of crop yield annually (Parker 2009). Of 43 identified species of *Striga*, three members, *S.hermonthica, S. asiatica* and *S. gesnerioides*, are well-known agronomic pests for farmers in the infected area. While *S. hermonthica* and *S. asiatic* are parasitic to Poaceae like maize, millet, sorghum and rice, *S. gesnerioides* preferentially attacks dicotyledonous plants, including wild and cultivated legumes and several members of Convolvulaceae, Euphorbiaceae and Solanaceae (Parker & Riches 1993). Among all the hosts for *S. gesnerioides*, the most agronomically affected is cowpea (*Vigna unguiculata*). About 93% of global cowpea production occurs in sub-Saharan Africa where the farmland is heavily *Striga* infested (Spallek et al. 2013; Fatokun et al. 2000).

Controlling *Striga* is especially difficult, because a mature *Striga* plant can produce ~50,000 tiny seeds which can remain viable in soil up to 20 years waiting for favorable condition to germinate. In addition to the massive produce and long period of seed dormancy, these root parasites damage the host crop even before emerging above ground. It causes practical difficulty of preventing parasite before it does the damage (Parker & Riches 1993). The current control methods include hand-pulling, crop rotation and the use of "trap crops" or chemicals that induce germination of the parasite before planting. However, none of them can provide a complete wide-scale effective solution to *Striga* infestation.

Resistant varieties have long been considered as the ideal strategy for control of *Striga*. A number of wild and cultivated crops have been discovered and bred in rice (Bennetzen et al. 2000; Gurney et al. 2006), sorghum (Maiti et al. 1984; Mbuvi et al. 2017), maize (Lane et al. 1997; Amusan et al. 2008) and cowpea (Lane, Bailey, et al. 1993; Omoigui et al. 2017). However, resistance to *Striga* is limited and is often overcome due to adaptation plasticity of the parasite. Thus, continuous discovery of new resistance resources is required to overcome these challenges.

However, from a more positive perspective, years of intense investigation of *Striga* has greatly increased our understanding of the cellular, molecular and genetic mechanisms underlying the interaction of this root parasite with hosts. Undoubtedly, this information will eventually contribute to the development of novel strategies for controlling *Striga* and thereby enhance the food security.

1.3.1 Striga life cycle and host effect

As an obligate parasite, the *Striga* life cycle depends heavily on its host. The germination of *Striga* is initiated by reception of strigolactone secreted from plant roots (Akiyama & Hayashi 2006). The germinated *Striga* seeds start to search for initial host contact and transform their elongated radicals into parasitic organ 'haustoria'. Haustoria penetrate root cell layers of host and establish vascular connection with host to enable nutrient transfer for parasite growth. After emerging from the underground, *Striga* starts to grow like an independent organism and using its own resources for growth (Riopel & Timko 1995).

Most of *Striga* species inevitably deplete nitrogen and carbon sources from either the soil or from the host through vascular connections. However, *Striga* infection has tremendous detrimental effects on the host far beyond the simple removal of resources. It was noticed that *S. hermonthica* infected cereals have significant change in root:shoot balance. The growth of root system of infected hosts is stimulated to increase the chance of attachment, while the shoot system is stunted due to growth reduction (Parker & Riches 1993). *S. gesnerioides* infected cowpea also exhibit stunted shoots. However, in contrast to promoting root growth in *S. hermonthica* infected cereal, root mass of infested cowpea is also reduced remarkably. In addition to the changes in root and shoot mass, *S. gesnerioides* infected cowpea often have a smaller leaf area and visible pale patches or chlorosis on their leaves (Parker 2009). The symptoms affecting the leaves result in considerably reduced the host photosynthetic efficiency, which has been considered another main effect on the host growth (Parker & Riches 1993).

The success of *Striga* is also due to its high variation within or between populations. *S. hermonthica* is a highly variable species due to its obligate out-crossing requirement. It can attack a wide range of crops in grass family but shows distinct host preference from population to population. For example, the localities that adapted to attack pearl millet are poor parasites of sorghum or vice versa (Parker & Riches 1993). However, re-adaptation of *Striga* for new hosts has been observed, where *Striga* was able to switch from barley to wheat after several years (Parker & Riches 1993). This explains the importance of crop rotation in *Striga* control.

Similar to *S. hermonthica*, *S. gesnerioides* has quite high variation within the species and has extreme host specificity among the populations. The Florida population that was developed from *Indigofera* triggers incompatible interaction with cowpea although it was stimulated to germinate by cowpea (Musselman & Parker 1981). Moreover, even cowpea-adapted *S. gesnerioides* can

trigger different resistance responses in various cultivars of cowpea. Based on the resistance variation and geographic difference, Botanga and Timko (2006) grouped cowpea-*parasitic S. gesnerioides* into seven races among which races SG1 (from Burkina Faso) and SG5 (from Cameroon) are the most closely related (Botanga & Timko 2006). One isolate of the *S. gesnerioides* from Zakpota in the Republic of Benin successfully overcome the B301 resistance and is genotypically distinct from other populations of race SG4. It was designated as a separate race SG4z.

1.3.2 multiple-level host resistance

Host resistance is usually mounted at multiple levels and various developmental stages during the life cycle of *Striga*.

Striga relies on a stimulus from the host to germinate and form invading tissue haustoria before attachment. Mutations in biosynthesis or secretion of those stimuli can reduce *Striga* viability and prevent the damage at the beginning. The best known germination stimulant is strigolactone which is widely distributed in plants (Akiyama & Hayashi 2006). Low germination stimulant genotypes of sorghum showed enhanced resistance to *Striga* (Vogler et al. 1996; Haussmann et al. 2001). In addition, the formation of haustoria requires inducing factor 2,4-dimethoxy-p-benzoquinone (DMBQ) released from host cell walls (Albrecht et al. 1999). Thus inhibiting the production and release of haustoria inducing signal also has the potential to prevent the initiation of parasitism.

Most examples of host resistance were found during cortex penetration. Forming physical barriers at endodermis is one of resistance strategies employed by the host plant. The electron dense phenolic/polyphenolic compounds accumulates at attachment interface between *S. hermonthica* and sorghum within 24 to 72 hours after inoculation (Olivier et al. 1991). It is often associated with deposition of callose, lignin and suberin and results in thickened cell walls at endodermis of plants. The histological evidence supports that haustorial ingress of *S. hermonthica* is terminated at the endodermis on the resistant maize inbred derived from Zea diploperennis (Amusan et al. 2008). In addition to the physical barrier, a HR was also observed at attachment site of resistant cowpea cultivar attacked by *S. gesnerioides* (Li & Timko 2009; Lane, Moore, et al. 1993). HR results in the browning at localized host cells and prevents further penetration of *Striga* to reach vascular cylinder. HR displays similar features as immune responses in other plant-pathogen interactions and effectively limits pathogen growth at the cost of local cell death.

Resistance is also frequently seen during the establishment of the vascular connection between *Striga* and its host plant. Cortex penetration is required for *Striga* to reach host vascular cylinder, but penetration alone will not necessarily always result in a connection to host vascular system. In the rice resistant cultivar Nipponbare, *S. hermonthica* successfully traversed the cortex. However, instead of ingress endodermis, it grew around the vascular cylinder and failed to access host water and nutrients (Gurney et al. 2006). This phenotype was observed at different frequencies in other rice cultivars (Yoshida & Shirasu 2009), suggesting that some deficiency in signaling pathway is involved in building vascular continuity (Timko et al. 2012; Yoder & Scholes 2010). Similar resistance due to failure of vascular connection has been also observed between cowpea and *Indigofera*-adapted *S. gesnerioides*. The development of *Striga* is arrested at tubercle, which swells but is internally disorganized (Botanga & Timko 2005). No vascular differentiation and development was observed between host and parasite, implying improper signaling misleading *Striga* development.

Typically, the successful establishment of the vascular connection with host cylinder is a sign of compatible interaction between parasite and host. However, there is one case that resistance occurs following successful vascular connection. When the wild maize *Tripsacum dactyloides* is infected with *S. hermonthica*, the parasite makes connections to the host xylem but still die (Gurney et al. 2003). Improper primary haustorial tissue differentiation on *T. dactyloides* prevents the subsequent secondary haustoria formation. It is proposed that some toxic compound is translocated from host to parasite inhibit the development of haustorial root system and halt the parasite growth (Gurney et al. 2003).

1.3.3 Genetic basis for host resistance

Over last few decades, the genetic basis for host resistance to *Striga* has been explored from different perspectives.

The study of resistance heritability by crossing ‘resistant’ and ‘susceptible’ cultivars or wild relatives have been done in many agronomically important crops including sorghum (Haussmann et al. 2001; Haussmann et al. 2004; Mohamed et al. 2003; Mohamed et al. 2010), maize (Gurney et al. 2003; Amusan et al. 2008), rice (Gurney et al. 2006; Swarbrick et al. 2009; Rodenburg et al. 2015) and cowpea (Omoigui et al. 2017; Singh & Emechebe 1990; Atokple et al. 1995). It was found that the resistance to *Striga spp.* in grass family is associated with multiple quantitative trait

loci (QTLs) and exhibits high phenotypic variations among populations. The low germination stimulant associated resistance to *S. hermonthica* is controlled by a major recessive gene and several minor genes in sorghum (Haussmann et al. 2004). Similarly, using backcross inbred lines (BILs), seven major QTLs have been identified and explain 31% of the phenotypic variation in rice post-attachment resistance to *S. hermonthica* (Gurney et al. 2006). In contrast, monogenic dominant inheritance is demonstrated from the progeny of the crosses between resistant and susceptible cowpea lines. The single dominant gene in cowpea cultivar B301 confers the resistance to *S. gesnerioides* race SG1, SG2 and SG3 respectively (Touré et al. 1997), while the resistance to SG3 in cultivar IT82D-849 was conferred by a single recessive gene (Li et al. 2009). Using various molecular marker techniques (AFLPs, RAPDs and SSRs), researchers were able to find that *Striga* resistance genes clustered at two linkage groups, LG1 and LG6, which are responsible for resistance to *Striga* races SG1, SG2 and SG3.

In addition to the genetic mapping, the molecular genetics has also been studied by examining gene expression changes during the *Striga*-host interactions. Some resistance-related candidate genes have been examined even before high-throughput technology. *NRSA-1*, which was *R* gene homolog, was up-regulated in the roots of non-host *Tagetes erecta* during invasion by the incompatible parasite *S. asiatica* (Gowda et al. 1999). Similarly, the expression of pathogenesis-related protein 5 (PR-5) was induced in roots of resistant cowpea cultivar compared to uninfected roots or roots of susceptible cultivar (Li et al. 2009). With the advent of microarray technologies, researchers have begun to characterize the global patterns of gene expression change between the resistant and susceptible hosts. Comparison between *S. hermonthica* resistant rice cultivar Nipponbare and susceptible cultivar IAC165 reveals a large number of defense genes that are upregulated in resistant hosts. They include PR proteins, pleiotropic drug resistance ABC transporters, genes involved in phenylpropanoid metabolism and WRKY transcription factors (Swarbrick et al. 2008). Similarly, multiple defense pathways in cowpea were induced in incompatible interaction (B301-SG3) compared to compatible interaction (B301-SG4z). These pathways include signal transduction, defense response to biotic and abiotic stress, programmed cell death and apoptosis, lignin biosynthesis and secondary cell wall modifications. These expression analyses provide a handful of candidate defense related genes that define components of a unique resistance mechanism.

Both genetic mapping and high-throughput expression analysis pioneered the molecular research in host resistance to parasitic plants. SSR-1 marker was found to co-segregate with *S.*

gesnerioides race 3 (SG3) resistance among all known cowpea cultivars. It was later found to be located within a cowpea EST sequence encoding canonical R gene (RSG3-301) with a coiled-coil (CC) protein-protein interaction domain at the N terminal, a nucleotide binding site (NBS) and a leucine-rich repeat (LRR) domain at the C terminal. RSG3-301–silenced B301 plants fail to mount a HR when challenged with SG3 while still exert resistance to SG2 and SG5 (Li & Timko 2009). This implies that RSG3-301 functions in a race-specific manner and exerts gene-for-gene resistance to *S. gesnerioidies*. OsWRKY45, from the other hand, was identified through genome-scale RNA sequencing expression analysis in *S. hermonthica*-infected rice roots. It modulates a cross talk between JA and SA pathways and enhances the resistance to *S. hermonthica* by positively regulating both pathways (Mutuku et al. 2015).

Although the genetic basis for host resistance to *Striga* has been explored for several decades, very little research is performed from parasite perspective. In my PhD dissertation, I was trying to uncover the molecular factors of *S. gesnerioides* that are involved in compatible (susceptible) and incompatible (resistant) interactions with cowpea. In Chapter 2, I examined the global transcriptional difference among the three races (SG3, SG4 and SG4z) of *S. gesnerioides* and revealed several candidate genes responsible for virulence and avirulence. In Chapter 3, I performed functional analysis of one candidate gene SHR4z and examined the molecular mechanism underlying SG4z overcoming host resistance. In Chapter 4, I inspected the transcriptional changes of *S. gesnerioides* during compatible (susceptible) versus incompatible (resistant) interactions with cowpea. Collectively, my work elucidates, for the first time, the molecular mechanism underlying the virulence of parasitic plant over host resistance. It could potentially contribute to *Striga* control and maintenance of crop security.

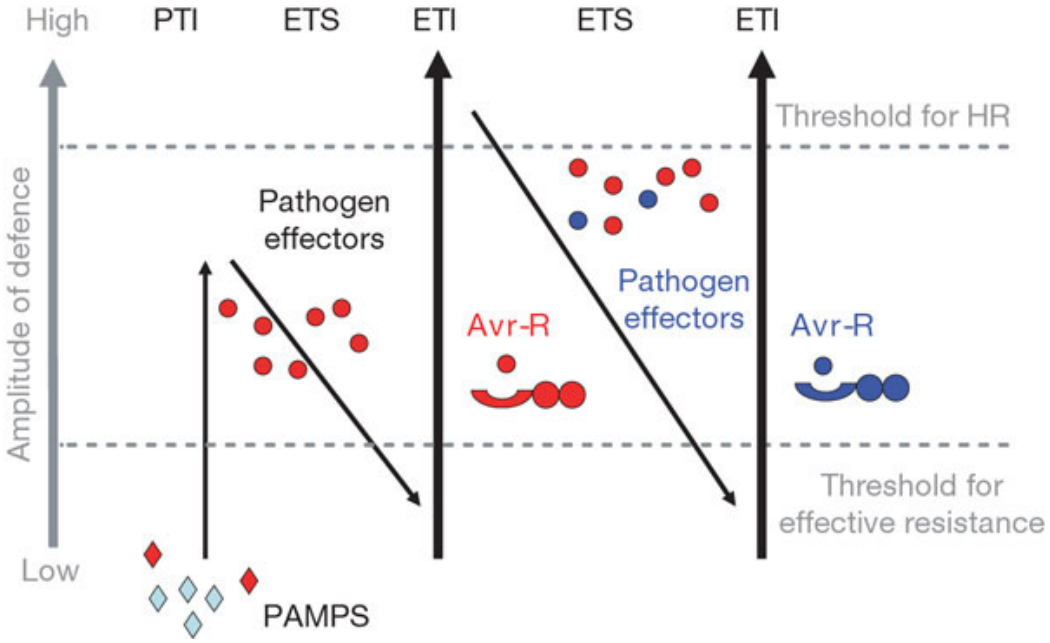


Figure 1.1. A zigzag model illustrates the multiple levels of plant immune system. Pathogen-associated molecular patterns (PAMPs) or microbe-associated molecular patterns (MAMPs) are recognized by pattern recognition receptors (PRRs), resulting in PAMP-triggered immunity (PTI). Adaptive parasites deploy effectors to block PTI and cause effector-triggered susceptibility (ETS). To overcome this, specific resistance (R) proteins are evolved to recognize parasitic Avr effectors leading to effector-triggered immunity (ETI) and HR. In response to ETI, mutations in effector genes results in losing Avr effector or generating new virulence effectors to re-gain ETS. Finally, the selection drives evolution of new R-genes that can recognize one of the newly acquired effectors, resulting again in ETI. (adapted from Jones and Dangl 2006)

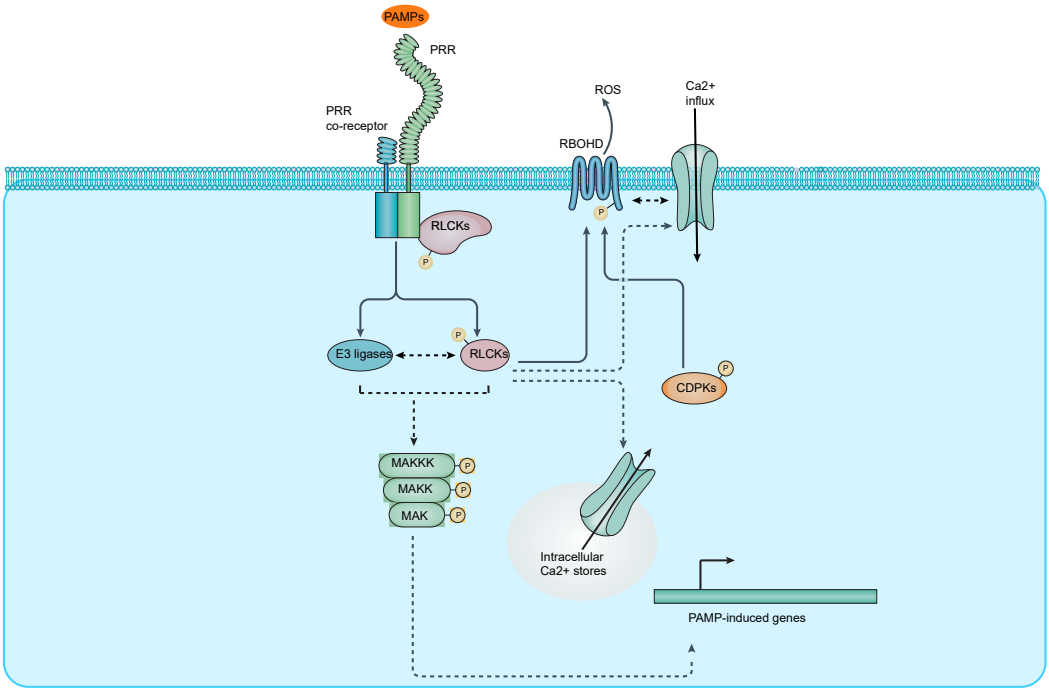


Figure 1.2. Early branching of PRR-triggered immunity signaling. Pathogen-associated molecular pattern (PAMP) perception by pattern recognition receptors (PRRs) induces immune signaling that branches immediate downstream of the PRR complex with facilitation of PRR co-receptors. The recognition of PAMPs triggers downstream signaling pathways mediated by the receptor-like cytoplasmic kinase (RLCK), E3 ligase, Ca2+-dependent protein kinase (CDPKs), mitogen-activated protein kinase (MAPK) cascades, and eventually results in reactive oxygen species (ROS) production and transcriptional re-programming of PTI genes (adapted from Couto and Zipfel 2016).

Comparative Transcriptome Analysis on Three Races of the Root Parasitic Angiosperm *Striga gesnerioides* that Differ in Virulence

2.1 Abstract

In this chapter, by using high throughput RNA-seq analysis, I carried out transcriptome profiling on tissue isolated from parasite haustorium during compatible and incompatible response of resistant (B301) and susceptible (Blackeye) cowpea cultivars with three races of *S. gesnerioides* (SG3, SG4 and SG4z). Subsequent de novo assembly reveals the first *S. gesnerioides* transcriptome consisting 14,507 contigs and 114,231 unigenes. Differential expression comparison between two evolutionarily closed races SG4 and SG4z reveals 267 race differentially expressed contigs. Based on expression pattern of transcripts and predicted structures of proteins, we were able to *in silico* identify 34 parasite effector candidates that are differentially expressed between SG4 and SG4z. A total of 12,588 contigs are differently expressed during incompatible interaction with B301 in both SG3 and SG4. 71 contigs are specifically upregulated in SG3 and account for the Avr candidates in gene-for-gene resistance interaction between B301 and SG3. This finding provides a manageable candidate list for further characterization of virulence profiles in *S. gesnerioides* and enhance our understanding of the evolution of virulence in parasitic weeds.

2.2 Introduction

Witchweed (*Striga spp.*) is a genus of devastating root parasitic plants that affect millions hectares of arable farmland in over 25 countries in sub-Saharan Africa (Parker 2009; Ejeta 2007). The 43 identified species of *Striga* are generally classified into two main groups based on their host preference. The first group, including *S.hermonthica* and *S. asiatica*, mainly parasitizes crops in Poaceae family. The other group, represented by *S. gesnerioides*, specifically attacks dicotyledonous plants, including wild and cultivated legumes, a number of *Tephrosia spp.* and *Indigofera spp.* in West Africa and some *Convolvulaceae* species in North Africa (Parker & Riches 1993). Cowpea (*Vigna unguiculata L. Walp*.), as a major source of dietary protein for humans and domesticated livestock in sub-Saharan West and Central Africa, is among the most economic significant hosts for *S. gesnerioides*.

Despite a wide range of hosts for the species as a whole, isolates of *S. gesnerioides* from different locations exhibit extreme host-specificity. Populations of *S. gesnerioides* harvested from tobacco

fields in Zimbabwe are unable to parasitize cowpea (Parker & Riches 1993), while the Florida population specifically infected *Indigofera* (Musselman & Parker 1981). Even among the isolates from cowpea fields, the compatibility of different *Striga* races with different cowpea cultivars varies dramatically. Lane et al. (1996) differentiated five races of S. gesnerioides from seven West Africa countries by performing pot infections on a series of four cowpea cultivars (Lane et al. 1996). A broadly virulent isolate, SG3, is pathogenic on three cultivars except most resistant cowpea cultivar, B301. However, this cowpea cultivar B301 is highly susceptible to *S. gesnerioides* isolates from Zakpota, a town in southern part of the Republic of Benin (Atokple et al. 1995; Singh & Emechebe 1990). Interestingly, even though one cowpea cultivar is resistant to multiple *Striga* races, the host resistant response is not identical among different races. The interaction between SG4 and B301 is characterized by limited growth of parasite (Lane et al. 1996), while SG3 infection results in necrosis on cowpea roots, which was later found to be caused by a race-specific resistance (*R*) gene *RSG3-301* in the B301 cultivar (Li & Timko 2009). The *Striga*-cowpea interaction appears to follows a gene-for-gene resistance mechanism, where the host R gene and parasite effector gene (*Avr*) are paired to trigger resistance defense.

The first evidence of association between race-specific virulence and genetic variety was demonstrated using Amplified Fragment Length Polymorphism (AFLP) profiles of seven distinguishable *S. gesnerioidies* races (Botanga & Timko 2006). Botanga and Timko (2006) found that the genetic variability is exceedingly low within isolates from the same country, with SG4 from Republic of Benin as an exception. Nineteen selective AFLP markers identified a single fragment polymorphic difference between the hypervirulent isolate SG4z and all other SG4 populations throughout Benin. It is likely that a recent mutation on Avr from SG4 in response to B301 resistance was fixed by host selection and gave rise to new variants of S. gesnerioides SG4z. However, the lack of genome-wide studies limited discovery of the molecular basis of virulence factors contributing race-specific genetic variance.

In last decade, the advancement of sequencing technologies enabled high-throughput studies on a tremendous number of non-model organisms including parasite plants. Yang et al. (2015) identified more than 100 putative parasitism genes using comparative transcriptome sequencing on three root parasites from *Orobanchecea* (Yang et al. 2015). A large-scale exchange of transcripts between the stem parasitic plant *Cuscuta pentagona* (dodder) and its hosts was characterized by tissue-specific sequencing on both host and parasite plants at attachment site (Kim et al. 2014). These studies pioneered the genomic research of parasitic plants and gained

insight into the evolution of parasitism among different plant species. However, the intra-species differences among parasitic plants with differential virulence profiles have yet to be examined. *S. gesnerioides* provides a great model for this study due to its considerable variation and its high host-specificity within population.

In this study, I characterized the virulence variation among three races of *S. gesnerioides* during the interaction with the cowpea cultivar B301 at subterranean stage. We identified putative virulence factors that assist *Striga* with overcoming host resistance in SG4z by comparing the expression profile of SG4z to its evolutionarily closest race SG4. This analysis revealed 267 contigs differentially expressed at post-attachment stages. Thirty-four of those contigs are predicted as virulence candidates because they are induced after host contact. Comparative transcriptomic analysis between SG3 and SG4 reveals 12,588 contigs undergo expression changes at incompatible interaction with B301 and 71 of them are specifically upregulated in SG3 and account for the Avr candidates in the gene-for-gene resistance interaction between B301 and SG3. This comparative study not only reveals direct transcriptome-scale genetic basis for population variety in *S. gesnerioides* but also provides a manageable candidate list for the functional study of *S. gesnerioides* virulence.

2.3 Materials and Methods

2.3.1 Plant Tissues, libraries and sequencing

S. gesnerioides seeds were collected from plants grown on cowpea cultivars from the following locations: SG3 (Maiduguri, Nigeria - 2008), SG4 (Cana, Benin - 2009) and SG4z (Zakpota, Benin -2006). Seeds from cowpea cultivar B301 were kindly provided by Dr. Lucky Omoigui (University of Agriculture Mukurdi, Nigeria) and California Blackeye no. 5 (BE) seeds were purchased commercially (Burpee Seeds). The resistance phenotypes in response to different races of *S. gesnerioides* has been previously described (Botanga & Timko 2005; Botanga & Timko 2006).

For preparing material for RNA sequencing, cowpea seeds were disinfected by first rinsing in 75% (v/v) EtOH, followed by a 5-10 min wash in 5% (v/v) hypochlorite, and finally extensive rinsing in sterile distilled water. Seeds were imbibed for 2 d at 30 C, germinated on sterile moistened filter paper, and the young seedlings were transferred to rhizotrons containing sterile granulated rockwool, and grown for 16 d at 30-33ºC under a 14 h light/10 h dark cycle (Mellor et

al. 2012). For parasite inoculations, *S. gesnerioides* seeds were sterilized, water pre-conditioned for 12 d and germinated by exposure to cowpea root exudates for 2 d at 30 C. The germinated seeds were either harvested at the germination stage (0 days post-inoculation (dpi)) or used to inoculate cowpea roots. Roots were inoculated by brushing the germinated *Striga* seedlings along host roots. Following inoculation *Striga* seedlings were allowed to initiate host contact and penetration and at 3 dpi (early attachment) and 10 dpi (late attachment), when the rhizotrons were opened, the parasite seedlings were collected from the host roots into polystyrene tubes, quickly frozen in liquid Nitrogen, and stored at –80˚C until used for RNA extraction. Experiments were conducted in triplicate for all parasite races for 2-host genotypes and 3 time points.

Total RNA was extracted from the various parasitic seedlings using an RNeasy Plant Mini Kit (QIAGEN cat# 74904) according to the manufacturer's protocol. The extracted RNAs were treated with DNase (QIAGEN cat#79254), and further purified using the RNA Clean & Concentrator-5 Kit (Zymo cat#R1016). PolyA^{+} -RNAs were purified and library preparation for subsequent Illumina single-end sequencing was carried out as previously described (Yang et al. 2015). 300 millions 150bp single-end reads were generated using Illumine mRNA sequencing.

2.3.2 Virulence assessment

The surface of *S. gesnerioides* seeds were sterilized and pre-conditioned as described above, and then gently transferred to B301 cowpea roots using a fine paintbrush. At 10-day post-inoculation (dpi) the occurrence number of all parasite-host root interaction events (i.e., attachment, tubercle swelling, cotyledon expansion and hypersensitive response) were scored and the percentage for each interaction type were determined. At total of 10 individual plants were used per treatment and the two-tailed t-test was used to determine statistical significance.

2.3.4 De novo assembly and post assembly process

Adaptors, leading and trailing bases quality below than Q30 were removed from libraries using Trimmomatic (v0.32, -SE, -phred33 2:30:10 LEADING:3, TRAILING:3). Only reads longer than 36bp (MINLEN:36) and average quality for continuous 4 bases (SLIDINGWINDOW:4:15) is above Q30 were retained for following analysis. All strand-specific libraries from different conditions were combined and de novo assembly was performed using the software, Trinity (Release v2.0.6, --seqType fq –single –SS_lib_type R) (Haas et al. 2013). Since we assumed that effectors are a group of small peptides, only protein coding sequences were subjected to post-

assembly cleanup, as performed in Yang et (2015) (Yang et al. 2015). The cleanup steps include removing potential non-plant contamination, excluding host contamination and retaining transcripts that are homologous to unigenes in other Orobanchaceae family members. CDS regions and protein sequences were predicted using ESTscan 3.0 with Arabidopsis as the reference (Iseli et al. 1999). The final transcriptome was obtained by removing redundant transcripts which share exactly same nucleotide sequences and peptide sequences shorter than 40 aa. Individual race transcriptome was assembled by combining libraries in each race and further processed using the same cleanup steps as above.

2.3.5 Plant Ortholog Group Assignment and Transcriptome Annotation

Plant orthogroups were identified by assigning predicted protein sequences to pre-defined orthogroup derived from 26 genomes from fully sequenced land plant genome using PlantTribe 26Gv2.0 (https://github.com/dePamphilis/PlantTribes) (Wall et al. 2008). Proteins predicted from the transcriptome were searched against previously constructed HMM orthogroup classification profiles using BLASTP (Evalue < 1e-5) and HMMscan (Evalue < 1e-10). Once hits were found, each transcript was assigned to an orthogroup which corresponds to the best hit. These translated transcripts were further annotated by searching homologs in TAIR10 and UniProtKB database using BLASTP (evalue < 1e-10). Putative functional annotations were assigned using the automated Assignment of Human Readable Descriptions (AHRD) pipeline (https://github.com/groupschoof/AHRD).

2.3.6 Read mapping and expression evaluation

The hybrid transcriptome was used as reference to obtain the relative expression using program eXpress v1.5.1 (Pachter 2011). First, high-quality Illumina reads from each library were mapped to reference transcriptome via bowtie2 (Langmead & Salzberg 2012) with recommended settings in eXpress (multiple alignments (-a) = T, multiseed aligment length = 20, maximum mismatches allowed in seed = 1. open gap penalties = 6, extend penalties = 5, minimun alignment score function = -0.6 + -0.4 * read length) (Roberts et al. 2011). Then transcript abundances were evaluated independently in each library by eXpress with default settings (except --r-stranded). The total count of reads which uniquely or ambiguously mapping to the transcripts were used as input for scaling sample sequencing depth in DESeq2 v1.6.3 (R version 3.2.2) (Love et al. 2013). Transcripts Per Million (TPM) was calculated to represent transcript abundance and used for expression visualization.

2.3.7 Differential expression analysis using DESeq2

Raw read counts from all libraries were normalized using the estimateSizeFactors function in DESeq2 (Love et al. 2013) and the normalized reads in each library were used to proceed with the differential expression analysis. Stage differential expression was examined in each *Striga*-host interaction (SG3-BE, SG4-BE, SG4z-BE, SG3-B301, SG4-B301, SG4z-B301). Genes were considered as differentially expressed when FDR < 0.01 and absolute value of log2FoldChange > 1 in pairwise comparison. Genes that dynamically expressed across three developmental stages were identified by comparing early attachment to germination and late attachment to early attachment respectively.

The two factor full generalized linear model (Race+Host+Race:Host) was used to examine race differential expression at each attachment stage (FDR < 0.01 and log2 of FoldChange > 1). Race differentially expressed genes were defined as differentially expressed at both host and pairwise race comparison.

2.3.8 *In silico* identifying effector candidates and expression clustering

To identify potential candidate effectors, genes differentially expressed across different developmental stages were first classified a priori based on the expected pattern of the stage pairwise comparison. We expected effectors to be highly expressed during the initial interaction of parasites with hosts. In addition, we anticipated three potential expression groups: genes with (i) continuously high expression during the two attachment stages, (ii) expression peaks at the early attachment stage; and (iii) expression peaks during the late attachment stage.

Following our expression analysis, all contigs with complete protein coding regions were examined for the presence of signal peptides and transmembrane domains using SignalP v 4.1 with default setting (Petersen et al. 2011) and TMHMM (Krogh et al. 2001). Contigs predicted to contain signal peptides and no transmembrane domain were further examined for the presence of ER localization sequences using the PROSITE database (Hofmann et al. 1999). Contigs containing signal peptides, lacking transmembrane domains, and lacking ER localization domains (i.e., predicted to be transport to extracellular region) were included in the “secretome” data set. Contigs in the secretome data set with expression patterns match our a priori assumptions for candidate effectors, as described above, were placed in the “effector candidates” group. A large number of contigs did not have complete coding regions and in particular could not be analyzed

for signal peptides at their 5' ends. The contigs that met all of the other criteria (i.e., lacking transmembrane domains, ER retention signals, and haustorial specific expression peaks in post-attachment stages of the host-parasite interaction) were also included in "effector candidate" group (Figure 2.3).

2.3.9 Enrichment Analysis

GO terms for each *Striga* contig were determined by GO terms associated with best hits from Arabidopsis (BLASTP evalue < 1e-10). The hypergeometric distribution was used to compare the number of genes were annotated for a specific GO term in the given set versus the number of genes expected to annotate to that specific GO term in a randomly drawn subset of the population. The overrepresented terms were determined by FDR from one-tail hypergeometric test (FDR < 0.05) and analyzed with the GOSlimViewer v2.00 (http://www.agbase.msstate.edu/cgi-bin/tools/goslimviewer_select.pl) to obtain a high level summary of functions for given gene set.

2.4 Results

2.4.1 Virulence variation among three races of S.gesnerioides on B301

The seeds of *S. gesnerioides* germinate in response to the presence of chemical signals (strigolactones) in the rhizosphere released from the roots of potential host plants (Bouwmeester et al. 2011; Bouwmeester et al. 2003). Two day-old germinated seedlings of the various *S. gesnerioides* races parasitize on cowpea crops in West Africa (e.g., SG4, SG4z, SG3) are phenotypically indistinguishable prior to host contact (Figure.2.1A). However, following attachment to the host root, differences in host and parasite phenotype become evident depending on the counterparts of the interaction (Lane et al. 1996; Botanga & Timko 2005). California Blackeye no. 5 (BE) is susceptible to all races of *S. gesnerioides* identified thus far (Botanga & Timko 2006) and by 3 days post-inoculation (dpi) BE roots show significant numbers of attached *Striga* seedlings from all three races. No obvious discoloration is found on the BE roots at the site of parasite attachment at this early stage. By 10 dpi, all three races of the parasite (SG4, SG4z and SG3) have successfully penetrated the BE root cortex and established connections with the host vascular system. *Striga* displays formation of an enlarged tubercle and evidence of cotyledon expansion, indications of a successful vascular connection between host and parasite (Fig 2.1B).

In contrast to BE, a range of phenotypic responses is observed in interactions between the *Striga* races and B301, a cultivar from Botswana resistant to all *Striga* races except SG4z (Botanga & Timko 2006). At 3 dpi, B301 roots show only very minor distinguishable phenotypic differences when challenged by *Striga* races SG4z, SG4 (from Benin) and SG3 (from Nigeria) with the most notable effect being the appearance of a slight discoloration or browning of the host root at the attachment interface of SG3 and SG4 (Fig. 2.1a). However, at 10 dpi, dramatically different phenotypic responses are observed among the races. Both SG4 and SG3 have elicited robust HR with associated browning and necrosis of the B301 root at the site of parasite attachment (Fig. 2.1a). There is also arrested growth of the parasite and a browning of the parasite. The reaction of B301 to SG3 challenge is more pronounced than that observed with SG4, with just slightly more arrested growth at the tubercle stage of parasite development than observed in SG4 (Fig. 2.1b). In stark contrast, SG4z seedlings fail to elicit a strong HR on B301 roots (Fig. 2.1). Similar to what is observed on the susceptible BE, SG4z-B301 interactions have significantly more tubercle swelling events and more frequent observed seedlings with evidence of cotyledon expansion (a characteristic associated with the formation of successful vascular connections with the host) by comparison to SG3 and SG4 –B301 interactions.

2.4.2 Overview of the *S.gesnerioides* transcriptome

To uncover alterations in gene expression in SG4z that could potentially underlie its ability to overcome or bypass the innate immunity / defense pathways present in B301 activated by the closely related SG4, we generated transcriptome profiles of SG4 and SG4z, and a more distanced race, SG3, during compatible and incompatible interactions with BE and B301 roots. Three developmental stages were sampled: two day-old germinated seedlings prior to host contact; parasite seedlings at an early host attachment stage (3 dpi); and parasite seedlings at a late host attachment stage (10 dpi). A total of 45 different libraries were generated and sequenced using an Illumina RNA HiSeq platform yielding more than 300 million 1 X 150 bp single-end reads (Appendix Table A1).

We built individual race-specific transcriptome assemblies for SG4, SG4z, and SG3 as well as a hybrid transcriptome assembly (combining all sequencing data for all stages from each race referred to hereafter as SGall) using the Trinity (r2014-07-17) assembly tools (Haas et al. 2013). The various transcriptome assemblies were annotated to coding proteins by ESTscan v3.0.3 (Iseli et al. 1999). A total of 145,407 protein-coding contigs were identified by the SGall hybrid assembly, and an average of 98,051 contigs were identified in each of the race-specific

assemblies (Table 2.1). The N50 contig length was 1,218 for SGall, and ranged from 1,459 (SG3) to 1,594 (SG4z) in the individual assemblies, which is comparable to the assembly statistics reported from other three haustorial transcriptome assemblies of parasitic plants (Yang et al. 2015).

We examined gene capture frequency using the three known sets of conserved single copy genes, namely the universally conserved orthologs (UCOs) (Der et al. 2011; Williams et al. 2014) (http://compgenomics.ucdavis. edu/compositae_reference.php, last accessed December 18, 2014), conserved single copy genes from COSII (Wu et al. 2006; Fulton et al. 2002; Williams et al. 2014) (http://solgenomics.net/documents/markers/ cosii.xls, last accessed December 18, 2014), and the set of conserved single copy genes in PlantTribes2 (Wall et al. 2008) (http://fgp.bio.psu.edu/tribedb/10_genomes/, last accessed 2015-8-18). The results indicate that gene coverage frequency ranges from 89.55% (COSII analysis) to 98.88% (UCO analysis) in our assembly, which is comparable to gene capture frequency reported for three haustorial transcriptome assemblies from other root parasitic plants including the closely related *S. hermonthica* (Yang et al. 2015). When the assemblies are compared to comparable developmental-stage specific transcriptome assemblies from *S. hermonthica* (i.e., Stages 1, 2 and 4) by reciprocal blastn (evalue < 1e-10), ~ 63.6% of *S. hermonthica* unigenes have homologous counterparts in the SGall assembly. Therefore, we believe that our assembly represents a majority of the expressed genes in the S. gesnerioides haustorial transcriptome.

To evaluate how representative the SGall hybrid assembly is compared to the individual race-specific assemblies, the predicted protein sequences of all four assemblies are combined to construct orthologous groups across the three races (SG4, SG4z, and SG3) using orthomcl (Li et al. 2003). Among all 91,756 clustered orthologous groups, 89.8% orthologous groups are shared by at least two of the race-specific assemblies. The SGall hybrid assembly represents 79.8%, 76.1% and 75.3% of orthogroups in the SG3, SG4 and SG4z assemblies, respectively (Appendix Figure 1). This is due to the assembly discrepancy of Trinity software using different pools of reads. Further nucleotide-wise comparisons reveal that an average of 92.3% of the race transcriptomes are represented by the SGall assembly, indicating that the SGall hybrid assembly is likely to capture a majority of the expressed genes in the individual race assemblies and can be used as a good reference for differential expression analysis across multiple races.

To determine which factor contributed most to expression profile differences among parasite, we map the reads from all 45 RNA-seq libraries to the SGall hybrid transcriptome assembly and then perform expression correlation clustering across the 15 conditions (parasite race X developmental time X host genotype). This analysis show that host genotype and parasite developmental stage contribute significantly to differences among the various parasite race expression profiles (Appendix Figure 2). At the germinated seedling stage, all three parasite races clustered together indicating that while differences exist among the races, such differences are not pronounced prior to host interaction. As might be expected, given the recent evolutionary derivation of SG4z from SG4 (Lane, Moore, et al. 1993; Botanga & Timko 2006), SG4 and SG4z samples are clustered closer to each other than either do to SG3. Similarly, samples from the same post-attachment developmental stage are clustered together, with the exception of those from the late attachment stages where compatible (susceptible) and incompatible (resistant) host-parasite pairing are being analyzed and clear distinctions are evident.

2.4.3 Differential Gene Expression between SG4z and SG4

We hypothesize that differences in gene expression may underlie variations in virulence among *S. gesnerioides* races (Whitehead & Crawford 2006) and that such differential expression would account for the differential resistance elicitation of SG4z compared to SG4 in its interaction with B301. Among the possible differences we anticipate are that SG4z might have lost or express significantly diminished levels of one or more gene products that function as avirulence factors that is recognized by B301. This precludes its detection by the host immunity system allowing SG4z to establish a compatible interaction with B301. Alternatively, SG4z may have altered the expression of an existing gene or gained new gene expression function (e.g., a new effector) that allows it to subvert the B301 defense pathways thereby yielding the host susceptible to parasitism.

To address these possibilities, we first examine the degree to which transcripts are differentially expressed between SG4z and SG4 at the various stages of interaction with BE and B301. Not surprisingly given their recent evolutionary divergence, relatively few contigs (135 in total) are found to be significantly differentially expressed (FDR < 0.01 log2FoldChange > 1) between the SG4 and SG4z transcriptomes of germinated seedlings (0 dpi) and of these 135 contigs, 50 are upregulated in SG4z and 85 are upregulated in SG4 (Appendix Table A2). A total of 259 contigs are differentially expressed between SG4 and SG4z at the early attachment stage (3 dpi) on B301, whereas a slightly larger number (300 contigs) are differentially expressed at 3 dpi on BE

(Appendix Table A2, Figure 2.2B and C). In addition, the magnitude of the expression changes between SG4z and SG4 is significantly increased at 3 dpi compared to 0 dpi (Figure 2.2A; t-test p-value < 2.2e-16). At the late attachment stage (10 dpi), 17,242 contigs (~11.9%) are differentially expressed between SG4z and SG4 when parasitizing B301 but only 191 contigs (0.13%) are differential expressed in the interaction with BE. This large difference in the number of differentially expressed genes at 10 dpi likely reflects the fact that one race (SG4z) is participating in a compatible interaction with B301 while the other (SG4) is in an incompatible interaction in which host penetration is arrested and it is dying (Appendix Table A2, Figure 2.2B and C). Combined, a total of 205 and 150 contigs are shared by both interactions and are significantly differentially expressed between SG4z and SG4 at 3 dpi and 10 dpi, respectively (Figure 2.2C). There are 267 contigs that have race-specific differential expression at either post-attachment stages. Hierarchical clustering of those 267 contigs reveals that 122 (45.7%) of the contigs are upregulated in SG4z and 131 (49.1%) have higher expression in SG4 (Figure 2.2 D).

To characterize what role these genes may be involved in, we focus on members from this list with known homologs in *Arabidopsis* and examine their functionally annotated categories using GO enrichment. The contigs highly expressed in SG4z are enriched for kinase activity (MF, FDR=8.36E-9), transferase activity (MF, FDR=3.21E-6), catalytic activity (MF, FDR=0.01), response to endogenous stimulus (BP, FDR=0.011) and response to oxygen containing (BP, FDR=0.011). The enrichment of molecular function of SG4 upregulated contigs is very similar to the contigs that are highly expressed in SG4z, including kinase activity (MF, FDR=0.012) and transferase activity (MF, FDR=0.048). However, the biological process of contigs upregulated in SG4 is quite different and enriched in proteolysis (BP, FRD=6.89E-009) and protein metabolic process (BP, FDR=1.46E-005). Interestingly, both race differentially expressed contigs are significantly enriched in extracellular region (FDR=0.04 for SG4 upregulated contigs, FDR=6.74E-006 for SG4z upregulated contigs), implying that the genes contributing to virulence difference between SG4 and SG4z are mainly present in secretome. This is consistent with the definition of pathogenic effectors that were secreted by parasites to aid infection of host.

2.3.4 Identification of effector candidates differentially expressed between SG4 and SG4z

Since effectors and avirulence factors are expected to be released from the haustorium upon the interaction of *Striga* with its potential host (serving to elicit host defense responses or functioning to suppress them), we hypothesize that there might be significant alterations in their expression

relative to germinated parasite seedlings. We also expect that proteins moving from parasite to host would likely contain secretion signals and other structural features to facilitate this movement (Goritschnig et al. 2012; Hacquard et al. 2012; Rafiqi et al. 2013).

To address these possibilities, we develop an *in silico* prediction pipeline based on those known features of effectors and applied it to transcriptome data (Figure 2.3A). We start by identifying contigs that are upregulated in at least one host-interaction stage (3 dpi and 10 dpi). There are 4,067, 4,952, 5,342 and 5,143 contigs following this expression pattern at SG4-B301, SG4-Blackeye, SG4z-B301 and SG4z-Blackeye interaction, respectively (Appendix Table A3). Following the initial expression analysis, all contigs with complete protein coding regions are examined for the presence of signal peptides, transmembrane domains, and ER retention sequences. Contigs meeting all three of the above criteria are included in the "secretome" data set. Contigs in the secretome dataset that are differentially expressed in the post-attachment stages of host-parasite interaction (as described above) are placed in the candidate effector group. Since a large number of contigs do not have complete coding regions, those contigs that meet all of the other criteria (i.e., lacking transmembrane domains, lacking ER retention signals, and haustorial specific expression peaks in post-attachment stages of the host-parasite interaction) are also included in effector candidates group. A total of 578 contigs are recognized as falling into the secretome group from all of the possible parasite-race:host genotype interactions (SG4-B301 (273), SG4-BE (334), SG4z-B301(280) and SG4z-BE (332)) and a total of 4635 contigs (SG4-B301 (1851), SG4-BE (2215), SG4z-B301(2338) and SG4z-BE (2308)) could be placed in the candidate effector group.

We hypothesize that the differential virulence of SG4z (which is derived evolutionarily from SG4), may have lost or significantly diminished levels of expression of one or more genes for effectors or avirulence factors recognized by B301 precluding its detection by the host immune system and allowing SG4z to establish a compatible interaction with the host (i.e., B301 now is susceptible). Alternatively, SG4z may have gained expression of a new gene or altered expression of a gene that allowed it to subvert host defense, thereby yielding the host susceptible to its parasitism. Therefore, the effectors that contribute to virulence difference are possibly among effector candidates that differentially expressed between SG4 and SG4z. Of all 4,635 effector candidates, 34 were differentially expressed between SG4z and SG4 at least at one stage (Figure 2.3B and Appendix Table A4). Of these 14 were upregulated in SG4 and constitutes avirulence effector candidates while 13 were induced in SG4z and may act to suppress the host's defense.

The rest 7 contigs are specifically highly expressed in SG4z at early attachment stage while upregulated in SG4 at late attachment stage. Among the contigs with homology to characterized gene, significantly up-regulated genes in SG4 include several protein kinases, FAD-binding Berberine family protein and protein with BED zinc finger. In contrast, the genes that specifically induced in SG4z include several proteins involved in cell wall modification (Peroxidase superfamily protein, beta-galactosidase and beta-D-xylosidase 4), Replication protein A 70 kDa DNA-binding subunit B and one leucine-rich receptor-like protein kinase family protein (Figure 2.3B).

To confirm that the candidate effectors are in fact differentially expressed in the post-attachment haustoria of SG4z compared to SG4, qRT-PCR is performed on six selected candidates that are homologous to functionally known proteins (Figure 2.3C), which confirm our RNA-seq results. One candidate, SGall_040908.1, reproducibly show no detectable expression in SG4 and significantly higher (100-fold or greater) expression in SG4z compared to SG3 at both 3 dpi and 10 dpi. Based on this dramatic difference in expression, we select this candidate for further detailed analysis in next chapter.

2.4.5 Global changes of gene expression from SG3 and SG4 during the resistant interaction with B301

Unlike to SG4z, both SG3 and SG4 trigger resistant response during the interaction with B301. In SG3, 5,958 contigs exhibit differential expression (4,298 upregulated, 1,660 downregulated) from 0 dpi to 3 dpi and 5,907 contigs (2,647 upregulated, 3,260 downregulated) from 3 dpi to 10 dpi. In SG4, the stage differentially expressed contigs include 3,328 (2,099 upregulated, 1,229 downregulated) from 0 dpi to 3 dpi and 5,229 (2,229 upregulated, 3,000 downregulated) from 3 dpi to 10 dpi. In total, there are 12,588 contigs responsive to host contacts in resistant interaction (Figure 2.4A).

To identify genes and pathways that are common to different races of *S. gesnerioides* during the resistant response, we examine the global expression changes that are shared by both races. There are 1,422 contigs upregulated and 814 downregulated shared by both races from 0 dpi to 3 dpi. A greater number of contigs are differentially expressed from 3 dpi to 10 dpi, with 1,075 contigs upregulated and 2,160 contigs downregulated (Figure 2.4A). All of these differentially expressed genes that are shared by both races are generally clustered to five groups of expression patterns (Figure 2.4B). They are downregulated at 3 dpi (749 contigs), downregulated at 10 dpi (1,106

contigs), transient expression at 3 dpi (851 contigs), upregulated at 10 dpi (1077 contigs) and up-regulated 3 dpi (1233 contigs). The pathways suppressed immediately after host contact (3dpi) are associated with seedling development and embryo development ending in seed dormancy (Appendix Table A5). This result fits with our expectations since early contact with resistant host start to inhibit *Striga* growth by interfering development pathway. The genes that are down-regulated later stage (10 dpi) are enriched in the biological processes associated with response to abiotic stress and cell wall repair (Appendix Table A5). In contrast to downregulated genes, the genes induced at 3 dpi are significantly enriched in pathways associated with defense response by callose deposition, response to endoplasmic reticulum stress and regulation of defense response (Appendix Table A5). These pathways are also enriched among the genes that are expressed at 10 dpi. In addition, the contigs annotated in the regulation of cell death are only induced at late attachment stage, which is consistent with the appearance of HR at 10 dpi. One group of genes are different from all above, and show transient expression at 3 dpi. Enrichment analysis reveals that those genes are significantly enriched in the processes of post-embryonic organ morphogenesis and regulation of nitrogen compound metabolic process (Appendix Table A5).

2.4.6 Identification of avirulence candidates that contribute to gene-for-gene resistance at B301

Although both SG3 and SG4 can trigger the resistant response at B301, the percentage and timing of HR occurrence is significantly different between those two races (Figure 2.1B). The cowpea-*Striga* interaction is an example of gene-for-gene (GFG) resistance (Li & Timko 2009) in which a specific *R* gene from host provides resistance to the pathogen that produces the corresponding *Avr* gene product (Flor 1971). R gene *RSG3-301* has been isolated and found to be responsible for race-specific HR when B301 cowpea is challenged with SG3 (Li & Timko 2009). To identify the counterpart of avirulence gene candidates to RSG3-301 in GFG model, we start by identifying the genes that are specifically expressed in SG3 compared to SG4 during the interaction with resistant host B301. There are 513, 1,704 and 632 contigs significantly upregulated in SG3 in relative to SG4 at 0 dpi, 3 dpi and 10 dpi, respectively (Figure 2.5A and Appendix Table A2). Among those, 1,668 (93.0%) contigs are differentially expressed at least one attachment stage. Since Avr genes are a special list of effectors that are induced at host interaction stages and secreted to apoplast or host cell to take effect, we apply the effector identification pipeline (Figure 2.3A) to narrow the list down to 260 candidates specifically upregulated in SG3 in at least one tested stage (Figure 2.5B). There are 201 candidates even showed higher expression at SG3 relative to SG4z expression profiles. Hierarchical clustering of correlation between candidate genes reveals that 71 Avr candidates are specifically expressed in SG3 compared to other two

races at all host interaction stages (Figure 2.5C). Although half of them do not contain domains or homology to genes with previously annotated functions, 32 of SG3 Avr candidates can be annotated with pre-classified plant orthogroups (Appendix Table A5). One interesting candidate (SGall_097996.2) is cysteine proteinases superfamily protein which has been previously identified as avirulence effector in several parasites (Shao et al. 2002; Axtell et al. 2001; Nimchuk et al. 2000). It degrades decoy protein RIN4 to trigger HR in resistance plant (Mackey et al. 2003).

2.5 Discussion

In this chapter, we used large-scale transcriptome sequencing to explore expression difference among the three genetic distinguishable races of *S. gesnerioides*. The large comparative analysis has made it possible to identify two sets of genes that we believe to be responsible for the hypervirulence of SG4z that overcomes resistance of cowpea cultivars and race-specific HR at SG3-B301 interaction. These candidate genes will be a valuable resource for future functional studies to understand the genetic changes that lead to the differentiation of distinct *Striga* races.

2.5.1 Gene expression difference between SG4z and SG4

SG4z and SG4 share high genetic similarity but interact differently with B301. SG4 fails to grow on B301 and triggers host resistance while SG4z overcomes the B301 resistance and successfully infests hosts' roots to deprive nutrient for its own growth (Botanga & Timko 2006). To detect transcriptomic difference that may contribute to the race-specificity, we identified 267 contigs differentially expressed between SG4z and SG4 at host interaction stages regardless of resistant and susceptible cowpea. The genes that are particularly induced in SG4 include those involved in proteolysis and protein metabolic process. These genes may directly lead to the breakdown of metabolism and cell death of SG4.

In a study of the incompatible interaction of Florida population SGFL with cowpea cultivar Blackeye, it has been found that, instead of rapid cell death caused by HR at host:parasite interaction site, the growth of SGFL is halted at the tubercle and the internal connection between parasite and host is disorganized (Botanga & Timko 2005). We observed similar phenotype between the SG4-B301 interactions in which the growth of SG4 ceasses at attachment without obvious HR at 10 dpi. It is proposed that an improper exchange of developmental cues misdirects *Striga* development and terminates *Striga* growth at certain stage (Botanga & Timko 2005). The

upregulation of cell communication and signal transduction in the evolutionarily closest race, SG4z, may reestablish the normal development of *Striga,* and enable SG4z to thrive on the resistance hosts and to successfully establish vascular integrity with host roots.

Regardless of induced expression in SG4 or SG4z, the genes that are differentially expressed between races are significantly enriched in extracellular region, which is a typical feature of pathogenic effectors. Using expression and structure criteria of known effectors in eukaryotic parasites, we annotated 34 contigs as effector candidates. Among 13 SG4z upregulated effector candidates, one contig encoding beta-D-xylosidase 4 is exclusively induced in SG4z relative to SG4 at both susceptible and resistant hosts. Beta-D-xylosidases has been known to be involved in the remodeling of xylans in vascular development (Arsovski et al. 2009) and is particularly expressed in tissue undergoing secondary cell wall thickening (Goujon et al. 2003). It is co-expressed with beta-galactosidase and works together to regulate cell wall metabolism (Montes et al. 2008). The induction of beta-D-xylosidases and beta-galactosidase in SG4z implies that hardening cell walls of *Striga* and inducing vascular development may directly assist the penetration of the parasite across the host root cortex and contribute to the vascular connection between parasite and host roots. In fact, enzymatic modification of the host cell walls has been suggested to be important in the invasion of the parasite towards host root cortex as it attempts to connect with host vascular tissues (Kuijt 1977). The genes encoding cell wall modifying enzymes are specifically enriched among pre-haustorial and haustorial stages of several parasitic plants (Ranjan et al. 2014; Yang et al. 2015). They include glycosyl hydrolase, pectate lyases, pectin methylesterase, cellulases, and expansins. The specific upregulation of cell wall modification protein in SG4z suggests hypervirulence of SG4z may partially depend on its ability to break host cell walls while harden its own cell walls to successfully penetrate host root cortex.

In addition to cell wall modification enzymes, one contig encoding Leucine-rich Receptor-like protein kinase (LRR-RLK) is also exclusively upregulated in SG4z during interaction with both Blackeye and B301. It shares high similarity with Somatic Embryogenesis Receptor Kinase (SERK) at conserved domain region but lacks predictable protein kinase domain. The tomato SERK homologs, SlSERK3A and SlSERK3B, interact in vivo with the Flagellin Sensing 2 receptor (FLS2) and are actively involved in plant innate immunity (Peng & Kaloshian 2014). It has been proposed that SlSERK3A and SlSERK3B are tomato orthologs of Brassinosteroid (BR) Insensitive 1-Associated Kinase 1 (BAK1). LRR-RLK, specifically expressed in SG4z, shares LRR domain (interaction domain) with SERK while differs in protein size and lacks kinase

domain (biochemical activity domain), suggesting it might function as a SERK mimic to interfere host resistance and to assist successful invasion of *Striga*.

In contrast to effector candidates that are highly expressed in SG4z, many of contigs that are dramatically induced in SG4 relative to SG4z have unknown functions and are not orthologous to any known plant proteins. It is consistent with many Avr effectors discovered in other parasites (Aggarwal et al. 2014; Park et al. 2012; Innes et al. 1993) that share no significant sequence similarity with known proteins. *Avr* genes are usually at front line of host-parasite arm race. They are highly diverse and evolve rapidly. The unknown effector candidates significantly upregulated in SG4 may constitute those Avr proteins that trigger B301 resistance and deserve further functional verifications.

2.5.2 Gene expression changes in SG3 and SG4 at incompatible interaction

Different from SG4 and SG4z which trigger different response on B301, both SG3 and SG4 exhibit the incompatible interaction with B301. We found that accompanying the rapid HR and arrested growth in those two races is the downregulation of development and biosynthesis pathways and upregulation of stress response which are featured by response to endoplasmic reticulum stress and defense response by callose deposition at both SG3 and SG4. The genes that regulate cell death and plant immunity are also upregulated at incompatible interactions in both races. This change in parasite is similar to the expression trend in cowpea, in which genes involved in cell wall biogenesis and processes leading to cell death are particularly activated in host side in incompatible interactions (Huang et al. 2012a). This is not unexpected since *Striga* and cowpea behave as counterparts in incompatible interaction. From one hand, the host immunity suppresses the *Striga* development by strengthening the physical barrier to prevent parasite ingress and blocking the nutrient transfer with local cell death at attachment site. Without nutrient supply, *Striga* stops its biosynthesis pathway and triggers cell death pathway. As a result, we observed the upregulation of calmodulin binding protein which activates Ca^{2+} dependent ROS production (Kawarazaki et al. 2013) and metacaspase-1 (SGall_066648.2 and SGall_066648.3) which directly regulates cell death (Coll et al. 2010). On the other hand, the plant immunity pathway is upregulated in SG3 and SG4 during incompatible interactions, implying that *Striga* as a plant may experience host resistance as invasion to trigger innate defense. Careful examination of the expression change in SG3 and SG4 reveals transient induction of NPR1 (SGall_062822.1 and SGall_062822.2) which is central to the activation of SA-dependent defense, and consistent upregulation of a list of ABC transporters PDR12 (SGall_068308.1, SGall_005274.1,

SGall_005274.2, SGall_025127.1) which are frequently induced by fungal and bacterial parasites and may be involved in transporting toxic secondary metabolites (Campbell et al. 2003).

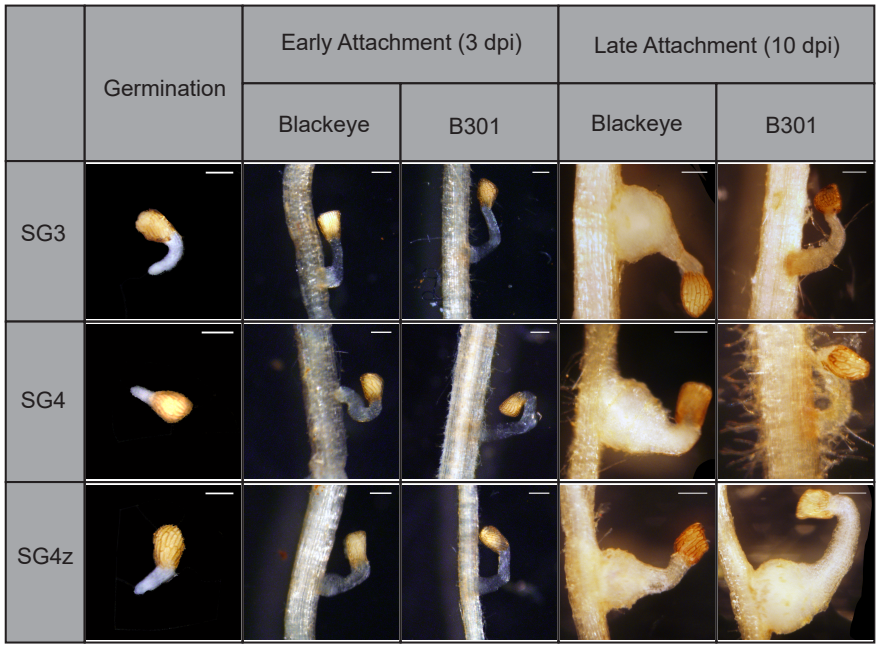
Besides common response shared by SG3 and SG4 at incompatible interaction, each race triggers race-specific response that acts as gene-for-gene resistance. The corresponding resistance gene RSG3-301 has been isolated and found to be responsible for race-specific HR when B301 cowpea is challenged with SG3 (Li & Timko 2009). Using race differential expression analysis, we found 71 avirulence candidates in SG3 which may directly or indirectly trigger *RSG3-301* gene-mediated HR. More than half of those candidates (60.5%) are functionally unannotated and show no homology to any known plant proteins. This is consistent with fact that Avr proteins share little sequence homology due to its rapid evolution reinforced by gene-for-gene relationship (Huang et al. 2014).

Cysteine proteases have been identified as Avr(s) in multiple parasites (eg. avrRpt2, YopT and AvrPphB) and trigger gene-for-gene resistance in plant containing certain R genes (Shao et al. 2002; Axtell et al. 2001). They cleave decoy proteins like PBS1 and RIN4 which are surveilled by R proteins and consequently activate HR. We observed one cysteine protease (SGall_097996.2) that is specifically expressed in SG3 and significantly induced after host contacts. Whether this cysteine protease is the counterpart gene of RSG3-301 in gene-for-gene model remains to be explored in future studies.

Interestingly, two proteins that are recurrently identified as plant immune genes are also specifically expressed in SG3 and significantly induced at host contact stages. Chitinase (SGall_091137.1) confers disease resistance by binding to fungal chitin and degrading fungal cell walls. It is induced to express in response to biotic stress and attack fungal parasites directly. The CC-NBS-LRR *R* gene (SGall_021378.1) mediates plant resistant response by direct or indirect interaction with Avr. *Striga*, like all other plants, is under attack from various parasites and triggers biotic stress response accordingly. The specific upregulation of those two genes in SG3 is likely to be associated with the particular defense response in those parasites or host resistance. Whether these race-specific immunity genes in SG3 are able to trigger gene-for-gene resistance requires additional study.

In conclusion, this study provides the first transcriptome wide comparison on three races of *S. gesnerioides* and reveals a list of candidate genes for future functional investigation on the virulence and avirulence of the parasitic plants.

A.


Germination
Early Attachment (3 dpi)
Late Attachment (10 dpi)
Blackeye
B301
Blackeye
B301
SG3
SG4
SG4z

B.

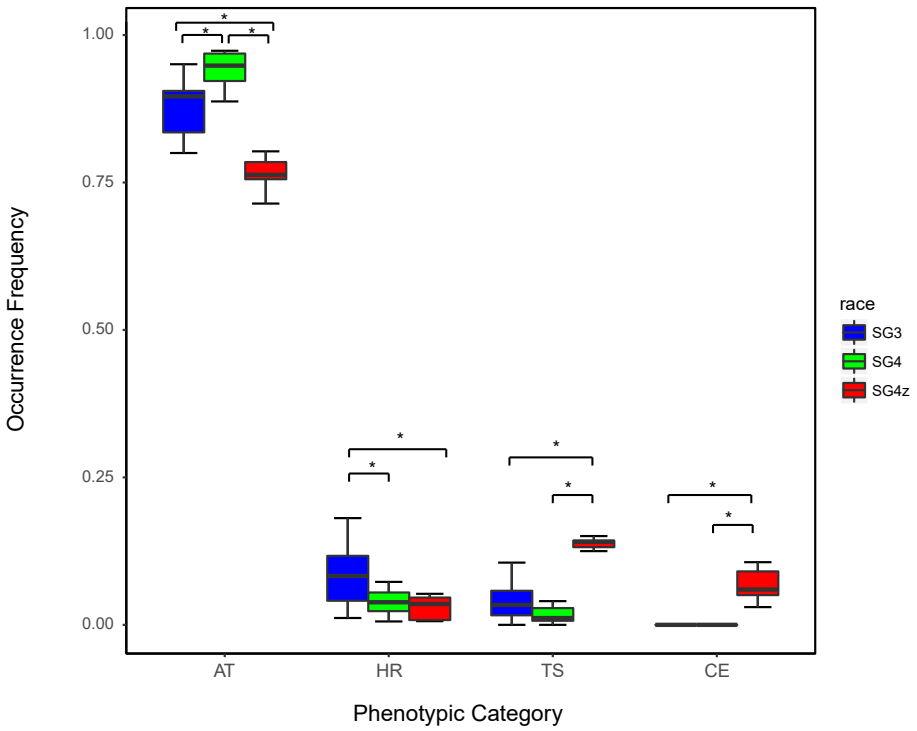
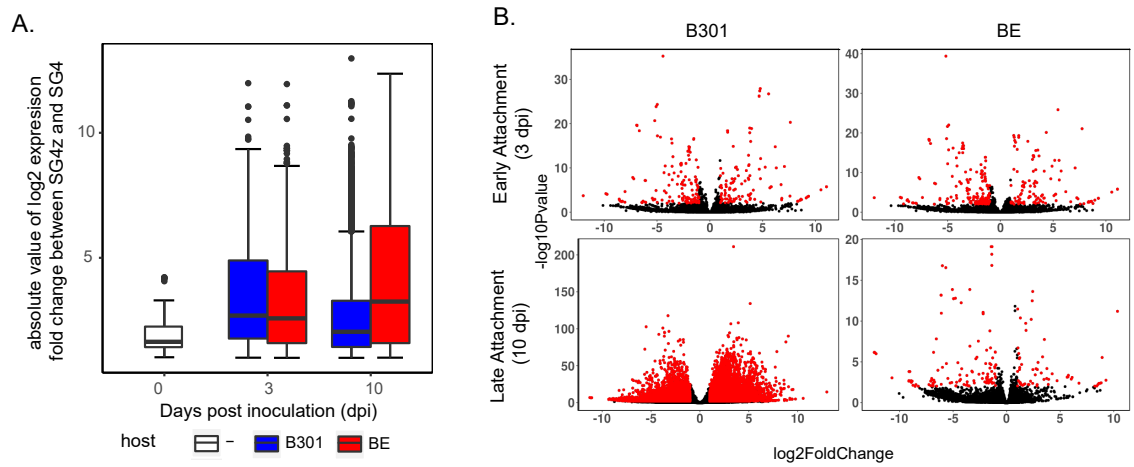
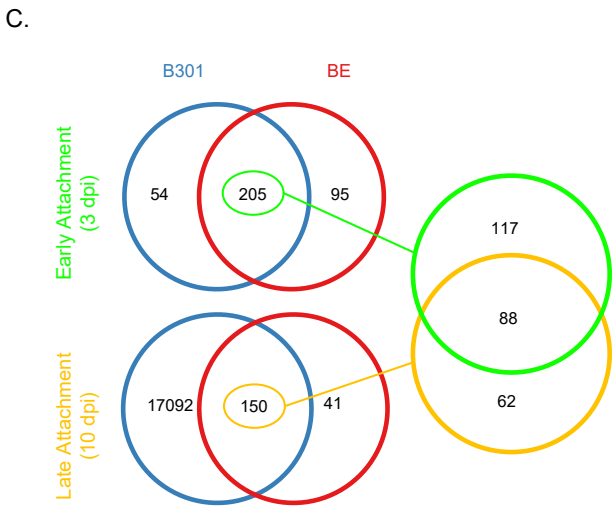

Occurrence Frequency
1.00
0.75
0.50
0.25
0.00
AT
HR
TS
CE
Phenotypic Category
race
SG3
SG4
SG4z

Figure 2.1. Differential response of cowpea cultivars B301 and California Blackeye No. 5 to parasitism by different races of *S. gesnerioides* races from West Africa. (A) Representative photographs illustrating the phenotypic response of B301, a multirace resistant cultivar, and California Blackeye No. 5 (BE) a susceptible cultivar, to parasitism by different races of *Striga gesnerioides*. Shown are the appearance of 2 day-old germinated SG4, SG4z and SG3 seedlings prior to host contact (0-day post-inoculation (dpi)) and the roots of cowpea cultivars B301 and BE at 3 dpi and 10 dpi. Scale bar represents 200 um. **(B)** Measured frequency of each different phenotypic event category during the interaction of the three *S. gesnerioides* races, SG3, SG4, and SG4z with resistant (B301) cowpea roots at 10 dpi. The abbreviation of the phenotypic event categories are as follows: AT, Attachment; HR, hypersensitive response; TS, tubercle swelling; and CE, cotyledon expansion. The interaction event ratio for each category was obtained by counting the number of each event category and dividing by the total number of phenotypic events occurring on each host plant. Statistical significance was determined using the t-test with a total of 10 independent host plants inoculated with equivalent amounts of the different races. An asterisk (*) indicates a p value < 0.05.

Table 2.1. Transcriptome assembly statistics and gene capture statistics

	SG3	SG4	SG4z	SGall
Transcriptome Size (Mbp)	111.5	95.7	90.5	115
Transcript Number	110298	98287	85568	145407
Unigene Number	81210	72487	61991	114231
Medium Length	679	631	732	475
Contig N50 length	1469	1526	1594	1218
Number of N50 contig	23567	19911	18171	27168
UCO (Ath)	2376	2375	2371	2382
UCO Proportion	82.82%	82.78%	82.64%	83.03%
COSII (220)	201	202	199	200
COSII Proportion	91.36%	91.82%	90.45%	90.91%
PlantTribe	901	904	901	905
Proportion	93.95%	94.26%	93.95%	94.37%


A.
absolute value of log2 expresison
fold change between SG4z and SG4
10
5
0
3
10
Days post inoculation (dpi)
host
B301
BE
B.
B301
BE
Early Attachment
(3 dpi)
Late Attachment
(10 dpi)
-log10Pvalue
log2FoldChange


C.
B301
BE
Early Attachment
(3 dpi)
Late Attachment
(10 dpi)
54
205
95
17092
150
41
117
88
62

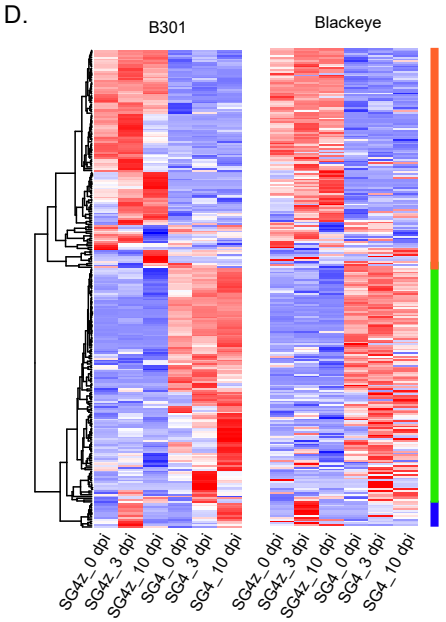

D.
B301
Blackeye
SG4z_0 dpi
SG4z_3 dpi
SG4z_10 dpi
SG4_0 dpi
SG4_3 dpi
SG4_10 dpi

Figure 2.2. Differential expression analysis between SG4 and SG4z. (A) The magnitude of expression difference between SG4z and SG4 across developmental stages. The magnitude of expression changes among differentially expressed contigs ($log_2FC > 1$ and $-log_{10}(FDR) > 2$) between SG4 and SG4z is represented as the absolute value of log2 fold change (log_2FC). The magnitude of expression changes is evaluated across three developmental stage: germination stage (0 dpi), early (3 dpi) and late (10 dpi) attachment stages, during the interaction with either host B301 or BE. **(B)** Volcano plots of gene expression comparisons between SG4 and SG4z. log_2 of fold change (log_2FC) between SG4 and SG4z is represented by x-axis. Log_2FC is positive if given contig is upregulated in SG4z compared to SG4. Statistical significance is indicated by $-log_{10}(FDR)$ on y-axis. Differentially expression genes (red) are indicated during the interaction with B301 or BE ($log_2FC > 1$ and $-log_{10}(FDR) > 2$) at early (3 dpi) and late (10 dpi) attachment stages. **(C)** Venn diagrams of race differentially expressed contigs at different stages during the interaction with B301 and BE. Numbers of contigs differentially expressed are used to make venn diagrams between early (3 dpi) and late (10 dpi) attachment stages during its interaction with resistant (B301) and susceptible (BE) cowpea cultivars. **(D)** The expression profiles of contigs show race-specific differential expression between SG4z and SG4. Scaled expression level (the mean of log_2TPM) is represented by color intensity from low (red) to high (blue) in heatmap. Upregulated contigs in SG4z (red bar) and in SG4 (green bar) are grouped into different clusters. Blue bar indicates a group of genes that are highly expressed in SG4z at early attachment stages while upregulated in SG4 at late attachment stages. Hierarchical clustering is performed based on pairwise correlation between contigs.

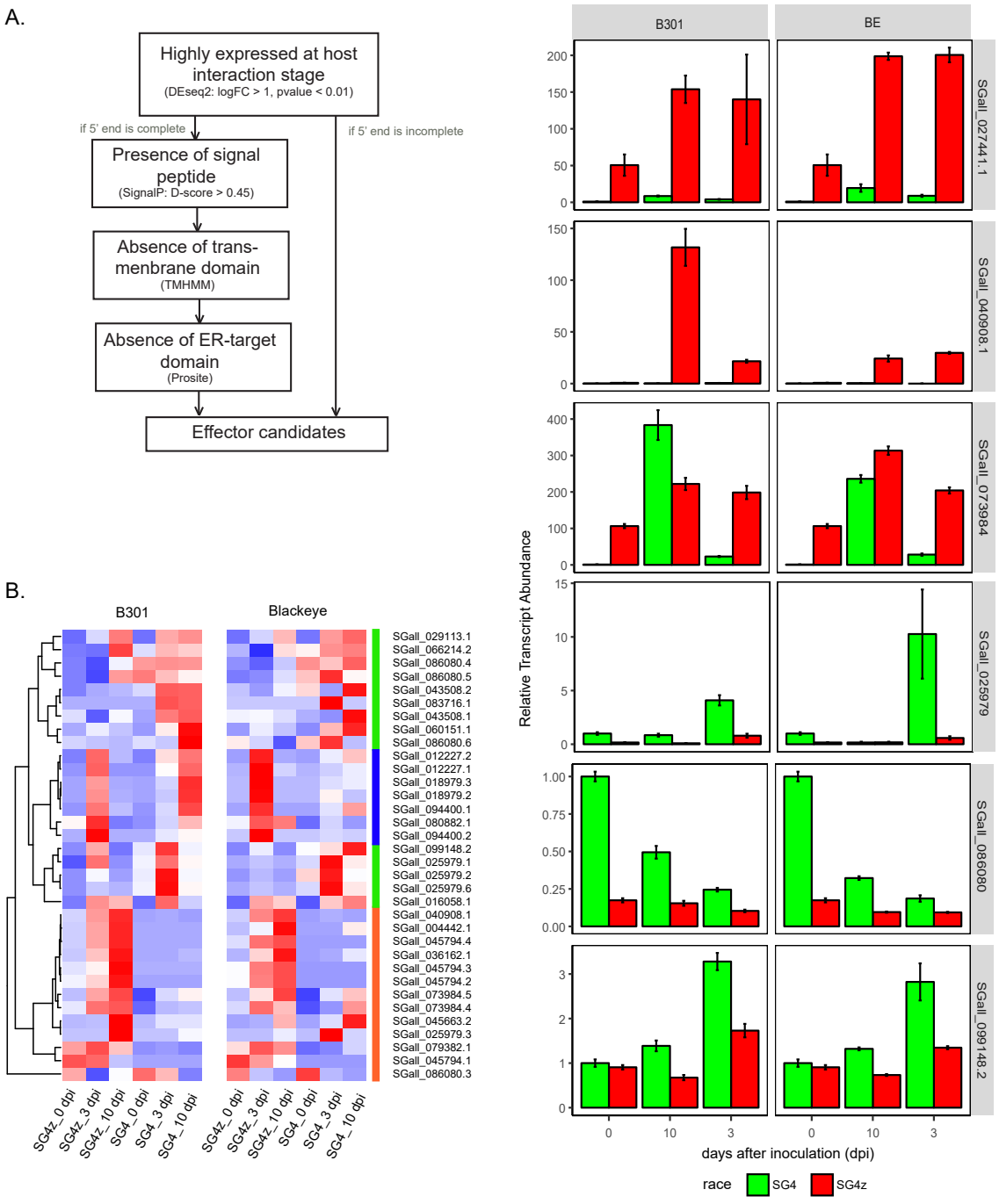
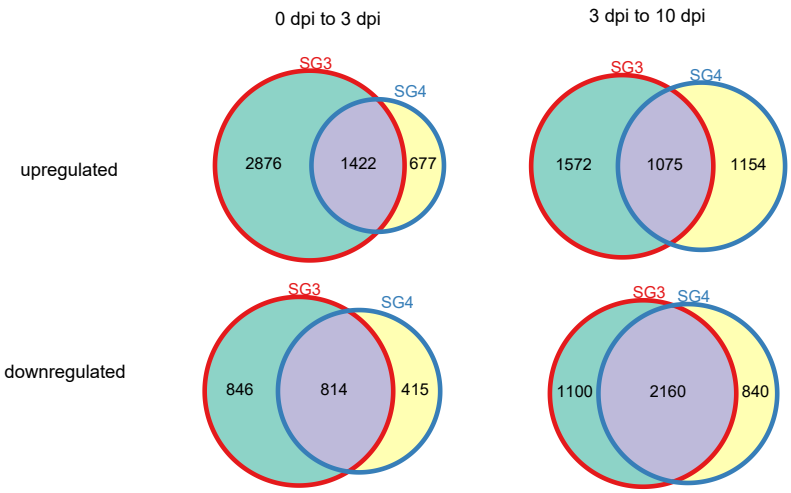

A.
Highly expressed at host interaction stage
(DEseq2: logFC > 1, pvalue < 0.01)
if 5' end is complete
if 5' end is incomplete
Presence of signal peptide
(SignalP: D-score > 0.45)
Absence of trans-menbrane domain
(TMHMM)
Absence of ER-target domain
(Prosite)
Effector candidates
B.
B301
Blackeye
SGall_029113.1
SGall_066214.2
SGall_086080.4
SGall_086080.5
SGall_043508.2
SGall_083716.1
SGall_043508.1
SGall_060151.1
SGall_086080.6
SGall_012227.2
SGall_012227.1
SGall_018979.3
SGall_018979.2
SGall_094400.1
SGall_080882.1
SGall_094400.2
SGall_099148.2
SGall_025979.1
SGall_025979.2
SGall_025979.6
SGall_016058.1
SGall_040908.1
SGall_004442.1
SGall_045794.4
SGall_036162.1
SGall_045794.3
SGall_045794.2
SGall_073984.5
SGall_073984.4
SGall_045663.2
SGall_025979.3
SGall_079382.1
SGall_045794.1
SGall_086080.3
SG4z_0 dpi
SG4z_3 dpi
SG4z_10 dpi
SG4_0 dpi
SG4_3 dpi
SG4_10 dpi
B301
BE
SGall_027441.1
SGall_040908.1
SGall_073984
SGall_025979
SGall_086080
SGall_099148.2
Relative Transcript Abundance
days after inoculation (dpi)
race
SG4
SG4z

Figure 2.3. Identification of effector candidates that differentially expressed in SG4 and SG4z. (A) The workflow of predicting effector candidates in *S. gesnerioides* using transcriptome data. **(B)** Expression profile of effector candidates over development. The expression level (the mean of $\log_2$TPM) is represented by color intensity from low (red) to high (blue) in heatmap. Upregulated effector candidates in SG4z (red bar) and upregulated effector candidates in SG4 (green bar) are grouped into different clusters. Blue bar indicates a group of candidates that are highly expressed in SG4z at early attachment stages while upregulated in SG4 at late attachment stages. Hierarchical clustering is performed based on pairwise correlation between contigs. **(C)** Results of qRT-PCR verification of selected effector candidates. Relative expression (abundance of the transcript) of selected contigs in SG4 (green) and SG4z (red) at 0, 3 and 10 dpi are represented in the bar plot. Three replications are performed to obtain statistical significance.

A.


0 dpi to 3 dpi
3 dpi to 10 dpi
upregulated
downregulated
SG3
SG4
2876
1422
677
1572
1075
1154
846
814
415
1100
2160
840

B.

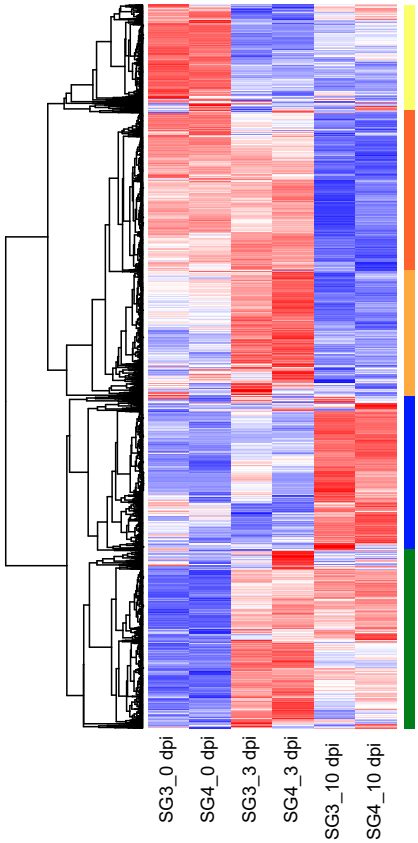
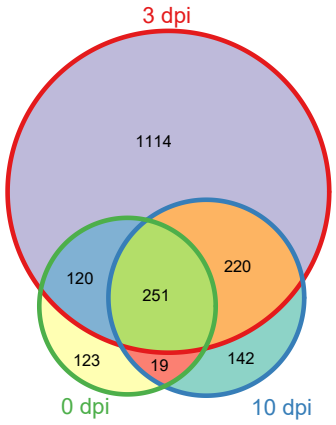
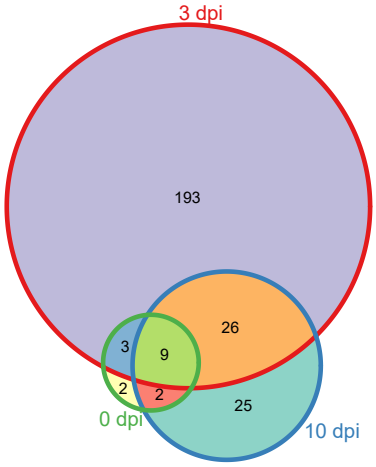

SG3_0 dpi
SG4_0 dpi
SG3_3 dpi
SG4_3 dpi
SG3_10 dpi
SG4_10 dpi

Figure 2.4. The expression profile changes in SG3 and SG4 in response to B301. (A) Venn diagrams of resistance responsive genes in SG3 and SG4. Numbers of contigs differentially regulated from 0 dpi to 3 dpi and from 3dpi to 10 dpi are listed in SG3 and SG4 during their interaction with B301. **(B)** The expression profiles of resistance responsive genes shared by SG3 and SG4 during the interaction with B301. The scaled expression level (the mean of $\log_2$TPM) is represented by color intensity from low (red) to high (blue) in heatmap. Hierarchical clustering reveals five distinguished genes clusters: downregulated at 3 dpi (yellow, 749 contigs), down-regulated at 10 dpi (red, 1,106 contigs), transit expression at 3 dpi (orange, 851 contigs), upregulated at 10 dpi (blue, 1,077 contigs) and upregulated 3 dpi (green, 1233 contigs).

A.


3 dpi
1114
120
251
220
123
19
142
0 dpi
10 dpi

B.


3 dpi
193
26
3
9
2
2
25
0 dpi
10 dpi

C.

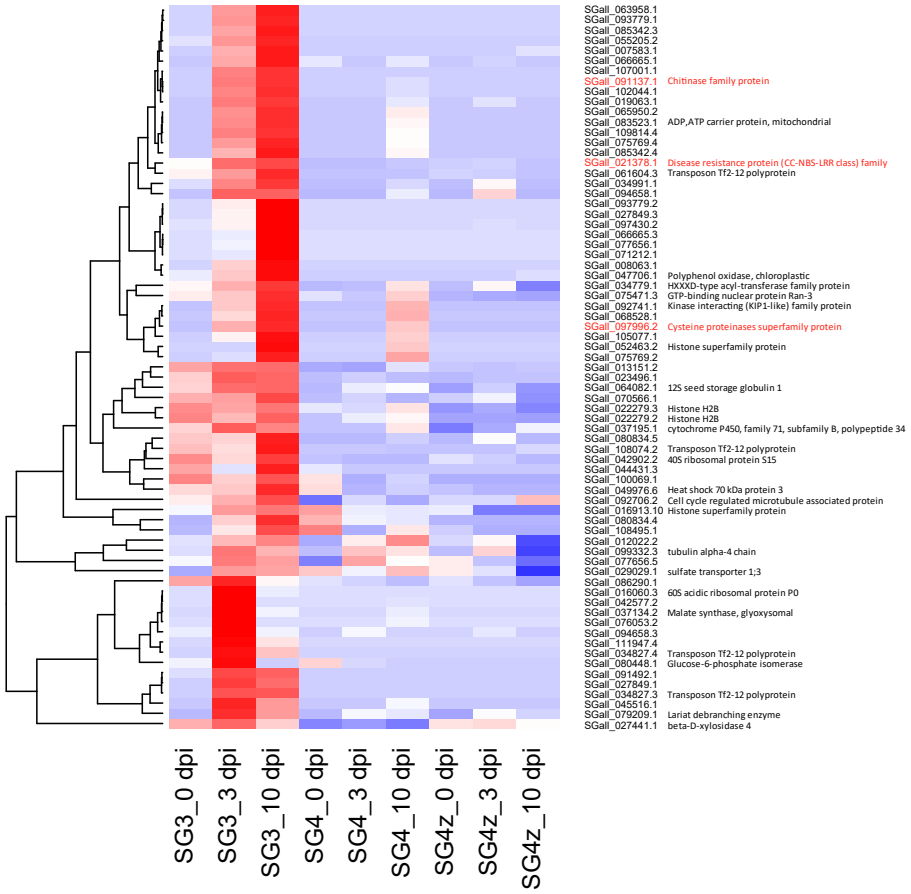

SGall_063958.1
SGall_093779.1
SGall_085342.3
SGall_055206.2
SGall_007583.1
SGall_066665.1
SGall_107001.1
SGall_091137.1 Chitinase family protein
SGall_102044.1
SGall_019063.1
SGall_065950.2
SGall_083523.1 ADP,ATP carrier protein, mitochondrial
SGall_109814.4
SGall_075769.4
SGall_085342.4
SGall_021378.1 Disease resistance protein (CC-NBS-LRR class) family
SGall_061604.3 Transposon Tf2-12 polyprotein
SGall_034991.1
SGall_094658.1
SGall_093779.2
SGall_027849.3
SGall_097430.2
SGall_066665.3
SGall_077656.1
SGall_071212.1
SGall_008063.1
SGall_047706.1 Polyphenol oxidase, chloroplastic
SGall_034779.1 HXXXD-type acyl-transferase family protein
SGall_075471.3 GTP-binding nuclear protein Ran-3
SGall_092741.1 Kinase interacting (KIP1-like) family protein
SGall_068528.1
SGall_097096.2 Cysteine proteinases superfamily protein
SGall_105077.1
SGall_052463.2 Histone superfamily protein
SGall_075769.2
SGall_013151.2
SGall_023496.1
SGall_064082.1 12S seed storage globulin 1
SGall_070566.1
SGall_022279.3 Histone H2B
SGall_022279.2 Histone H2B
SGall_037195.1 cytochrome P450, family 71, subfamily B, polypeptide 34
SGall_080834.5
SGall_108074.2 Transposon Tf2-12 polyprotein
SGall_042902.2 40S ribosomal protein S15
SGall_044431.3
SGall_100069.1
SGall_049976.6 Heat shock 70 kDa protein 3
SGall_092706.2 Cell cycle regulated microtubule associated protein
SGall_016913.10 Histone superfamily protein
SGall_080834.4
SGall_108495.1
SGall_012022.2
SGall_099332.3 tubulin alpha-4 chain
SGall_077656.5
SGall_029029.1 sulfate transporter 1;3
SGall_086290.1
SGall_016060.3 60S acidic ribosomal protein P0
SGall_042577.2
SGall_037134.2 Malate synthase, glyoxysomal
SGall_076053.2
SGall_094658.3
SGall_111947.4
SGall_034827.4 Transposon Tf2-12 polyprotein
SGall_080448.1 Glucose-6-phosphate isomerase
SGall_091492.1
SGall_027849.1
SGall_034827.3 Transposon Tf2-12 polyprotein
SGall_045516.1
SGall_079209.1 Lariat debranching enzyme
SGall_027441.1 beta-D-xylosidase 4
SG3_0 dpi
SG3_3 dpi
SG3_10 dpi
SG4_0 dpi
SG4_3 dpi
SG4_10 dpi
SG4z_0 dpi
SG4z_3 dpi
SG4z_10 dpi

Figure 2.5. Identification of Avr candidates in SG3. (A) Venn diagram of SG3 upregulated contigs in relative to SG4. Numbers of contigs significantly upregulated in SG3 were summarized at 0, 3 and 10 dpi during the interaction with B301. **(B)** Venn diagram of effector candidates upregulated in SG3 compared to SG4. *In silico* effector identification pipeline is applied to the contigs that are differentially expressed between SG4 and SG3. Numbers of effector candidates significantly upregulated in SG3 are summarized at 0, 3 and 10 dpi during the interaction with B301. **(C)** The expression profiles of SG3 Avr effector candidates. The expression level (the mean of log2TPM) is represented by color intensity from low (red) to high (blue) in heatmap. Hierarchical clustering is performed based on pairwise correlation between contigs.

SHR4z, an effector protein secreted from the haustorium of the parasitic weed *Striga gesnerioides* suppresses cowpea host plant innate immunity

3.1 Abstract

The resistance of cowpea to *S. gesnerioides* is remarked by hypersensitive response (HR) at attachment site. A hypervirulent race SG4z rising in the Republic of Benin successfully overcomes this resistance and parasitizes the most resistant cowpea accession B301. Comparative transcriptomics and *in silico* computational analysis in the last chapter uncovers a leucine-rich receptor (LRR)-protein kinase (PK) homolog dubbed *SHR4z* (Suppressor of Host Resistance 4z) that is highly expressed in SG4z haustoria and secreted into the host root. Overexpression of *SHR4z* in transgenic B301 roots leads to suppression of HR elicitation and loss of host innate immunity. SHR4z binds a host BTB-BACK domain containing ubiquitin E3 ligase homolog (VuPOB1) and silencing of *VuPOB1* expression in transgenic B301 roots lowers the frequency of HR and increases the levels of successful parasitism by SG4. In contrast, overexpression of *VuPOB1* results in decreased parasitism by SG4z suggesting VuPOB1 functions as a positive regulator of HR and plant innate immunity. These findings provide new insights into how parasitic weeds overcome host defenses and could potentially contribute to the development of novel strategies for controlling *Striga* and other parasitic weeds thereby enhancing crop productivity and food security globally.

3.2 Introduction

The capacity of plants to protect themselves against various parasites depends on their sophisticated defense mechanisms. The basal defense is triggered by recognition of evolutionarily conserved microbe or pathogen-associated molecular patterns (M/PAMPS) while secondary defense is initiated through specific gene-for-gene recognition events to activate R gene mediated resistance responses. One feature of the resistance response is the HR which involves a highly localized programmed cell death (PCD) at the infected region to constrain pathogen spread.

Cowpea (*Vigna unguiculata L.*), a major food and forage in West Africa, is exposed to a variety of biotic pests. Among the major biotic constraints is a parasitic plant *S. gesnerioides*. It invades the roots of cowpea and robs nutrient and water from host to survive. Although most cowpea plants are susceptible to *Striga* parasitism, some local cultivars and wild accessions have been

identified to be resistant to *Strig*a. Cultivar B301 is among the most resistant cowpea and induces HR when challenged by *Striga*. One canonical R gene (*RSG3-301*) from B301 was isolated and characterized to specifically mediate HR triggered by *Striga* race SG3 (Li & Timko 2009), implying gene-for-gene resistance in *Striga*-cowpea interactions. However, no effector counterpart has ever been discovered at the side of *Striga*.

Although B301 is resistant to a wide variety of *S. gesnerioides*, SG4z, from Zokpota at Republic of Benin, is able to overcome host resistance after B301 was introduced to this area 20 years ago (Lane, Moore, et al. 1993). *Striga* race SG4z can successfully suppress the HR, establish vascular connections with host, and expand its cotyledons leading to subsequent aboveground growth and flowering. Nonetheless, the mechanism of resistance suppression is still unknown in the context of parasitic plant and their hosts. Since SG4z is recent directive from its closest race SG4, these natural populations of *S. gesnerioides* provide a convenient system to inquiry the molecular mechanism underlying the virulence of the parasitic plant over host resistance.

In this chapter, we characterized a leucine rich receptor (LRR) protein kinase (PK) homolog dubbed *SHR4z* (suppressor of host resistance 4z) that is highly expressed in SG4z haustoria, secreted into the host root, and capable of suppressing the host innate immunity by binding to a host BTB-BACK domain containing ubiquitin E3 ligase homolog (VuPOB1). We show that silencing of *VuPOB1* expression in transgenic B301 lowers the frequency of HR and increases the levels of successful parasitism by SG4, whereas overexpression of *VuPOB1* resulted in decreased parasitism by SG4z suggesting VuPOB1 functions as a positive regulator of the HR response. These studies provide a novel insight into the role of secreted effectors as a part of the strategy used by parasitic weeds to overcome host immunity and complete their life-cycles.

3.3 Materials and Methods

3.3.1 Cloning from genomic DNA and individual seeds

To test genetic variation among the races, *S. gesnerioides* from each of seven races were growing on blackeye in pot for 60 days.10 individual above-ground seedlings were collected and grounded to fine powder in liquid nitrogen for CTAB based DNA extraction. The seedling DNA from each race was diluted to 50 ng/ul and used as template for PCR with 5 pairs of primers (Appendix Table A7).

PCR-based target DNA detection for individual seeds was performed using Phire Plant Direct PCR kit (Thermo Science). Each pre-germinated Individual seed was dried on a filter paper and added to Phire PCR mix using sharp forceps. Touchdown PCR was used to amplify target gene with primer pair LRR_F and LRR_R (Appendix Table A7). The annealing temperature (69 °C) was gradually lowered (0.5 °C per cycle) to 62.5 °C. A twenty-cycle amplification was then performed with an annealing temperature of 62 °C. The PCR products were separated by electrophoresis in 1.0% agarose gel and subjected to ethidium bromide staining. 20 individuals were tested for each race with three replicates. ANOVA was performed on detection ratio to test the genetic difference of target gene among races.

3.3.2 3' and 5' RACE

Total RNA was isolated from *Striga* SG4z seedling growing in contact with B301 10 dpi and two gene-specific forward and three reverse primers (Appendix Table A7, LRR_5RACE_R1, LRR_5RACE_R2, LRR_3RACE_F1, LRR_3RACE_F2) were designed for 3'RACE and 5'RACE respectively based on the partial transcript sequence of *SHR4z* found in the RNA-seq transcriptome assembly. 5'-RACE and 3'-RACE reactions were carried out using the FirstChoice RLM-RACE Kit (Ambion). For 5'RLM-RACE, 10 ug of total RNA was ligated to the 5' RNA oligo adapter after treatment with calf intestinal phosphatase and tobacco acid pyrophosphatase. cDNAs were synthesized using 6-mer random primers and further PCR reactions were carried out using three nested primers and 5'-adaptor primer supplied by the manufacturer. For 3'-RLM-RACE, 1 ug of total RNA was reversely transcribed using 3RACE oligo d(T) adapter. PCR was performed using 3' adapter primers and forward primers specific for SHR4z. RACE products were gel purified and blunt end cloned into pJet1.2 vector (CloneJET PCR Cloning Kit, Thermo Fisher) for sequencing. Fragment sequences were further assembled using CAP3 web service to obtain the full-length sequence of SHR4z.

3.3.3 Construction of transgene expression plasmids and generation of ex vitro composite plants.

The full length SHR4z protein was predicted based upon the first ATG at 5' end and longest open reading frame (ORF) to an in-frame stop codon and this coding region was cloned and mobilized into Gateway® pDONR™221 Vector (Invitrogen) using Gateway BP Clonase Enzyme mix (Invitrogen). A truncated version of the SHR4z full length protein in which the predicted signal peptide was removed (SHR4z**Δ**SP) was also cloned into the Gateway® pDONR™221 Vector. In SHR4z**Δ**SP an ATG codon was inserted at 5' end of truncated sequence to allow for translation initiation. Both pDONR-SHR4z and pDONR-SHR4z**Δ**SP were further recombined with

pK7WG2D (Karimi et al. 2002) using the LR recombination reaction in the Gateway LR Clonase Enzyme mix (Invitrogen, Grand Island, NY). pK7WG2D expressing small FLAG-tag was used as control. All pK7WG2D plasmids (pK7WG2D-SHR4z, pK7WG2D-SHR4z**Δ**SP and pK7WG2D-FLAG) were transformed into *Agrobacterium rhizogenes* strain R1000 for heterogeneously expression in cowpea cultivar B301.

B301 seeds were sterilized with 10% (v/v) bleach for 10 min and pre-germinated on sterile rockwools. *Ex vitro* composite plants were then generated as described by Mellor et al (2012) using *A. rhizogenes* strains carrying the pK7WG2D-SHR4z, pK7WG2D- SHR4z**Δ**SP and pK7WG2D-FLAG plasmids (Mellor et al. 2012). Seedlings with regenerated roots were moved to rhizotrons 20 d after transformation and grown at 30 C for 14 d before Striga inoculation as described below.

3.3.4 Striga-host root interactions assay and statistical analysis.

Seeds of the *S. gesnerioides* race SG4 and SG4z were surface sterilized with 10% bleach for 10 minutes and pre-conditioned at 30 C for 9 days. Seeds were pre-germinated by incubation with cowpea root exudates for 2 days (Mellor et al. 2012) and then gently transferred to transgenic and non-transgenic cowpea roots using a paintbrush. At 10, 20, and 30 days post-inoculation (dpi) the occurrence number of all parasite-host root interaction events (i.e., attachment, tubercle swelling, cotyledon expansion and hypersensitive response) were scored and the percentage for each interaction type determined for non-transgenic (no visible GFP) and transgenic (GFP expressing) roots. At total of 10 individual plants were used per treatment and the occurrence percentage for each interaction event were subject to paired t-test for statistical significance.

3.3.5 Yeast Two Hybrid (Y2H) screening

B301 cowpea roots were infected with SG4z and SG4 and at 3 and 10 dpi the parasite seedlings were removed and the host roots were harvested. Total RNA was isolated from ~10 g of root using PureLink® Plant RNA Reagent (Invitrogen) according to the manufacturer's protocol and a 1:1 mixed sample of RNA was used to construct a normalized cDNA 'prey' library in a modified pDEST22 vector by a commercial vendor (Bio S&T Inc, St. Laurent, Qubec, Canada).

Y2H screening was performed using ProQuest system (Invitrogen). PDONR-SHR4z**Δ**SP was recombined with pDEST32 to generate DNA binding domain 'bait' fusion (pDEST32-SHR4z**Δ**SP). PDEST32-SHR4z**Δ**SP was transformed to yeast strain MaV203 and positive

transformants selected by planting on Leu deficient YNB plates (Leu- YNB). A single verified transformant was selected for the preparation of competent cells and the competent cells were then transformed with the cowpea host root cDNA 'prey' library. The resulting transformants were plated on triple amino acid deficient YNB plates (Leu-/Trp-/Ura- plates) to select interaction-positive clones. The putative positive colonies were recovered on YAPD plates and blotted to Nylon membrane. The membrane with colonies blot was incubated with X-gal assay buffer (Z-buffer) containing 1mg/ml XGal to measure relative ß-galactosidase activity.

The plasmids from colonies giving strong positive XGal expression were sequenced and analyzed to retrieve in-frame insertion. The plasmids with relatively long in-frame insertions (more than 30 amino acids) were considered as putative host interacting protein candidates and were re-transformed into MaV203 containing either pDEST32 empty control vectors or pDEST32-SHR4z**Δ**SP expression vectors. This will rule out the false positive caused by the interaction between DNA binding domain and prey candidate. The pDEST22 with small FLAG-tag was used as control and transformed to above two MaV203 lines to rule out false positives due to interaction between activation domain and SHR4z**Δ**SP.

3.3.6 BiFC and confocal microscopy

A transcript (UP12_14781) encoding the full length cowpea VuPOB1 protein was retrieved from Cowpea EST database (Version 1.42 of "HarvEST:Cowpea"). Coding frame was determined by selecting the longest CDS with 100% match to Y2H in-frame insertion and cloned from a B301 root cDNA library using gene specific primers (Appendix Table A7, BTB_ORF_F and BTB_ORFnoT_R). The purified PCR products were sub-cloned into split YFP vector SPYCE via restriction enzyme digest (XbaI and XhoI) to generate VuPOB1-YC protein fusion. SHR4z**Δ**SP was inserted into split YFP vector SPYNE using the same method and form SHR4z**Δ**SP-YN.

All expression constructs were verified by sequencing and transformed into *A. tumefaciens* strain GV3101 for transient protein expression. Positive Agrobacterium colonies were cultured in YEP media for 24 h at 28C and used for Agroinfiltration into three of the largest leaves of a 6 week-old *N. benthamiana* plant as previously described (Wydro et al. 2006). At 3-day post infiltration infiltrated leaves were harvested and examined under Leica sp8 scanning confocal microscope. Split-YFP was excited with 488 nm laser with emissions collected between 497 – 544 nm.

3.3.7 Ectopic overexpression and RNAi silencing of VuPOB1 in B301 roots

A ~250 bp fragment within the coding region of VuPOB1 was amplified and mobilized into the Gateway® pDONR™221 Vector (Invitrogen) and then recombined into pK7GWIWG2D(II) (Karimi et al. 2002) to generate inverted repeat construct pK7GWIWG2D-VuPOB1-RNAi. The full length PCR amplicon of VuPOB1 was mobilized into a Gateway® pDONR™221 Vector (Invitrogen) generating pDONR-VuPOB1 (with C-terminal stop codon) and subsequently recombined with pK7WG2D (Karimi et al. 2002) to generate overexpression construct pK7WG2D-VuPOB1. The sequence verified vectors were then transformed into *A. rhizogenes* R1000 and used to generate *ex vitro* composite B301 plants as previously described (Mellor et al. 2012). Transgenic and non-transgenic roots were harvested separately and then analyzed by qRT-PCR to confirm silencing of VuPOB1 in B301 roots. qRT-PCR was performed as previously described (Schmittgen & Livak 2008) using Takara Bio SYBR Green Master Mix (cat#RR820B). Relative transcript levels were obtained from calibrating its threshold cycles of target genes with that of cowpea 18S. ΔΔC t analysis was used to calculated relative expression to control sample (non-transgenic roots). All experiments were performed with three independent replicates.

3.3.8 Phylogenetic relationship analysis

BTB domain (PF07707) and BACK domain (PF00651) were searched against annotated protein database of Arabidopsis (TAIR10), Soybean (Glycine max v1.1, Gmax189) and cowpea (v1.1) using profile Hidden Markov Models (HMM, E-value < 1e-5). Gene models that contain both BTB and BACK domain were used for phylogenetic analysis (AtNPR which contains BTB and Ankyrin repeat-containing domain was used as outgroup).

Similarly, LRR domain (PF00560) and Protein kinase domain (PF00069) were searched against cowpea (v1.1) protein database to identify the LRR-PKs in cowpea. The sequences with 3-4 LRR domains were collected and combined with five Arabidopsis SERK protein sequences (AT1G34210.1, AT1G71830.1, AT2G13790.1, AT2G13800.1 and AT4G33430.2) for phylogenetic analysis (AtFLS2 which contains LRR domains was used as outgroup).

Amino acid alignments were generated with MAFFT v7.245 (--maxiterate 1000 --localpair) (Katoh et al. 2002) and trimmed with trimAL v1.4 (Capella-Gutiérrez et al. 2009) to remove sites with less than 10% of the taxa (-gt 0.1). Maximum likelihood (ML) phylogenetic trees of amino acid alignments were generated using RAxML v8.2.2 (-f a –x 12345 –p 12345)(Stamatakis 2006) with the WAS and JTT amino acid substitution model (-m). To evaluate the reliability of the

branches on the tree, 100 pseudo-samples for the alignment were generated to estimate branch support using the bootstrap method (-# 100).

3.4 Results

3.4.1 SHR4z encodes an extracellular targeting protein that was highly expressed in SG4z

In the previous chapter, comparative transcriptomic analysis revealed 34 effector candidates that are differentially expressed between SG4 and SG4z at post-attachment stages. One candidate coding a Leucine Rich Repeat Receptor-like protein (LRR-RLP, SGall_040908.1), reproducibly shows higher (100-fold or greater) expression in SG4z compared to two other races at both 3 dpi and 10 dpi (Figure 3.1). Based on the dramatic expression difference, we selected this candidate for further detailed analysis.

Contig SGall_040908.1 contains only a partial coding sequence and therefore we perform 5'- and 3'-RACE using total RNA isolated from SG4z haustoria growing on B301 roots to obtain the full length 942 nt transcript encoding a 195 amino acid protein (Figure 3.2A). The encoded full length protein bears homology to known leucine-rich repeat (LRR) containing receptor protein kinases (RPK) and has a 25 amino acid extracellular (apoplastic) targeting signal peptide at the N-terminal. There are four LRRs in the protein with one located adjacent to the signal sequence at the N-terminal and three arranged in tandem near the C- terminal of the protein (Figure 3.2A). The predicted 3-D structure of the protein indicates that three tandem LRRs forms a stack of beta-bend ribbons which serve as potential protein interaction domain (Figure 3.2B). Based on the speculated role of this secreted effector in mediating the loss of HR in SG4z-B301 interactions we designate the gene as *Suppressor of Host Resistance 4z* (*SHR4z*).

To confirm the presence and the exon-intron structure of the *SHR4z* gene in the SG4z genome, 5 pairs of primers are designed across the full length coding region and used these in PCR amplifications with SG4z genomic DNA. The *SHR4z* gene is found to consist of 6 exons and 5 introns (Figure 3.2C). PCR amplifications using genomic DNA from the various races of *S. gesnerioides* in West Africa show that while *SHR4z* transcripts are only highly expressed in SG4z seedlings, homologs of the *SHR4z* genes are present in the genomes all of the other races of *S. gesnerioides* tested (Figure 3.2D and E). While the reason is not yet known, the ability to detect SHR4z in the genome of all individuals within populations representing each of the races varied, with SG4z, SG3 and SG5 having a significantly higher proportion of individuals giving an

amplicon (ANOVA p-value = 0.000102) (Figure 3.2E). This may simply reflect minor variation in sequences among the genes present and the ability to be recognized by the primers used.

3.4.2 Ectopic overexpression of SHR4z in the roots of B301 suppresses HR-PCD following parasite attack

To test the hypothesis that SHR4z is a secreted effector involved in manipulating host root innate immunity, we ectopically overexpress C- terminal mCherry fusions of the full length SHR4z protein and a truncated version of SHR4z lacking the apoplastic targeting signal peptide (SHR4zΔSP) in B301 roots (Appendix Figure 3A and B) using an *ex vitro* composite plant transformation system for cowpea (Mellor et al. 2012). We rationalize that the full length and truncated SHR4z proteins may function differently depending on whether they are secreted to the cell apoplast or retained in the cytoplasm. As shown in Figure 3.3A transgenic B301 roots expressing the full length SHR4z protein exhibit mCherry expression on the outer surface of the cowpea root cell, which is consistent with the secretion signal targeting the protein to the apoplast. In contrast, transgenic roots expressing SHR4zΔSP, the truncated secretion signal-less version of SHR4z, do not exhibit a similarly strong mCherry signal suggesting that either the truncated protein is turned over rapidly or is present at only low levels in the cytoplasm.

Transgenic and non-transformed control B301 roots expressing the SHR4z and SHR4zΔSP proteins are then challenged with *Striga* race SG4 known to elicit a strong HR in B301 roots (Figure 3.3B and C). Transgenic B301 roots expressing SHR4zΔSP shows a significantly lower frequency of HR events than the control non-transgenic roots at 10 dpi with SG4 (t-test p value < 0.05) and significantly more tubercle swelling events (t-test p value < 0.05) indicating that the overexpression of SHR4zΔSP in host root is suppressing the host resistance response. We did not observe any cotyledon expansion events on the transgenic B301 roots expressing SHR4zΔSP challenged with SG4. Cotyledon expansion is a well characterized phenotypic indicator of successful vascular connection between *Striga* and it host. The absence of cotyledon expansion on SHR4zΔSP expressing transgenic B301 roots suggests that despite its ability to suppress HR, SHR4z alone may not be sufficient to completely overcome host innate immunity to the extent that it fully allows successful parasite growth.

In contrast, B301 roots expressing the full length SHR4z does not show a corresponding decrease in HR frequency (Appendix Figure 4). We infer that targeting the SHR4z to the apoplast of the host cell makes it unavailable to influence host defense pathways. Collectively, these findings

support a role for SHR4z as a suppressor of the host resistance response in the host cells as we previously speculated.

To test whether SHR4z overexpression increases the susceptibility of B301 to SG4z parasitism, we compared the frequency of successful parasite growth on transgenic B301 roots expressing SHR4zΔSP to that of non-transgenic control roots (Figure 3.3D). A slight reduction in frequency of HR events is found (t-test p value < 0.05), but in general, no significant differences are found in the frequency of tubercle swelling or cotyledon expansion events between transgenic and non-transgenic roots challenged with SG4z. Therefore, we infer from this finding that SHR4zΔSP overexpression in B301 does not enhance host susceptibility to the already hypervirulent SG4z.

Since resistance to *Striga* in B301 is conferred in a race-specific manner involving a number of different race-specific R genes, we also tested to see whether the resistance response of B301 to other races of *S. gesnerioides* is altered by SHR4zΔSP overexpression. To this end we inoculated non-transgenic control and transgenic B301 roots expressing SHR4zΔSP with SG3. Intriguingly, transgenic B301 roots overexpressing SHR4z**Δ**SP have a slightly lower frequency of HR events than non-transgenic at 10 dpi but no dramatically reduction on HR frequent at 20 and 30 dpi (Figure 3.3E). This may reflect subtle differences among the parasite's abilities to differentially activate resistance mechanisms due to the nature of their Avr factors.

3.4.3 SHR4z physically interacts with cowpea POB1 E3 ligase

To identify host targets of SHR4z, we carried out a yeast two-hybrid (Y2H) screen using SHR4zΔSP as bait and a cDNA prey library constructed from RNA isolated from B301 roots parasitized by SG4 and SG4z. From a screen of ~2 $X10^6$ colonies, we identified 6 clones that gave a reproducibly strong positive signal (Figure 3.4A). The inserts from these clones were recovered, sequenced, and annotated and all six were found to contain the same insert encoding a BTB (Bric-a-Brac, Tramtrack, and Broad Complex)-BACK (BTB and C-terminal Kelch) domain-containing protein (Gingerich et al. 2007) with a high level of similarity (76% amino acid sequence identity) to the Arabidopsis AtPOB1 and *Nicotiana benthamiana* NbPOB1 proteins (76% and 74% amino acid sequence identity, respectively) (Appendix Figure 5A). Both AtPOB1 and NbPOB1 have been previously identified as essential ubiquitin-protein ligases (E3s) involved in plant immune responses where they function to degrade downstream protein PUB17 (Orosa et al. 2017).

Using HMM domain searching against the recently published cowpea reference genome assembly (Muñoz-Amatria et al. 2017), we identified three more BTB-BACK proteins in cowpea. Phylogenetic comparisons (Appendix Figure 5B) indicate that the cowpea BTB-BACK protein identified in our screen clusters with the other cowpea BTB-BACK proteins and all three fall in a clade along with the *Arabidopsis* AtPOB1/AtLRB2, AtLRB1, and tobacco NtPOB1 proteins. Based on its high level of sequence similarity and phylogenetic relatedness, we confidently designated the cowpea interacting protein as VuPOB. The two additional cowpea BTB-BACK domains proteins that also fall within this cluster we have designated as VuPOB2 (Vigun10g140200.1) and VuPOB3 (Vigun06g120500.1).

The predicted macromolecular structures of SHR4z and VuPOB1 indicates a strong likelihood for interaction between the two proteins (Appendix Figure 5C). To confirm a physical interaction between VuPOB1 and SHR4z**Δ**SP *in vivo*, bimolecular fluorescence complementation (BiFC) assays were performed using a split YFP system in *N. benthamiana* leaves. Single and paired plasmids expressing the SHR4z**Δ**SP-YN and VuPOB1-YC constructs were Agro-infiltrated into *N. benthamiana* leaves and at 3 d post infiltration confocal microscopy was performed. Reconstituted YFP fluorescence could only be seen in the cytoplasm of the transgenic cells expressing both constructs (VuPOB1 and SHR4z**Δ**SP) indicating that these two proteins interact with the host cell cytoplasm (Figure 3.4B).

3.4.4 VuPOB1 is a positive regulator of cowpea HR upon Striga parasitism

To uncover how VuPOB1 may be functioning in cowpea innate immunity, we first measured the levels of VuPOB1 transcripts in the various compatible and incompatible cowpea-*Striga* interactions. qRT-PCR analysis of VuPOB1 transcript levels in B301 roots parasitized by SG4z and SG4 showed a substantial but transient increase in VuPOB1 transcript levels at 3 dpi compared to 10 dpi. In contrast, no significant expression change was observed in BE roots subjected to similar parasitism (Figure 3.5A). Interestingly, in the microarray analysis of differential gene expression during compatible and incompatible *Striga*-cowpea interactions, Huang et al. (2012) observed that among the genes most significantly up-regulated (5.57- and 6.04-fold induced at 6 dpi and 13 dpi, respectively) in B301 roots parasitized by *S. gesnerioides* race SG3 was *VuPOB2*, the paralog of *VuPOB1*(Huang et al. 2012a)(Appendix Figure 5 B). Cumulatively, these findings suggest a role for VuPOB1 (and perhaps the other paralogs) the cowpea resistance response to *Striga*.

To test whether VuPOB1 is necessary for the resistance response of B301, we used RNAi silencing to knock out *VuPOB1* expression in B301 roots on parasitism by SG4 and SG4z. As shown in Figure 3.5B, *VuPOB1*-silenced B301 roots challenged with SG4 showed a significantly lower frequency of HR events and a greater frequency of tubercle swellings events than non-transformed control B301 roots indicating that VuPOB1 is a necessary component of innate immunity. *VuPOB1*-silenced B301 roots challenged with SG4z showed slightly fewer HR events, significantly more tubercle swelling events, but comparable frequency of cotyledon expansion events (Appendix Figure 6). These data suggesting that silencing of VuPOB1 may only marginally enhance susceptibility of B301 to SG4z.

When we examined the effects of ectopically overexpressing *VuPOB1* on *Striga*-host interactions, a different outcome was observed in the normally successful parasitism of B301 by SG4z parasitism. Rather than the normal compatible interaction, B301 roots overexpressing *VuPOB1* display a higher frequency of HR events and no clear evidence of cotyledon expansion events among attached SH4z parasites **(**Figure. 3.5C**).** These data clearly indicate that overexpression of VuPOB1 results in decreased parasitism by SG4z suggesting that VuPOB1 functions as a positive regulator of HR and host immunity in cowpea.

3.5 Discussion

The use of secreted effectors to modulate host resistance is widespread in nature and has been described in interactions between plants and a wide variety of pathogenic microbes, fungi, and nematodes. This work extends the use of secreted effectors for host immunity manipulation to parasitic weeds thereby opening a new understanding of the mechanism of the intriguing plant-plant interactions. Our findings indicate that the SHR4z effector is formed in the parasite haustorium prior to host attachment and deployed in the early post-attachment stages of the host-*Striga* interaction. The 25-amino-acid secretion signal targets the SHR4z to the apoplast of the haustorial cell where it is available for entry into the host cell to carry out its suppressive function. Our data show that when the full length SHR4z protein is expressed in host roots (following *Agrobacterium*-mediated transient transformation), the apoplastic targeting signal at N-terminal of SHR4z moves the protein to the host cell apoplast where it is not functional in suppression. Thus, retention in the host cell cytoplasm is required for function. How transfer occurs from parasite apoplast to host cytoplasm and what molecules are involved remains to be determined

In oomycetes and fungus, it has been long known that the typical virulence effectors are among pathogenic secretome and have clear N-terminal signal peptide to allow translocating from pathogen cells to apoplastic region. They can function in apoplast and engage cell entry by overrepresented host-targeting motifs following N terminal signal peptide. The known host-targeting motifs, like RXLR, LFLAK, and CHXC amino acid sequences, were first discovered by examining sequence alignment of known AVR proteins (Rehmany et al. 2005) and further tested by molecularly manipulating motifs for translocating detection (Whisson et al. 2007). SHR4z functions in host cells suppress HR but it is still unclear which region is responsible for the translocation from apoplast to host cells. Expect for extracellular targeting signal, we did not detect typical fungal/oomycetes host-targeting motifs in SHR4z. In addition, with limited examples of known effectors in parasitic plants, it is challenging to pin down conserved motifs responsible for those cross-host-membrane effectors. Furthermore, as a parasitic plant, *Striga* shared all the common features as a normal plant (plasmodesma) and connected to host plants through vascular integrity. It is likely that effectors with small size were delivered through plant-specific protein transfer route that is totally independent of normal pathogenic molecular signals. We believe that, with advance of genetic manipulation in parasitic plants, the structural basis for protein trafficking between parasitic plant and host will become more explicit in near future.

In addition to the translocating signal to extracellular of *Striga* cells, SHR4z contains three Leucine-rich repeat motifs at C-terminal of protein, which typically is responsible for protein-protein interaction and present in many receptor kinases (RK) or receptor-like kinases (RLK) involved in plant immunity. The LRR motifs of SHR4z shared high sequence and structural similarity to a subgroup of LRR-RLKs called SERK (Somatic Embryogenesis Eeceptor-like Kinase) family. Arabidopsis contains 5 members of SERKs that participate in multiple signaling pathways including brassinosteroid signaling (Li et al. 2002; Nam & Li 2002), male sporogenesis (Albrecht et al. 2005; Colcombet et al. 2005) and plant immunity (Kemmerling et al. 2007; Chinchilla et al. 2007; Heese et al. 2007; He et al. 2007). AtSERK3/BAK1 interacts with FLS2 immediately after perception of bacterial flagellum and trigger plant innate immunity. It also works redundantly with AtSERK4/BKK1 to control cell death. The rice homolog of *SERK1* (*OsSERK1*) is known to be involved in defense response and significantly induced at the RNA level by rice blast fungus (H. Hu et al. 2005). Considering the high sequence and structural similarity of SHR4z to those SERK protein and the evidences of SHR4z suppressing host resistance in cowpea-*Striga* interaction, we suggest that SHR4z acts as a mimic of SERKs to interfere SERK-mediated plant immunity pathways.

More interestingly, although LRR motifs shared high similarity with that of SERK family, SHR4z lacks the kinase domain that was typically responsible for phosphorylation activity in SERKs. With the perception that LRR is the protein-protein interaction motifs, we inferred that SHR4z may bind to a protein that usually interacts with SERKs, but lack the ability to employ the biochemical activity as SERK normally acts. Yeast Two Hybrid screening reveals a BTB-BACK domain protein directly interacts with SHR4z and homologous to a conserved E3 ligase POB1. E3 ligase has been long found to be a common interactor with RK or RLKs. Arabidopsis Plant U-Box E3 ubiquitin ligases 12 (PUB12) and PUB13 have been shown to be BAK1 phosphorylation targets and degrades FLS2 to attenuate PTI in a flg22-dependent manner (Lu et al. 2011). The RING-type E3 ligase XB3 is required for the accumulation of the receptor-like kinase (RLK) protein XA21 and promotes downstream R-mediated HR against bacterial blast *Xanthomonas oryzae pv. Oryzae* (Wang et al. 2006). Two Plant U-Box (PUB) family of putative E3 ubiquitin ligases (Os08g01900 and Os01g66130) were found to interact with four phylogenetically distant defense-related RLKs (Os04g38480, Os07g35580, Os07g35260, and Os08g03020) in rice (Ding et al. 2009). It seems that E3 ligases have been shown to be regulating or being regulated by RLK in all steps of plant immune responses and constitute the central of plant defense signaling pathways (Duplan & Rivas 2014). There is no doubt that they become the targets of pathogenic effectors like SHR4z. In fact, the cases of manipulating host E3 ligase by effectors have been reported in several parasites before. The effector AVR3a from the oomycete *Phytophthora infestans* targets and stabilizes the U-box-type Ub-ligase CMPG1/PUB20 which controls *P. infestans* elicitin INF1 triggered cell death (Bos et al. 2010). Effector AvrPiz-t from the rice blast fungus *Magnaporthe oryzae* interacts and inhibits the rice RING-type Ub-ligase APIP6 which is normally a positive regulator of PTI. The binding of AvrPiz-t and APIP6 results in degradation of both proteins in *N. benthamiana* and suppresses flg22-induced ROS production (Park et al. 2012). Although how the interaction of SHR4z and VuPOB1 results in suppression of development of the HR in B301, we speculate that SHR4z interferes VuPOB1-mediated defense pathway by disrupting its normal interaction with certain member(s) of SERK family (Figure 3.6).

In tobacco, the NtPOB1/NbPOB1 homolog was recently found to target and degraded downstream U-box E3 ligase PUB17 which activates *Cf*-mediated HR-PCD (Yang et al. 2006; Orosa et al. 2017). In *Arabidopsis*, there are 49 members in PUB E3 ligase group (Azevedo et al. 2001; Mudgil et al. 2004; Mazzucotelli et al. 2006) and several members have been shown to be involved in plant immunity. In addition to PUB12 and PUB13 that attenuate PTI by

ubiquitination of the flg22 receptor FLS2 (Lu et al. 2011), PUB 22, 23, and 24 also act in combination to negatively regulate plant immunity by suppressing oxidative burst, the MPK3 activity, and transcriptional activation of defense regulatory genes (Trujillo et al. 2008). They are all negative regulators of plant defense and may be targeted and degraded by other E3 ligases like POB1. VuPOB1, however, appears to be a positive regulator of HR and therefore which members of the PUB E3 ligase family it interacts with remains to be determined.

It is clear that to mount a successful HR in B301 requires a number of downstream components that must be sensitive to the presence or activity of VuPOB1 (Figure 3.6). Previous studies from our group (Huang et al. 2012b; Mellor 2013) have shown that the transcriptional activator VuGRF is required for resistance response and that RNAi silencing of VuGRF leads to a loss of the SG3 induced HR. Similarly, silencing of a number of genes involved in cell wall modification (e.g., peroxidases, galacturonases, etc), general defense (e.g., narbonin, oxylipin biosynthesis, etc) and ROS release also decrease the ability to mount an effective HR that stimulates abiosis and leads to parasite death (Mellor 2013).

Finally, it is worth noting that the discovery of a secreted effector involved in mediating the interaction of *S. gesnerioides* with its cowpea host raises questions about the generality of this strategy among other parasitic weeds. Examination of the haustorial transcriptomes of other parasitic members in the Orobanchaceae (Yang et al. 2015) revealed presence of homologs of SHR4z in the transcriptomes of *S. hermonthica* and *Phelipanche aegyptiaca*, the former being a close relative to *S. gesnerioides* while the latter being much more evolutionarily diverged. Like SHR4z these homologs were also highly expressed in haustorial stages coincident with host contact. Whether these homologs are functional and acting in a similar manner to SHR4z remains to be determined.

Given that despite years of research, wide-scale effective control methods for *Striga* have remained elusive. The work described here not only enhances our understanding the molecular mechanism of how *Striga* overcomes host resistance, but our findings could potentially contribute to the development of novel strategies for controlling *Striga* and other parasitic weeds and thereby enhancing plant productivity and food security worldwide.

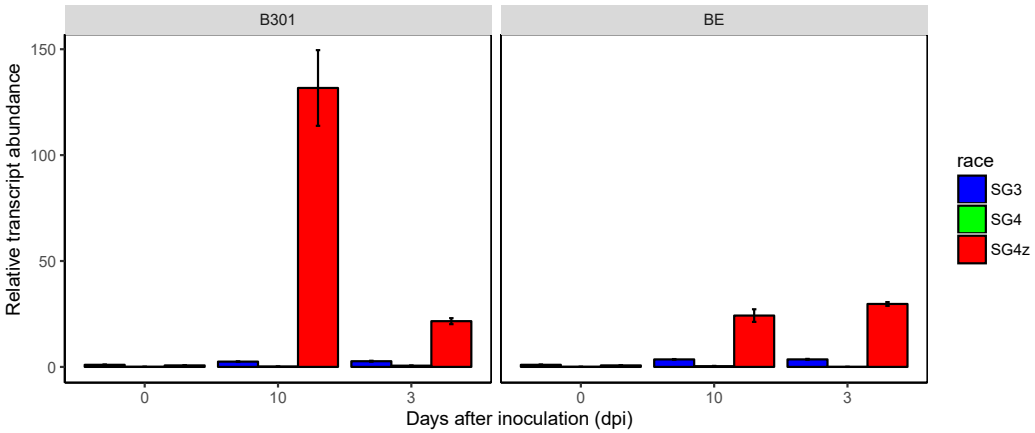

B301
BE
Relative transcript abundance
150
100
50
0
0
10
3
0
10
3
Days after inoculation (dpi)
race
SG3
SG4
SG4z

Figure 3.1. Shown is the relative abundance of the SGall_040908.1 transcript in germinated seedlings and haustoria of the three different *Striga* races at 3 dpi and 10 dpi as determined by qRT-PCR. The diagrams are coded as follows: SG3, blue; SG4, green; and SG4z, red. Three replications were performed to obtain statistical significance.

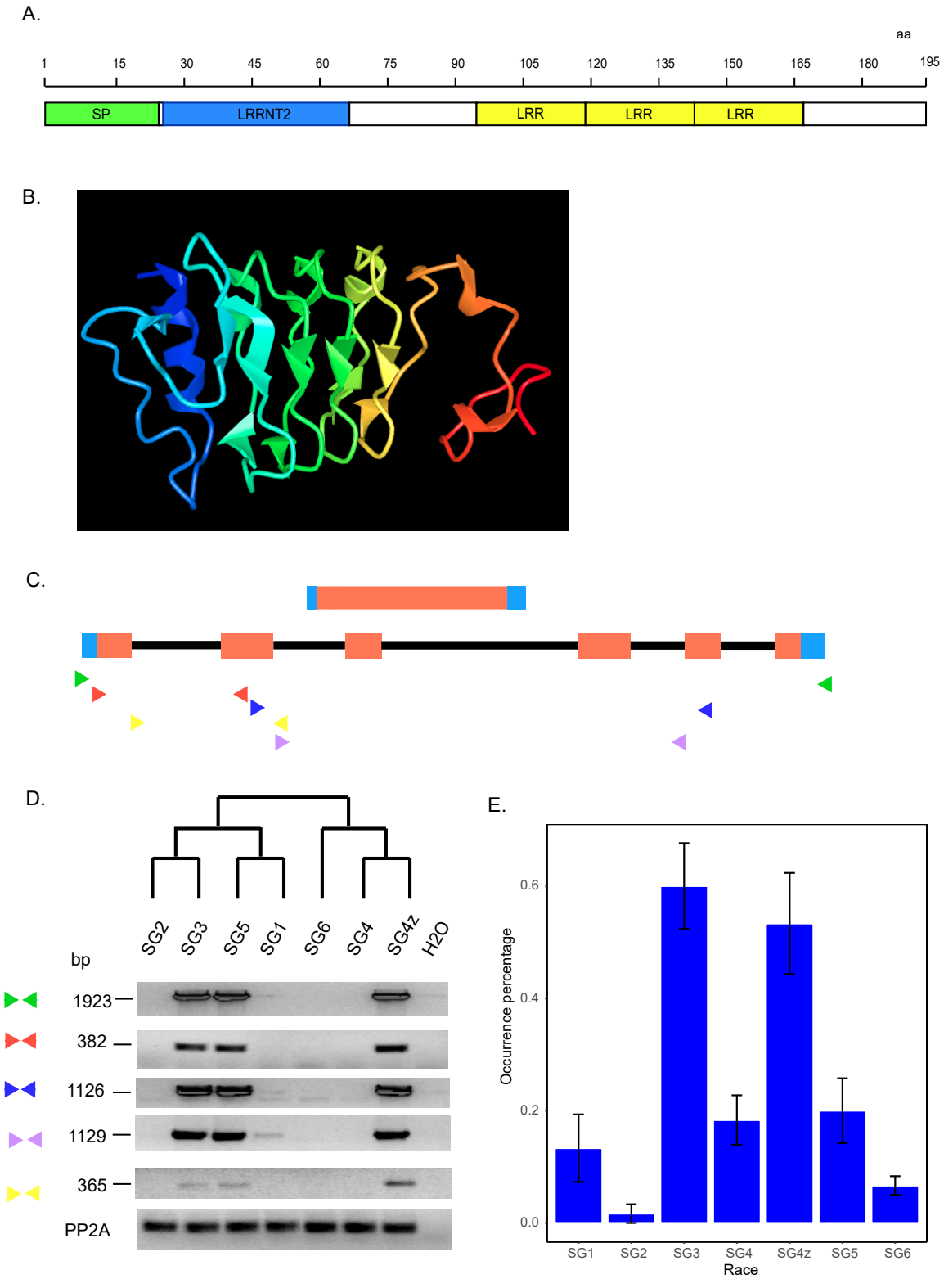

A.
aa
1
15
30
45
60
75
90
105
120
135
150
165
180
195
SP
LRRNT2
LRR
LRR
LRR
B.
C.
D.
SG2
SG3
SG5
SG1
SG6
SG4
SG4z
H2O
bp
1923
382
1126
1129
365
PP2A
E.
Occurrence percentage
0.0
0.2
0.4
0.6
SG1
SG2
SG3
SG4
SG4z
SG5
SG6
Race

Figure 3.2. SHR4z encodes a LRR-RPK protein with extracellular targeting signal peptide. (A) Motifs of the full length SHR4Z protein from *S. gesnerioides*. The location of the apoplastic targeting signal and leucine rich repeat regions within the SHR4z protein are indicated. **(B)** Predicted 3D protein structure of SHR4z. SWISS-MODEL was used to predict macromolecular structures of SHR4z based on sequence homology (Biasini et al. 2014). The 3D structure was viewed by Web-based 3D structure viewer iCn3D. **(C)** Structural organization of the SHR4z gene showing the location of exons and introns. Location of five pairs of gene specific primers (indicated by arrows with different colors. Green: LRR_F3R3; red: LRR_F2R2; blue: LRR_FR; yellow: LRR_F6R6; purple: LRR_F5R5, refer to Appendix Table A7) were used for PCR amplification of SHR4z homologs from the genomic DNA of the seven known races of *S. gesnerioides* found in West Africa. **(D)** A photograph of various PCR amplicons generated using five pairs of gene specific primers for the SHR4z gene. PCR amplifications were carried out using genomic DNA of the seven known races of *S. gesnerioides* found in West Africa and the various gene specific primers for SHR4z. Primers for the housekeeping gene PP2A was used as control. **(E)** Frequency of detection of the SHR4z gene within populations representative of the various races of *S. gesnerioides* in West Africa. PCR amplification was carried out directly on DNA of germinated seedlings isolated using the Phire Plant Direct PCR Kit (Thermo Scientific) and gene specific primer pair LRR_FR (blue in Figure 3.2C). The occurrence frequency was calculated from sample of 20 individuals in each population and three replications were performed to obtain statistical significance.

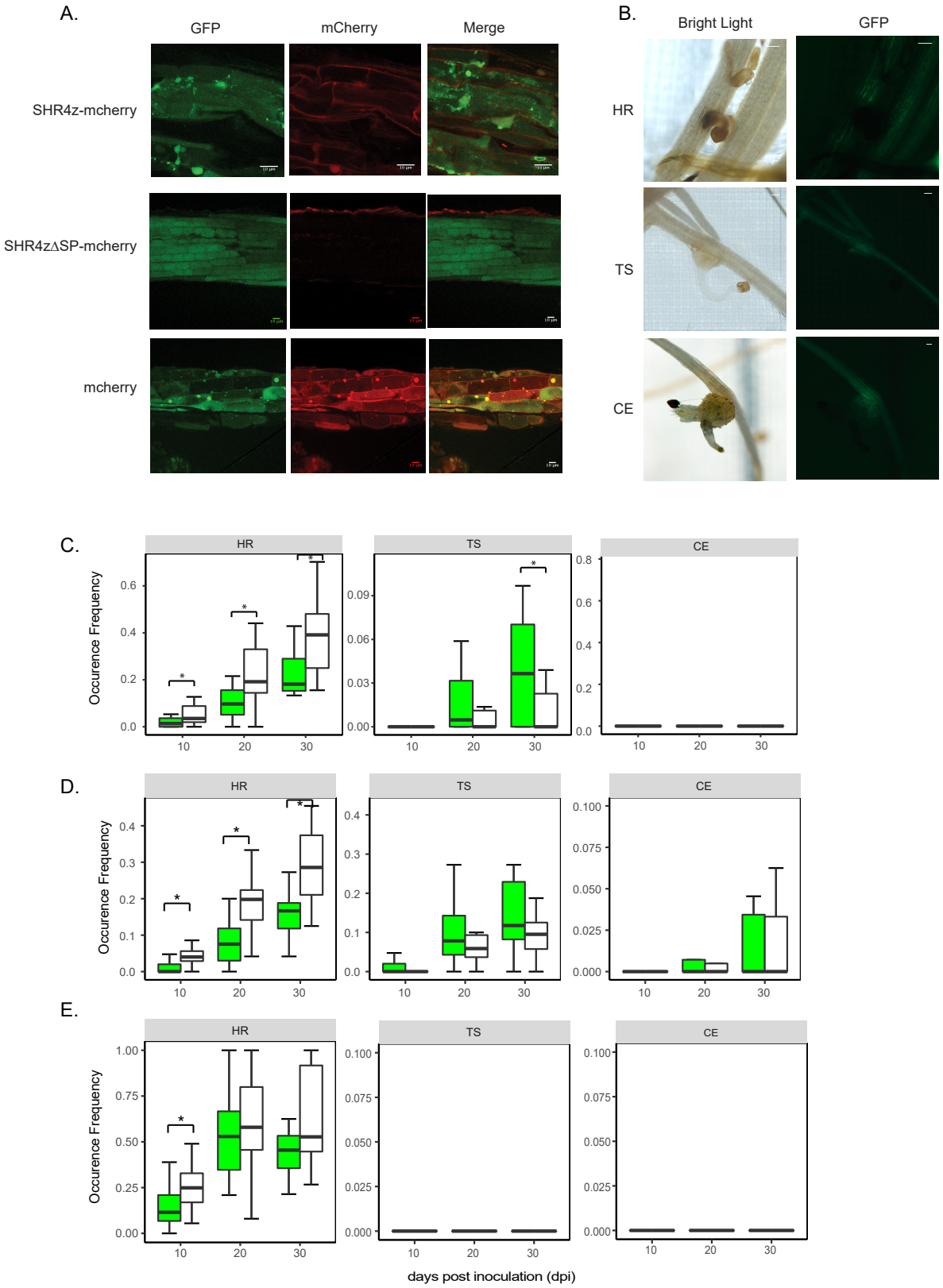
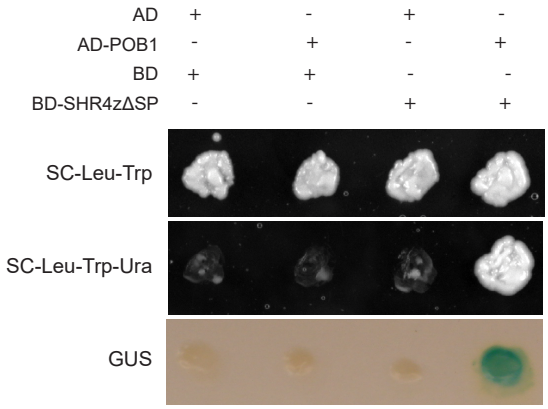

A.
GFP
mCherry
Merge
SHR4z-mcherry
SHR4zΔSP-mcherry
mcherry
B.
Bright Light
GFP
HR
TS
CE
C.
D.
E.
HR
TS
CE
Occurence Frequency
days post inoculation (dpi)

Figure 3.3 Ectopic overexpression of SHR4z suppresses cowpea host innate immunity. (A) Shown are representative photographs of the subcellular localization of SHR4z- and SHR4zΔSP-mCherry fusion proteins in transgenic B301 cowpea roots at 20 days post-transformation as viewed by confocal microscopy. B301 roots were transformed using A. rhizogenes containing the pK7WG2D-SHR4z-mcherry and pK7WG2D-SHR4z**Δ**SP-mcherry plasmids and the pK7WG2D-mcherry control constructs. fluorescence mCherry visualization was performed with a 564-630 nm filter range and GFP (filter range 497-544 nm) was used to label transformed root cells. The scale bar is 10 μm.

(B) Representative images of HR and tubercle swelling (TS) on composite plants. B301 roots from composite plants (pK7WG2D-SHR4z**Δ**SP construct) were inoculated with pre-germinated *S. gesnerioides* (SG4 and SG4z) at 30 days after transformation. Interactions (HR, TS and CE) were examined under light and fluorescence microscopy. The transgenic roots were indicated by GFP labeling. The scale bar is 200 um.

(C-E) Effects of ectopic overexpression of SHR4zΔSP on suppression of B301 innate immunity. *Ex vitro* composite B301 plants were generated by ectopically overexpressing the full length VuPOB1 transcript. Transgenic and non-transgenic roots were inoculated with two-day germinated SG4 (C), SG4z (D) and SG3 (D) seedlings, and the phenotypic responses of the roots were scored at 10 dpi, 20 dpi, and 30 dpi. The abbreviation of the phenotypic event categories are as follows: HR, hypersentive response; TS, tubercle swelling; and CE, cotyledon expansion. The interaction event ratio for each category was obtained by counting the number of each event category and dividing by the total number of phenotypic events occurring on transgenic (green bar) and non-transgenic (white bar) roots of each host plant. Statistical significance was determined using the t-test on th more than 10 independent host plant replicates. An asterisk (*) indicates a p value < 0.05.

A.


AD + - + -
AD-POB1 - + - +
BD + + - -
BD-SHR4zΔSP - - + +
SC-Leu-Trp
SC-Leu-Trp-Ura
GUS

B.

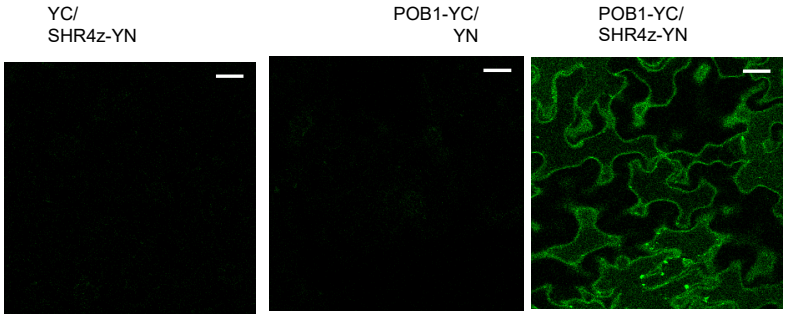
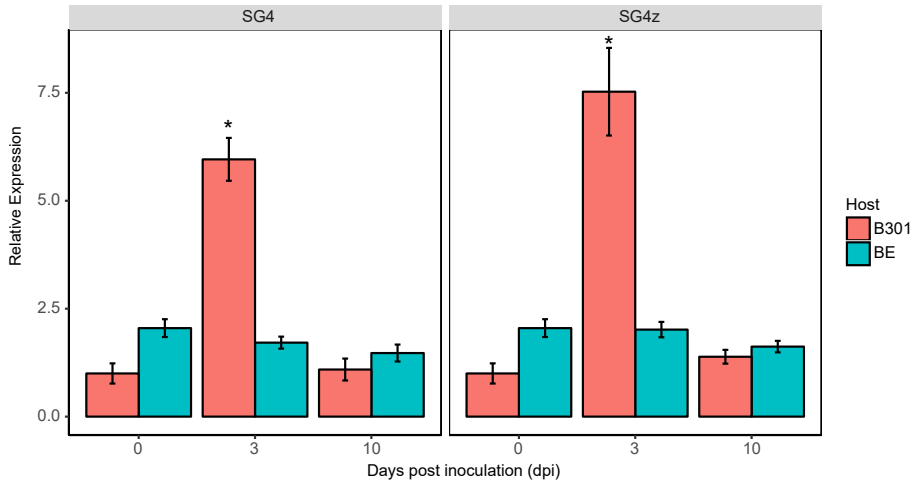
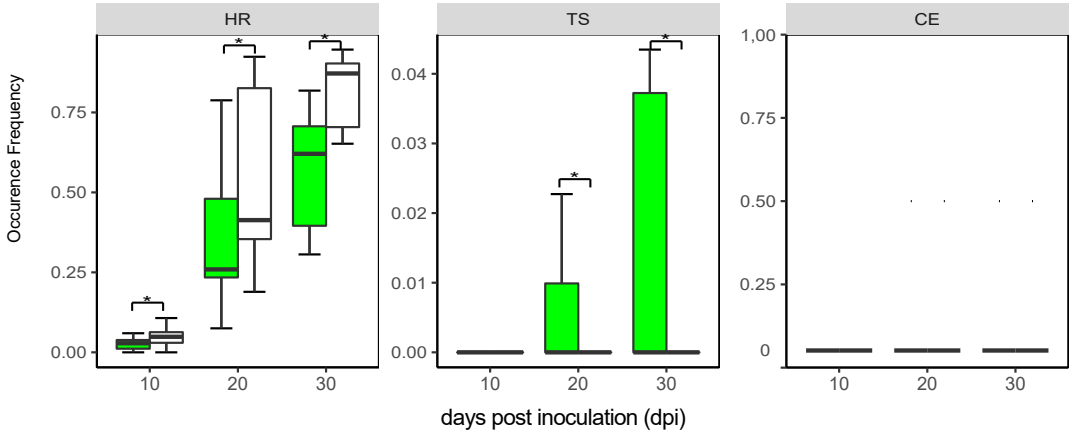

YC/
SHR4z-YN
POB1-YC/
YN
POB1-YC/
SHR4z-YN

Figure 3.4. SHR4z interacts with a BTB-BACK E3 ligase (POB1) in cowpea. **(A)** Shown are the results of yeast-2-hybrid assays indicating that SHR4z without signal peptide (SHR4zΔSP) interact with VuPOB1. The various indicated proteins were expressed as AD- and BD- fusions in MaV203 yeast cells. Transformed MaV203 cells were grown on selective media (Leu-, Trp-, Ura-) and X-gal media to test for interacting partners and on nonselective media (Leu-, Trp-) to test for the transformation efficiency. BD-FLAG and AD-FLAG fusion proteins were used as negative controls for interaction in yeast.
(B) Bimolecular fluorescence complementation (BiFC) assay demonstrating the *in planta* interaction of SHR4z with VuPOB1. Confocal microscopy was used to visualize the interaction of SHR4z with VuPOB1 following transient co-expression of plasmids containing YN-VuPOB1 and YC-SHR4z**Δ**SP in *N. benthamina* leaves. The scale bar is 25 μm.

A.


SG4
SG4z
Relative Expression
7.5
5.0
2.5
0.0
0
3
10
Days post inoculation (dpi)
Host
B301
BE

B.


HR
TS
CE
Occurence Frequency
0.75
0.50
0.25
0.00
0.04
0.03
0.02
0.01
1,00
10
20
30
days post inoculation (dpi)

C.

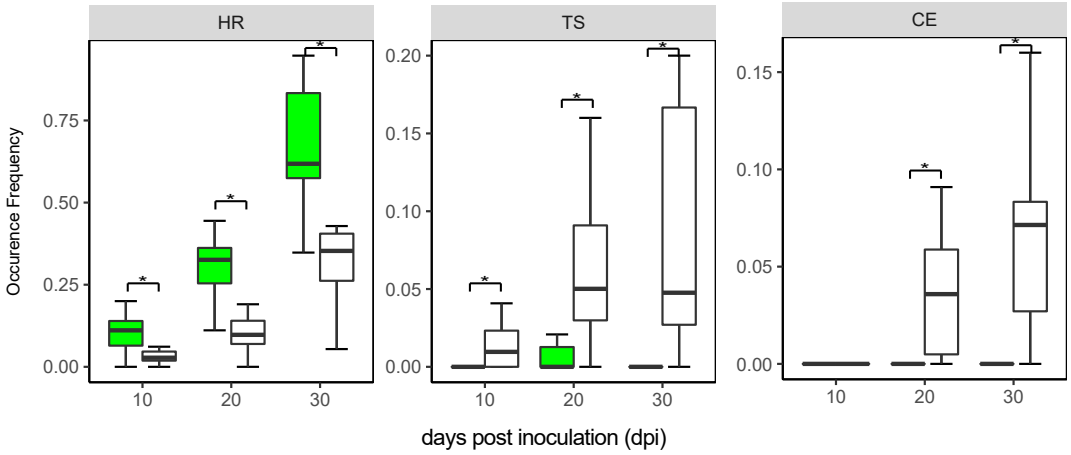

HR
TS
CE
Occurence Frequency
0.75
0.50
0.25
0.00
0.20
0.15
0.10
0.05
10
20
30
days post inoculation (dpi)

Figure 3.5. VuPOB1 functions as a positive regulator of plant defense.

(A) Shown is the relative transcript abundance (level of expression) of VuPOB1 in B301 and BE roots before parasite attachment and at 3 days-post inoculation (dpi) and 10 dpi with *Striga* races SG4 and SG4z. Relative expression was determined by qRT-PCR as described in the Materials and Methods. T-test was performed on three replications . The asterisk (*) indicates $p < 0.001$

(B) Effect of RNAi-silencing of VuPOB1 on SG4-triggered host resistance response. Ex vitro composite B301 plants were generated by expressing constructs capable of RNAi silencing of VuPOB1 expression. Transgenic and non-transgenic roots were inoculated with two-day germinated SG4 seedling, and at 10 dpi, 20 dpi, and 30 dpi the phenotypic responses of the roots scored. The abbreviation of the phenotypic event categories are as follows: HR, hypersensitive response; TS, tubercle swelling; and CE, cotyledon expansion. The interaction event ratio for each category was obtained by counting the number of each event category and dividing by the total number of phenotypic events occurring on each host plant. An asterisk (*) indicates a p-value < 0.05.

(C) Effects of ectopic overexpression of VuPOB1 on SG4z mediated suppression of B301 innate immunity. *Ex vitro* composite B301 plants were generated by ectopically overexpressing the full length VuPOB1 transcript. Transgenic and non-transgenic roots were inoculated with two-day germinated SG4z seedlings, and at 10 dpi, 20 dpi, and 30 dpi the phenotypic responses of the roots scored. The abbreviation of the phenotypic event categories and scoring of phenotypic categories are as above. An asterisk (*) indicates a p-value < 0.05.

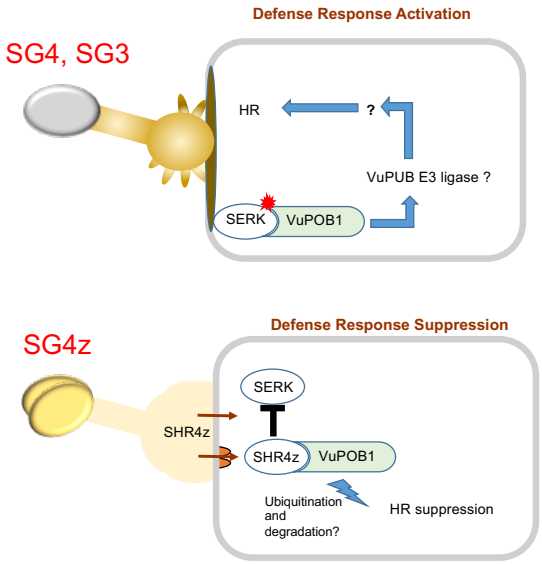


Figure 3.6. Model illustrating suppression of host defense response by SHR4z. SHR4z was translocated to host cells with assistance of specific transporters or directly via plasmodesmata connections established after parasite attachment. It directly targets VuPOB1 to prevent the interaction between SERK and VuPOB1. The binding of SHR4z to VuPOB1 results in subsequently ubiquitation and degradation reactions which may involve Plant U-box E3 ligase (PUB) and leads to suppression of host defense. In contact, SERK binds to VuPOB1 when SHR4z is absent in SG4 and SG3. The SERK-VuPOB1 complex activates the defense response and leads to HR.

Global expression changes of *Striga gesnerioides* during compatible and incompatible interactions with cowpea

4.1 Abstract

In this chapter, we compare expression profile of *S. gesnerioides* growing on resistant host B301 and susceptible host Blackeye (BE) to access the differentially expressed genes in response to different hosts. Different hosts induced distinct expression changes in SG3 and SG4. The disease resistance genes (NBS-LRR) are significantly upregulated in *Striga* haustorium in response to resistant host compared to susceptible host, while the expression of genes responsible for development are suppressed at resistant interaction. SG4z, which has adapted to parasitize B301, triggers the same response from B301 and Blackeye. However, SG4z still exhibits differential expression of 180 contigs at late infection stages in response to different host genotypes. Among those 64 contigs were suppressed in SG4z but induced in SG4, suggesting they may be associated with avirulence in *Striga*-cowpea interaction. These results enhance our understanding of plant-parasite interactions and the evolution of *Striga* in response to host selections.

4.2 Introduction

S. gesnerioides, as widely distributed witchweed, has an extensive host range including several wild and domestic legumes, members of the Conolvulaceaes, Euphorbiaceae and Solannaceae (Parker & Riches 1993). They are morphologically variable and have a number of strains within the species with a much narrower host range. Due to its autogamous reproduction lifestyle, *S. gesnerioides* successfully maintained the genetically-isolated host-specific morphotypes among various populations. The strain of *Striga* that adapted to tobacco fails to parasitize cowpea (Parker & Polniaszek 1990) while the Florida population that was developed on Indigofern fails to attack cowpea as well (Musselman & Parker 1981).

Local isolates of *Striga* that are growing in cowpea field exhibit extreme host specificity. For various cultivars, landraces, and breeding lines of cowpea, different populations of *S. gesnerioides* trigger different physiological responses, which are generally categorized as either ‘resistant/incompatible’ or ‘susceptible/compatible’ interaction. While susceptible interaction results in formation of a very large tuberous haustorium, the resistant interaction leads to arrest growth at attachment or HR at attachment site. Based on the different responses to cowpea and

genetic variability within and among populations of the parasite, Botanga and Timko (2006) distinguished seven races of *S. gesnerioides* in West Africa (Botanga & Timko 2006). One hypervirulent race SG4z in Zakpota from Republic of Benin diverges from its next closest race, SG4, and overcomes the resistance of cultivar B301 after the resistant cowpea grew in the field for 20 years. It suggests that the strong host selection is driving the differentiation of the *Striga* population. However, how *Striga* responds to the host selections is still unknown.

Here, I addressed the global gene expression changes of *Striga* during the compatible and incompatible interactions with cowpea. Using the *Striga* growing susceptible host BE as control, I first identified genes differentially expressed in each examined race during the interaction with B301. In SG4 and SG3 which trigger the resistance response in B301, common and race-specific resistance responsive genes were examined and analyzed for Gene Ontology (GO) term enrichment. The disease resistance genes (NBS-LRR) are significantly upregulated in *Striga* haustorium in response to resistant host compared to susceptible host, while the expression of genes responsible for development are suppressed at resistant interaction. In SG4z which infects both B301 and BE successfully, I still found 180 contigs differentially expressed at late infection stages in response to different hosts. The comparison of host-responsive genes between SG4 and SG4z revealed gene expression variation that may contribute to SG4z hypervirulence.

4.3 Methods and Materials

Methods and materials were generally same as described in Chapter 2. The differential expression analysis was performed using DESeq2 (Love et al. 2013). The two factor full generalized linear model (Host+Race+Race:Host) was used to examine host effect on differential expression of *Striga* genes. Using the interaction with Blackeye as control, differentially expressed genes due to host effect were identified in SG3, SG4 and SG4z, respectively (FDR < 0.01 and absolute value of log2FoldChange > 1). Since SG3 and SG4 both trigger HR after attaching to B301 roots while successfully grow on BE roots, differentially expressed genes shared by both races were subject to GO enrichment as described in Chapter 2 to obtain the common pathways or biological processes in response to resistant interactions.

In contrast, SG4z trigger same susceptible response from both B301 and BE. The differential gene expression in SG4z during interaction with different hosts is mainly caused by host effect instead of dramatic interaction response difference. Those genes were further subjected to effector identification pipeline (Figure 2.3A) and race-differential expression analysis described

in Chapter 2. Following the flowchart in Appendix Figure 7, we were able to identify the hyper-virulent effectors in SG4z that facilitate suppression of B301 resistance and Avr in SG4 that may trigger the HR in B301.

4.4 Results

4.4.1 The resistant responsive genes in SG3 and SG4

SG3 and SG4 both trigger HR after attaching to B301 roots while successfully grow on BE roots. To find host effect on gene expression profiles of *Striga*, we examined the expression changes in response to resistant host B301 compared to susceptible host BE. Very little expression change was detected at 3 dpi, with only 46 contigs differentially expressed in SG3 and no differentially expressed genes in SG4 (pvalue < 0.01, absolute value of $\log_2$foldchange > 1). However, at 10 dpi when phenotypic difference becomes obvious between resistant and susceptible interactions, I found 19,496 and 18,427 contigs differentially expressed during interaction with different hosts in SG3 and SG4, respectively (p-value < 0.01, $\log_2$foldchange > 1). 14,753 differentially expressed contigs were shared by both races (75.7% in SG3 and 80.1% in SG4), with 8,167 contigs induced and 6,583 suppressed at the resistant interaction. Functional annotation of these differentially expressed contigs revealed that genes that were suppressed in the resistant response were significantly enriched in the biological processes such as 'cell cycle', 'cell differentiation', 'cellular protein modification process', 'DNA metabolic processes' and 'embryo development'. The genes that were induced in resistant interaction were mainly responsible for 'response to stress', 'response to external stimulus' including 'response to biotic stimulus' and 'response to abiotic stimulus' (Figure 4.1). Interestingly, 32 contigs, which were annotated as the CC-NBS-LRR class disease resistance proteins, were significantly upregulated in the resistant interaction compared to susceptible interaction (Appendix Table A8). This CC-NBS-LRR family was overrepresented in the upregulated genes at resistant interaction (one tail hypergeometric test p-value = 1.103×10^{-5}), implying that host resistance may trigger the defense response in *Striga*.

4.4.2 Host effect on SG4z

In contrast to SG4 and SG3, SG4z seedlings fail to elicit HR on B301 roots. Similar to what is observed on the susceptible BE, the interaction is compatible between SG4z and B301. Thus we hypothesized that the expression profile of SG4z growing on B301 roots should be very similar to the one growing on BE. As expected, no significant expression difference was observed between

SG4z-BE and SG4z-B301 interactions at 3 dpi. However, at 10 dpi, there were 180 contigs significantly differentially expressed in SG4z during the interaction with different hosts (Appendix Table A9). It suggests that the genetic variations of hosts still subtly affect expression profile of the parasite although the phenotypic interactions are identical. GO analysis reveals that GO terms like 'immune effector process' (FDR = 0.028), 'plant-type cell wall loosening' (FDR = 0.028) and 'regulation of root development' (FDR = 0.028) are significantly enriched among genes that are differentially expressed in SG4z. More interestingly, those genes were significantly enriched at 'extracellular region' (FDR=0.022), 'cell wall' (FDR=0.022) and 'external encapsulating structure' (FDR=0.022), suggesting that the proteins encoded by these genes may be secreted. This is consistent with the definition of pathogenic effectors that were secreted by parasites to aid infection of host. Since one essential genetic difference in B301 relative to BE is the presence of *R* genes resistant to several *Striga* races (including SG4 which is evolutionarily close to SG4z), we hypothesized that some of these differentially expressed genes from SG4z may be hyper-virulent effectors that facilitate the suppression of B301 resistance.

To identify candidate genes that may contribute to the suppression of B301 resistance by SG4z, we examined how these differentially expressed genes in SG4z were regulated differently in SG4 in response to the host effect. The 'host effect' was represented by the expression changes at the interaction with B301 relative to the interaction with Blackeye (log2FoldChange(B301/BE)). No significant expression changes were observed at 3 dpi, suggesting that the host effect on SG4 and SG4z is not obvious at early attachment. At 10 dpi, 107 out of 180 contigs differentially expressed in SG4z are also differentially expressed in SG4z in response to the different host. Hierarchical clustering on expression fold changes due to host effect on SG4 and SG4z reveals that those 180 contigs were clearly clustered to 4 groups (Figure 4.2A). There were 69 contigs upregulated and 31 downregulated in both SG4 and SG4z while the other two groups were regulated differently. Sixty-four (64) contigs were significantly suppressed in SG4z but not in SG4, while 16 contigs were induced in SG4z but not in SG4 during interactions with B301 (Figure 4.2B, Appendix Figure 7).

Among 64 contigs of which expressions were suppressed at 10 dpi in SG4z, we found that 20 contigs were significantly induced at host interaction stages in SG4 and/or SG4z (Appendix Figure 7). However, at 10 dpi, the expression levels of these contigs are downregulated in SG4z while remained the same or was continuously increasing in SG4 during the interaction with B301 (Figure 4.2C). I also found that 19 out of these 20 contigs were race differentially expressed

and were significantly upregulated in SG4 relative to SG4z (Appendix Figure 7). The expression pattern implies that those genes may encode Avr factors which trigger the resistance response in SG4 while their expressions are suppressed in SG4z to escape the host immunity surveillance. Further gene annotation revealed that 14 out of 19 were classified to 11 plant gene orthogroups (Appendix Figure 7). These 11 gene orthogroups include CBL-interacting protein kinase, WRKY DNA-binding protein, transposon Tf2-12 polyprotein, ribonuclease H-like superfamily protein, phosphate transporter, IAA-amino acid hydrolase ILR1-like 3, glutathione S-transferase family protein, glutamate synthase, formin homology 14, disease resistance protein (CC-NBS-LRR class) family and calcium-transporting ATPase.

In contrast, among 16 contigs upregulated in SG4z, 11 were induced at host attachment stages and significantly upregulated at SG4z in relative to SG4 (Figure 4.2C, Appendix Figure 7). Seven (7) of them were annotated to plant gene orthogroups including DNA-directed RNA polymerase subunit alpha, protein ycf2 (ATPase with unknown function), chloroplastic 30S ribosomal protein 3, cox 19 like CHCH family protein, trichome birefringence-like (TBL) protein family, DNA glycosylase superfamily protein, and eukaryotic initiation factor 4A-III.

4.5 Discussion

In this study, we used large-scale transcriptome sequencing to investigate the nature of the gene expression changes in multiple races of the parasitic plant, *S. gesnerioides*, in response to compatible and incompatible interactions with cowpea.

During the incompatible interaction between *Striga* and cowpea, the growth of *Striga* halts at the initial attachment. We found that arrested growth accompanies down-regulation of genes involved in cell cycle and embryo development in *Striga* races SG3 and SG4 during the interaction with resistant cowpea cultivar B301. However, gene expression associated with abiotic stress and external stimulus response was significantly activated at incompatible interaction relative to the compatible interaction. These include genes involved in plant stress response and negative regulation of HR. The down-regulation of development genes and up-regulation of stress response genes may be a result of *Striga* shifting more of its energy to deal with the host resistance rather than growth.

Not only in *Striga*, defense genes show elevated expression in the haustoria of many other parasitic plants during the interaction with their hosts (Yang et al. 2015). It was proposed that

they were obtained by horizontal gene transfer (Yang et al. 2016) and co-opted by the parasitic plants for defense against other pathogens. In our data, 32 contigs encoding CC-NBS-LRR class disease resistance proteins were significantly upregulated at resistant interaction. Plant NBS-LRR proteins are one of the largest ancient plant gene families and are known to be the master regulators in disease resistance. NBS-LRR proteins trigger HR which imposes a heavy cost at plant development and therefore is likely to be tightly regulated to avoid tissue damage (McHale et al. 2006). Many known NBS-LRR proteins are only induced in perception of parasites (Yang et al. 2008; Li et al. 2013; Navarro et al. 2004; Zipfel et al. 2004). The upregulation of CC-NBS-LRR in *Striga* during incompatible interaction suggests that *Striga* may employ its defense pathways against host resistance and attempt to attenuate the attack of the host immunity system. However, the mechanism of these NBS-LRR genes in *Striga* defense against host immunity remains to be investigated.

In contrast to SG4 and SG3, SG4z show compatible interactions with both B301 and Blackeye. No significant difference was observed on the expression profile of SG4z raised on either B301 or Blackeye at 3 dpi, suggesting that the host effect on SG4z is not obvious at early developmental stages. However, at 10 dpi, we found 180 contigs differentially regulated in response to different hosts. This result implies that although both B301 and Blackeye are susceptible to SG4z, the different genetic background of hosts may affect the survivability of SG4z. In fact, we observed that SG4z seedlings were more likely to emerge from the underground when inoculating on BE than on B301, suggesting B301 is still more resistant to SG4z than Blackeye. Since B301 still exerts selective pressure on SG4z, the genes differentially regulated in SG4z in response to B301 compared to BE are likely to compensate the B301 constraint and contribute to the suppression of host resistance.

Comparing the host effect on SG4z and SG4, we found that more than half (100 out of 180 contigs, 55.6%) of the host responsive genes showed similar expression changes in SG4z and SG4. It suggests that SG4 is still adapting to overcome the host resistance. Among the host-responsive genes that were differentially regulated in SG4z and SG4, there were more genes suppressed compared to genes induced (64 vs. 16) in SG4z during the interaction with B301. This is consistent with our previous hypothesis that SG4z may lose Avr genes or decrease expression of Avr to escape the host immune surveillance. One of Avr candidates was annotated as IAA-amino acid hydrolase ILR1-like 3, which hydrolyzes certain amino acid conjugates of the plant growth regulator indole-3-acetic acid (IAA). Hydrolases have been identified as Avr/effectors in

many parasites (Win et al. 2012; Kong et al. 2015), and are responsible for degrading cell walls, modifying their host targets and eventually promoting virulence of parasites. Some of them, however, trigger the HR when R genes are present (Kong et al. 2015). We assumed that the transient upregulation of ILR1-like 3 in SG4z may maintain its virulence activity at early attachment stage while avoid its recognition by *R* gene at late attachment stage. Thus, *ILR1-like 3* will be an ideal candidate Avr in SG4. However, to test whether *ILR1-like 3* is a genuine *Avr* and to understand how it is regulated remains further characterization.

In conclusion, this chapter, from the perspective of host affect, examined expression changes of *Striga* transcriptome during the compatible and incompatible interactions. It not only provides an alternative method to identify effector candidates that may trigger or attenuate host resistance, but also enhances our understanding of the host selection on *Striga* populations.

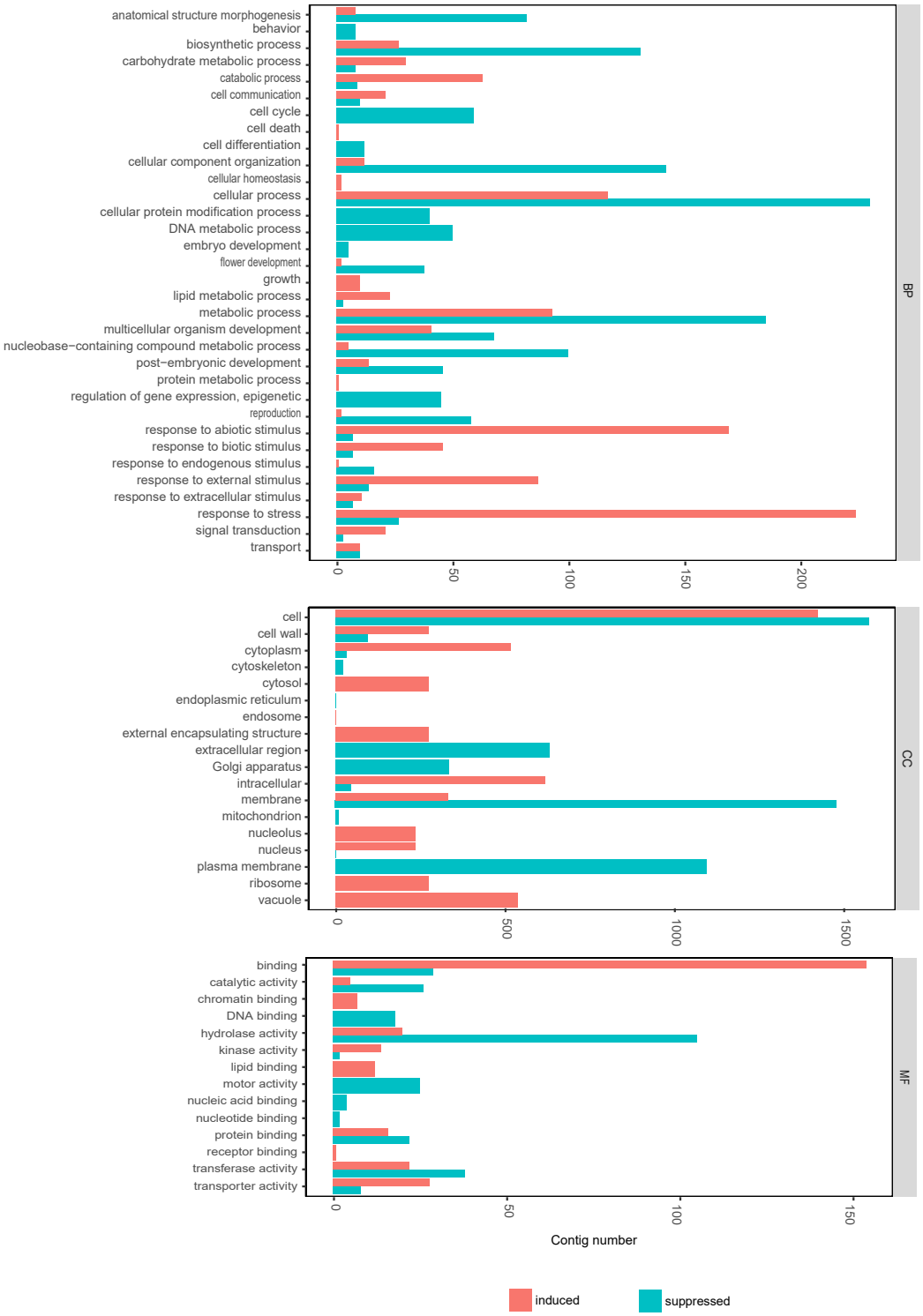

anatomical structure morphogenesis
behavior
biosynthetic process
carbohydrate metabolic process
catabolic process
cell communication
cell cycle
cell death
cell differentiation
cellular component organization
cellular homeostasis
cellular process
cellular protein modification process
DNA metabolic process
embryo development
flower development
growth
lipid metabolic process
metabolic process
multicellular organism development
nucleobase−containing compound metabolic process
post−embryonic development
protein metabolic process
regulation of gene expression, epigenetic
reproduction
response to abiotic stimulus
response to biotic stimulus
response to endogenous stimulus
response to external stimulus
response to extracellular stimulus
response to stress
signal transduction
transport
BP
0
50
100
150
200
cell
cell wall
cytoplasm
cytoskeleton
cytosol
endoplasmic reticulum
endosome
external encapsulating structure
extracellular region
Golgi apparatus
intracellular
membrane
mitochondrion
nucleolus
nucleus
plasma membrane
ribosome
vacuole
CC
0
500
1000
1500
binding
catalytic activity
chromatin binding
DNA binding
hydrolase activity
kinase activity
lipid binding
motor activity
nucleic acid binding
nucleotide binding
protein binding
receptor binding
transferase activity
transporter activity
MF
0
50
100
150
Contig number
induced
suppressed

Figure 4.1. GO enrichment summary of contigs suppressed and induced in resistance response. GO enrichment was performed on the genes that were suppressed (green) and induced (red) in response to host resistance, respectively. GO categories were summarized in terms of Molecular Function(MF), Biological Process (BP) and Cellular Component (CC) and number of genes from significantly enriched terms (FDR < 0.01) were presented in bar graph.

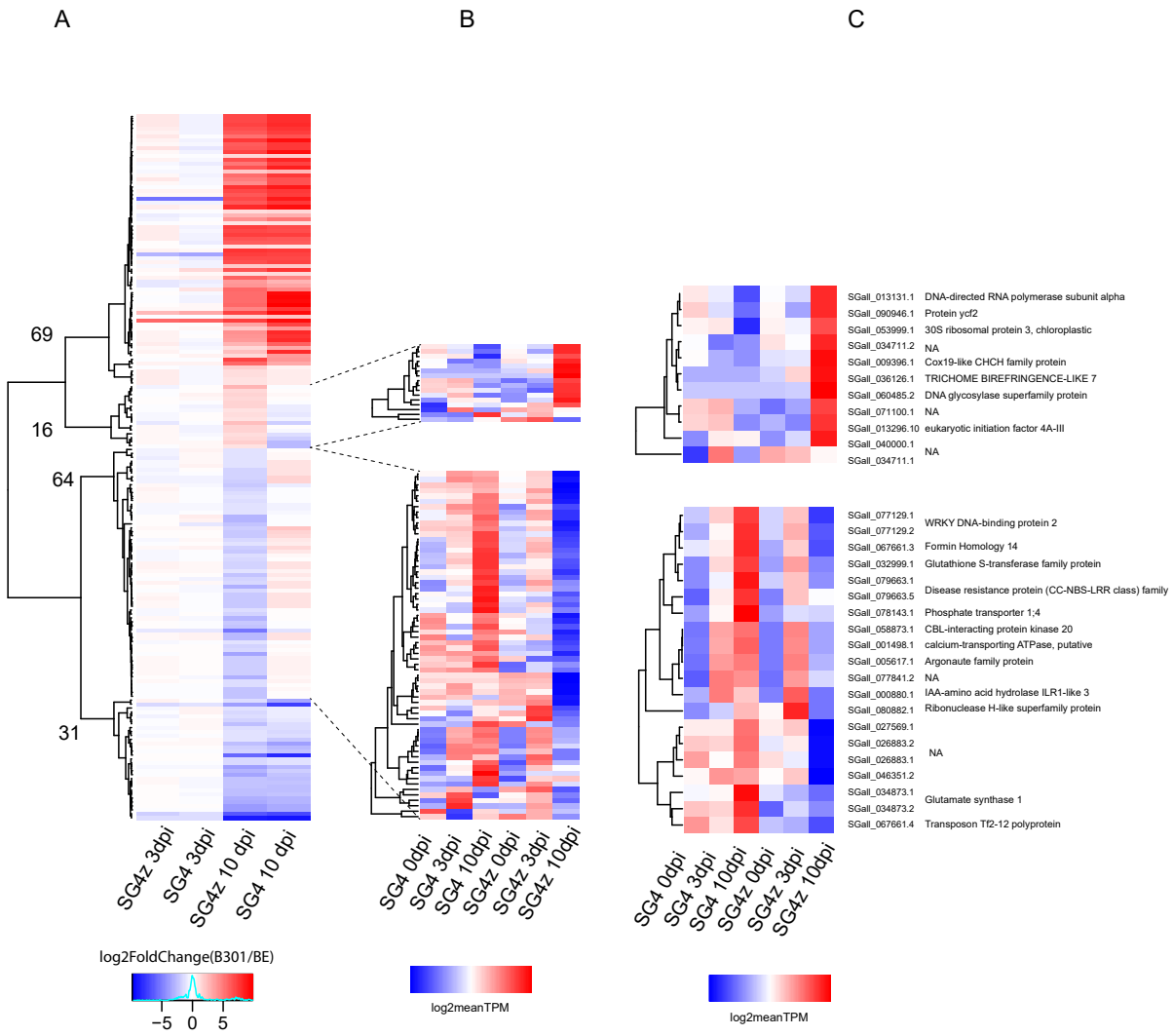

A
B
C
69
16
64
31
SG4z 3dpi
SG4 3dpi
SG4z 10 dpi
SG4 10 dpi
log2FoldChange(B301/BE)
−5 0 5
SG4 0dpi
SG4 3dpi
SG4 10dpi
SG4z 0dpi
SG4z 3dpi
SG4z 10dpi
log2meanTPM
SGall_013131.1 DNA-directed RNA polymerase subunit alpha
SGall_090946.1 Protein ycf2
SGall_053999.1 30S ribosomal protein 3, chloroplastic
SGall_034711.2 NA
SGall_009396.1 Cox19-like CHCH family protein
SGall_036126.1 TRICHOME BIREFRINGENCE-LIKE 7
SGall_060485.2 DNA glycosylase superfamily protein
SGall_071100.1 NA
SGall_013296.10 eukaryotic initiation factor 4A-III
SGall_040000.1 NA
SGall_034711.1
SGall_077129.1 WRKY DNA-binding protein 2
SGall_077129.2
SGall_067661.3 Formin Homology 14
SGall_032999.1 Glutathione S-transferase family protein
SGall_079663.1 Disease resistance protein (CC-NBS-LRR class) family
SGall_079663.5
SGall_078143.1 Phosphate transporter 1;4
SGall_058873.1 CBL-interacting protein kinase 20
SGall_001498.1 calcium-transporting ATPase, putative
SGall_005617.1 Argonaute family protein
SGall_077841.2 NA
SGall_000880.1 IAA-amino acid hydrolase ILR1-like 3
SGall_080882.1 Ribonuclease H-like superfamily protein
SGall_027569.1
SGall_026883.2 NA
SGall_026883.1
SGall_046351.2
SGall_034873.1 Glutamate synthase 1
SGall_034873.2
SGall_067661.4 Transposon Tf2-12 polyprotein
SG4 0dpi
SG4 3dpi
SG4 10dpi
SG4z 0dpi
SG4z 3dpi
SG4z 10dpi
log2meanTPM

Figure 4.2. Effect of host genotypes on gene expression in SG4z and SG4. (A) Host effect on the expression changes in SG4 and SG4z at attachment stages. Host effect is represented by log2 fold change of expression during the interactions with B301 compared to BE (log2FoldChange(B301/BE)). The induced expression in B301, no expression changes at different hosts, the suppressed expression in B301 are indicated by heatmap color red, white and blue, respectively. The host effect of 180 genes that are differentially expressed in SG4z during the interaction with different hosts (B301 vs. BE) are clustered and presented by heatmap. The number labeled on the branches of each cluster indicates the gene number in that group. **(B)** The expression profiles of contigs that are differentially affected by host genotypes in SG4z and SG4. The scaled expression level (the mean of log2TPM) is represented by color intensity from low (red) to high (blue) in heatmap. The up panel shows expression profile of contigs that are induced in SG4z but not in SG4 during the interaction with B301 at 10 dpi. The bottom panel shows the expression profile of contigs that are suppressed in SG4z but not in SG4 during the interaction with B301 at 10 dpi. **(C)** The expression profiles of effector candidates that are differentially affected by host genotypes. *In silico* effector identification pipeline (Figure 2.3A) was applied to the genes from each group in (B). The scaled expression level (the mean of log2TPM) was represented by color intensity from low (red) to high (blue) in heatmap. The up panel shows expression profile of effector candidates that are induced in SG4z but not in SG4 during the interaction with B301 at 10 dpi. The down panel shows the expression profile of effector candidates that are suppressed in SG4z but not in SG4 during the interaction with B301 at 10 dpi.

General Discussion

The advancement in next generation sequencing technologies provides great opportunities to leverage genome-scale tools to gain a better understanding on the molecular biology of parasitic plants. In addition, years of intense investigation on other plant-pathogen system have uncovered many molecular components of the interactions and provide valuable reference information for the study of host-parasitic biology. In this chapter, I integrate my findings in light of previous research and interpret dynamic aspects of *Striga*-host interaction at molecular and genomic level to give a broad view about future in parasitic plant research.

5.1 Global identification of effector genes in the parasitic plants

Parasitic plants, similar to all plant parasites, directly interact with their host plants. To complete their parasitic life cycle, they must overcome sophisticated defense mechanisms of host plants and exhibit virulence for their survival at host-pathogen interactions. Based on previously developed two-tier plant immunity model, it was hypothesized that parasitic plants, like all other plant parasites, secrete an arsenal of effector molecules to suppress plant innate defense system and allow parasite to continue proliferating at susceptible hosts.

Pioneering research used genome-wide approaches to identify effectors in the three related species of parasitic plants within Orobanchaceae, during multiple stages of parasite growth (Yang et al. 2015). Over one hundred of unigenes were identified as core parasitism genes and upregulated in parasitic process (Yang et al. 2015). All those core parasitism genes were specifically enriched at extracellular region, suggesting that secretion of proteins is important for parasitic plants during the interaction with their hosts.

In fact, effector proteins are a subset of secretory proteins that alter host defense mechanisms to either facilitate or sabotage pathogen growth dependent on susceptibility of infected hosts. Computational prediction pipelines have been developed to identify fungal secreted effector in hundreds of fungal and oomycete species (Sonah et al. 2016). The defined criteria of candidate secreted effector proteins in fungi are proteins with a signal peptide for secretion, no transmembrane domains, no homology to other proteins, fairly small size and likely species-specific (Lum & Min 2011; Choi et al. 2010; Sonah et al. 2016). A similar set of criteria has been used to identify putative nematode effectors (Danchin et al. 2010). In addition to secretome

identification criteria analogous to fungal one, researchers focuses on the nematode effectors that specifically expressed in esophageal glands and at specific stages of parasitism to narrow down the list of putative nematode effectors (Maier et al. 2013).

Based on those criteria developed in these eukaryotic plant parasites (see chapter 2), I developed a computational prediction workflow adapted for global identification of effector candidates in parasitic plants using transcriptomic data of *S. gesnerioides*. Similar to the identification of parasitism genes described by Yang et al (2015), I started by identifying contigs that were upregulated during the host-interaction stage and restricted our analysis to genes with stage specific expression in parasitic plants. We then applied secretome prediction criteria, which include SignalP for the presence of signal peptide, TMHMM for identifying membrane proteins, and Phobius PS-Scan for identifying endoplasmic reticulum (ER) retained proteins. A total of 891 distinct contigs were identified as secreted effector candidates from all of the possible parasite-race:host genotype interactions (Appendix Table A3, SG3-B301 (449), SG3-BE (547), SG4-B301 (273), SG4-BE (334), SG4z-B301(280) and SG4z-BE (332)). More than one third of those candidates (37.1%) were functionally unknown proteins with no homologs to other plant proteins. This is consistent with observation that effector protein share little sequence homology and species-specific (Huang et al. 2014). Among 559 annotated effector candidates, 418 shared high sequence similarity with other parasitic plants in Orobanchaceae (namely *S. hermonthica, S. asiatica, Triphysaria versicolor* and *Phelipanche aegyptiaca*) and clustered to the same orthogroups. Interestingly, the majority of these orthologs found in other parasitic plants also followed similar expression patterns, implying conserved effector functions among Orobanchaceae family. In addition to those 891 secreted effector candidates, I included the genes that were upregulated at host interactions but lack complete sequences for secretome prediction to increase our identification sensitivity, as many of our contigs were incomplete and lacked the full length sequence. The less strict pipeline predicted 8,431 effector candidates from all of the possible parasite-race:host genotype interactions (Appendix Table A3). Although the list of 8,431 effector candidates contains more false positives due to less specific prediction criteria, it provides an extensive database for functional enrichment study and race-specific virulence study.

Global identification of effector genes enables us to examine functional annotation of those genes from a broad perspective. It has been found that gene families encoding cell wall modifying enzymes (e.g. xylanase, 1,4-ß galactosidase, glycosyl hydrolases, and pectin methylesterase and peroxidase enzymes) are among predicted effector candidates through our study and were also

identified by Yang et al. (2015). Consistent with their annotation, GO enrichment analysis of the cellular component revealed that cell wall and extracellular localization annotation terms were significantly enriched among these genes. These categorization suggests that the effector candidates encoded by these genes tend to be secreted and impact cell wall integrity of the host similar to the virulent effectors in other eukaryotic plant parasites (Bellincampi et al. 2014). In addition to cell wall modifying enzymes, genes annotated as proteases also account for a large proportion of effector candidates identified in our RNA-seq study. This is analogous to the bacterial protease effectors that are required for virulence and/or avirulence in plant-bacteria interaction (Axtell et al. 2003; Shao et al. 2002; López-Solanilla et al. 2004). Interestingly, analysis of selective constraints showed that protease effectors in Orobanchaceae have a greater dN/dS as compared with their orthologous genes in nonparasitic plants (Yang et al. 2015). Some particular sites within functional domain are positively selected (dN/dS > 1), suggesting that adaptation of these protease-coding genes are associated with the evolution of parasitism (Yang et al. 2015). Furthermore, several NBS-LRR genes were upregulated in the haustorial region at host interaction stage and identified at all parasitic plants studied (Yang et al. 2016; Yang et al. 2015). This discovery is unique to parasitic plants since no R protein-like effector has been previously identified in other plant parasites. Although the function of these expressed NBS-LRR genes is unclear, phylogenetic detection suggests that those NBS-LRR genes may be acquired by horizontal gene transfer (HGT) from host plants and associated with particular defense response to those parasites or host resistance (Yang et al. 2016). Finally, among the candidates that were annotated with unknown function, some of them contains sequences of mobile elements such as retrotransposons, implying that transposition may also generate new effectors in parasitic plants.

Therefore, like other secreted effector proteins in eukaryotic plant parasites, the global identification of predicted effector candidates not only serves as an excellent starting point to screen virulent and Avr factors for functional analysis, but also helps to understand the evolution and distribution of effectors among plant parasites.

5.2 Functional characterization of effector candidates in parasitic plants

Similar to all plant parasites, the effectors secreted by parasitic plants may function in apoplast or symplast of plant to alter host immunity. Depending on the ultimate effects on pathogen growth, effectors act in one of two perspectives, either the virulence or avirulence. Virulence factors function to suppress the host immunity and lead to increase the fitness of parasite while

avirulence results from a group of Avr factors that trigger ETI in plants containing corresponding *R* genes.

Most effector genes are functionally defined due to their virulence effect. When delivered to the host, they manipulate host immunity by targeting certain signaling components in defense pathway and facilitate the growth of parasites. For example, the bacterial type III effector AvrB from *P. syringae pv. glycinea* directly interacts with a cochaperone for HSP90 RAR1 and consequently promoting downstream MPK4 kinase activation for RIN4 phosphorylation (Cui et al. 2010). The phosphorylation of RIN4 disrupts hormone signaling and induces plant susceptibility in host genotypes lacking R protein RPM1 (Mackey et al. 2002). Similarly, the fungal example *C. fulvum* Avr2 directly interacts with and inhibits tomato proteases RCR3 and PIP1 and expression of Avr2 in tobacco causes enhanced susceptibility toward multiple strains and species of fungi (van Esse et al. 2008).

Heterogeneous expression of virulence effectors in host plants is usually the standard molecular method to test effector virulence and detect corresponding interacting host proteins, following the identification of effector candidates. Using this approach, I cloned a *Striga* effector candidate that contains LRR domain (SHR4z) and transformed it to host cowpea roots using *ex vitro* composite plant system. The expression of SHR4z significantly enhanced the host susceptibility which was assessed by phenotype of interactions between various races of *Striga* and resistant host landrace B301. In addition, I showed that SHR4z directly targets a host protein E3 ligase POB1 which was identified as a positive regulator of host defense in cowpea.

Besides genetically manipulating host plants by delivering virulent effector into host cells, biochemically inhibiting the effector activity is another method to study effector virulence in parasitic plants. Bleischwitz et al. (2010) discovered a cysteine protease (custutain) from *Cuscuta reflexa* through a comparative microarray approach. A inhibitor propeptide solution on *Cuscuta*-tobacco interaction significantly reduced viability of *Cuscuta* but did not influence on tobacco development (Bleischwitz et al. 2010). Compared to genetic manipulation, the biochemical inhibition method requires additional information of effector activities and discovery of appropriate inhibitors, thus it might not be feasible for effectors with unknown annotation and no homology to known proteins. However, it provides a convenient substitute to test functionally known effectors which target the hosts that are difficult to genetically manipulate.

However, not all the effectors only elicit virulence that leads to increased host susceptibility. Due to natural selection pressure on host plants, some effectors (Avr) are detected by host immunity system and trigger intense host defending process known as ETI. ETI usually happens in the hosts containing R protein, which, in many circumstances, is in gene-for-gene association with Avr. For example, AvrB from *P. syringae pv. glycinea* interacts with 'guardee' protein RIN4 and triggers R gene *RPM*-mediated ETI (Mackey et al. 2002). Avr2 from *C. fulvum* inhibits tomato proteases RCR3 and is surveilled by R gene *Cf-2* to trigger HR (Dixon et al. 1996). In the hosts of parasitic plants, we identified a NB-LRR R protein RSG3-B301 in cowpea which are specifically responsible for resistance to *S. gesnerioides* race SG3 (Li & Timko 2009). Transient expressing RSG3-B301 in susceptible cowpea plants triggered visible HR when challenging the host plants with SG3. Thus, based on gene-for-gene hypothesis, we hypothesized that corresponding SG3-specific Avr(s) may directly interact with or be indirectly surveilled by RSG3-B301 during *Striga*-cowpea interaction. In this thesis, I performed a comparative transcriptomic analysis on three races of *S. gesnerioides* and identified 71 effector candidates that were specifically induced in SG3 during interaction with B301. Co-expressing these effectors with RSG3-B301 in a transient expression system is anticipated to characterize avirulence effect of those effectors.

Besides the virulence and avirulence function of effectors, the site of action is also important to characterize their molecular mechanism. In oomycetes and fungi, typical effectors have clear N-terminal signal peptide to allow their translocation from pathogen cells to the apoplastic region. They function in either the apoplast or move into the host cells by over-representing host-targeting motifs, like RXLR, LFLAK, and CHXC amino acid sequences, at C-terminal to signal peptide. In this study, I found that the 25 amino acid secretion signal targeted the SHR4z to the apoplast of the haustorial cell. Retention in the cytoplasm of the host cell is required for its virulence function. However, the molecular mechanism responsible for transferring SHR4z from the parasite apoplast to host cytoplasm and what molecules or host-targeting motifs are involved in this translocation remains to be determined in parasitic plants.

Observing the translocation of cytoplasmic effectors has usually been achieved by genetically transforming parasites with effectors fused to fluorescent proteins. The movement can be visualized using immunofluorescence light microscopy and/or immune-localization electron microscopy. In filamentous fungi, translocation of fluorescent effectors into the host cytoplasm is routinely observed using a variety of effector proteins fused to various versions of green or red fluorescent proteins (Kemen et al. 2005; Khang et al. 2010; Giraldo & Valent 2013). In parasitic

plants, a transformation system has not been established for *Striga* that allows the transgenic plants to retain its parasitic capacity. To overcome this challenge, we attempted to express SHR4z in another root parasitic plant, *Tryphisaria versicolor,* of which transgenic roots remained competent to form haustoria and infect hosts like *Arabidopsis* and *Medicago truncatula* (Tomilov et al. 2007). When full length SHR4z tagged with mcherry was expressed in roots in *Tryphisaria*, we observed the red fluorescence in the apoplast of the parasite, suggesting that the N-terminal signal peptide targets the protein to the extracellular region as expected. However, when we infected *Arabidopsis* roots with transgenic *Tryphisaria* roots, the fluorescence was too weak to be clearly differentiated from host auto-fluorescence. It is difficult to assess whether SHR4z was translocated from apoplast to the host cell. In fact, the weak signal from translocated cytoplasmic effector fluorescence has been a consistent problem for filamentous fungal study as well. It may be caused by diffusion in the host cytoplasm. In the future perhaps brighter fluorophores or the addition of nuclear localization signals may help to overcome this challenge (Giraldo & Valent 2013). Characterizing the translocation of effectors will contribute to elucidating mechanism of specialized structure haustorium that facilitates effector movement in parasitic plants.

In addition to using host-targeting motifs of parasitic effectors or specialized structures to assist translocation, movement of small proteins may also occur by diffusion through plasmodesmata connections established after parasite attachment. Since parasitic plants share most cellular features in common with their hosts, unlike pathogenic microbes and fungi, movement of small proteins may not require specialized structures or protein transfer routes. The physical connection between hosts and parasitic plants may enable effectors to translocate simply by hijacking the route that small protein in host plants usually take. Thus, the structural basis for protein trafficking between parasitic plant and host is clearly an open area for future studies.

5.3 RNAs involved in parasitism

Besides translocation of effector proteins, the movement of RNAs has been described between parasitic plants and hosts. A large scale of mRNAs were moved bi-directionally between the parasitic plant *Cuscuta pentagona* (dodder) and its host and contribute to a large proportion of transcriptome in both *Arabidopsis* (45%) and dodder (24%) (Kim & Westwood 2015). The mobile RNAs tend to be abundantly expressed and significantly enriched in GO annotation like hydrolase activity and response to stimulus compared to non-mobile transcripts. Those enriched functions are similar to the functions of our effector candidates. Thus, it is possible that mobile

mRNAs are trafficking from *Striga* to cowpea and translated into effector proteins at host cells to facilitate parasitism.

Beyond the role of coding for protein, RNAs are also responsible for gene expression regulation. Long hairpin RNAs are processed into small RNAs, which silence mRNAs *in trans* by binding to complementary transcripts. Engineered host-induced gene silencing (HIGS) has proven to be effective countermeasure against fungi (Arias et al. 2015; Cheng et al. 2015; Nowara et al. 2010), nematodes (Huang et al. 2006), insects (Mao et al. 2007) and the parasitic plants (Alakonya et al. 2012; Tomilov et al. 2008; Bandaranayake & Yoder 2013; Runo 2011). The apparent ease of HIGS implies that host plants might naturally exchange small RNAs with their parasites. In fact, the necrotrophic fungal pathogen *Botrytis cinerea* sRNAs have been evidenced to hijack host RNA interference (RNAi) pathway and specifically silence host immunity genes like MAPK signaling proteins (MPK2 and MPK1), oxidative stress-related protein (PRXIIF) and cell wall-associated kinase (WAK) (Weiberg et al. 2013). Conversely, cotton plants induce and export two conserved microRNAs (miRNA), miR166 and miR159, in response to infection by fungal pathogen, *Verticillium dahliae*, and specifically silence two essential fungal virulence genes Ca^{2+}-dependent cysteine protease (Clp-1) and an isotrichodermin C-15 hydroxylase (HiC-15) (Zhang et al. 2016). Therefore, we hypothesize that sRNA trafficking can also be a parasitism or defense mechanism occurring between parasitic plants and their hosts.

Recent work on parasitic plant *Cuscuta campestris* revealed accumulation of miRNAs during the parasitism. Some of those miRNAs have clear targets in host *Arabidopsis*. Hosts with mutations in two of the targets (AFB3 and SEOR1) exhibit significantly increased growth of *C. campestris*, suggesting that miRNAs from parasitic plants may act as virulence factors in suppressing host immunity (Shahid et al. 2017). Similarly, we have recently performed small RNA-sequencing on three races of *S. gesnerioides* at the developmental stages corresponding to sequencing design in this thesis. Preliminary data analysis reveals 9 mobile miRNAs (8 novel miRNAs and 1 conserved miRNAs miR319) detected in the cowpea samples but predicted to be *Striga* miRNAs from the parasite genome sequence. Most of these miRNAs were abundantly expressed at host interaction stages and miR319, particularly, was specially induced in the hypervirulent race SG4z (refer to personal conversation). miR319 has been known to target the jasmonic acid biosynthesis transcriptional factor TCP (Schommer et al. 2008), which is involved in host immune defense and a common target of many pathogenic effectors (Weßling et al. 2014). Thus, those naturally

occurring trans-species miRNAs may be employed by parasitic plants as strategy of parasitism in plant-to-plant interactions.

In addition to immediate involvement of RNAs in parasitism, the massive movement of RNA also provides a likely route of mRNA-mediated horizontal gene transfer (HGT) and may contribute to the adaptive lifestyle of parasitic plants. A HGT event was reported in *Striga*, in which a nuclear monocot gene from the monocot host sorghum integrated into the genome of the eudicot parasite (Yoshida et al. 2010). It was proposed that the retro-processed RNA might mediate the HGT since the gene lacks intron and presents poly-A tail after the HGT. The integration of host genes in parasite genome provides the potential mechanism of generating host mimicry to hijack or adapt to host immune system. In fact, one example has been shown that plant-parasitic cyst nematode possesses two plant specific signaling molecules CLAVATA3/ESR (CLEs) which play a critical role in establishing and maintaining this feeding site by regulating cell division and differentiation (Wang et al. 2011). Similarly, we also found that the dicot parasite *S. hermonthica* contains another small signaling molecule, C-TERMINALLY ENCODED PEPTIDE (CEP), with a monocot-like domain. This *Striga* CEP phylogenetically clusters with CEP family from monocot species and lies next to the CEP of its monocot host sorghum. Although no genome sequence is available to study potential mRNA-mediated HGT events on the genes coding *Striga* CEP, it is likely that host CEP may be the target peptide mimicked by parasitic plant to promote parasitism.

In conclusion, the symplastic connection between parasitic plants and their hosts allows large scale of protein and RNA trafficking. Studying those molecules from parasitic plants will benefit from not only deeper understanding of haustorial function and parasitism evolution in plants, but also the potential of developing new control strategies for these agricultural pests. Current advancement in sequencing technology has brought great progress and opportunities in understanding adaptive lifestyles of parasitic plants at the molecular level. This molecular information allows us to ask several intriguing questions: which effector/Avr triggers the race-specific resistance? What is the molecular mechanism underlying resistance to *Striga*? What structures or molecules facilitate the movement of effectors? Do RNA and protein share the same trafficking route? What are the functions of RNA trafficking in plant parasitism? Answering these questions will greatly improve our understanding of parasitism mechanism in plants and provide a unified view of plant-plant interaction.

Reference

den Ackerveken, G., Marois, E. & Bonas, U., 1996. Recognition of the bacterial avirulence protein AvrBs3 occurs inside the host plant cell. *Cell*, 87(7), pp.1307–1316.

Aggarwal, R. et al., 2014. Avirulence effector discovery in a plant galling and plant parasitic arthropod, the Hessian fly (Mayetiola destructor). *PLoS One*, 9(6), p.e100958.

Akiyama, K. & Hayashi, H., 2006. Strigolactones: chemical signals for fungal symbionts and parasitic weeds in plant roots. *Ann. Bot.*, 97(6), pp.925–931.

Alakonya, A. et al., 2012. Interspecific RNA interference of SHOOT MERISTEMLESS-like disrupts Cuscuta pentagona plant parasitism. *Plant Cell*, 24(7), pp.3153–3166.

Albrecht, C. et al., 2005. The Arabidopsis thaliana SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASES1 and 2 control male sporogenesis. *Plant Cell*, 17(12), pp.3337–3349.

Albrecht, H., Yoder, J.I. & Phillips, D.A., 1999. Flavonoids promote haustoria formation in the root parasite triphysaria versicolor. *Plant Physiol.*, 119(2), pp.585–592.

Allen, R.L. et al., 2004. Host-parasite coevolutionary conflict between Arabidopsis and downy mildew. *Science*, 306(5703), pp.1957–1960.

Amusan, I.O. et al., 2008. Resistance to Striga hermonthica in a maize inbred line derived from Zea diploperennis. *New Phytol.*, 178(1), pp.157–166.

Arias, R.S., Dang, P.M. & Sobolev, V.S., 2015. RNAi-mediated Control of Aflatoxins in Peanut: Method to Analyze Mycotoxin Production and Transgene Expression in the Peanut/Aspergillus Pathosystem. *J. Vis. Exp.*, (106), p.e53398.

Arsovski, A.A. et al., 2009. AtBXL1 encodes a bifunctional beta-D-xylosidase/alpha-L-arabinofuranosidase required for pectic arabinan modification in Arabidopsis mucilage secretory cells. *Plant Physiol.*, 150(3), pp.1219–1234.

Asai, T. et al., 2002. MAP kinase signalling cascade in Arabidopsis innate immunity. *Nature*, 415(6875), pp.977–983.

Atokple, I.D.K., Singh, B.B. & Emechebe, A.M., 1995. Genetics of Resistance to Striga and Alectra in Cowpea. *J. Hered.*, 86(1), pp.45–49.

Axtell, M.J. et al., 2003. Genetic and molecular evidence that the Pseudomonas syringae type III effector protein AvrRpt2 is a cysteine protease. *Mol. Microbiol.*, 49(6), pp.1537–1546.

Axtell, M.J. et al., 2001. Mutational analysis of the Arabidopsis RPS2 disease resistance gene and the corresponding pseudomonas syringae avrRpt2 avirulence gene. *Mol. Plant. Microbe. Interact.*, 14(2), pp.181–188.

Azevedo, C., Santos-Rosa, M.J. & Shirasu, K., 2001. The U-box protein family in plants. *Trends*

Plant Sci., 6(8), pp.354–358.

Bandaranayake, P.C.G. & Yoder, J.I., 2013. Trans-specific gene silencing of acetyl-CoA carboxylase in a root-parasitic plant. *Mol. Plant. Microbe. Interact.*, 26(5), pp.575–584.

Barrett, L.G. et al., 2009. Diversity and evolution of effector loci in natural populations of the plant pathogen Melampsora lini. *Mol. Biol. Evol.*, 26(11), pp.2499–2513.

Bellincampi, D., Cervone, F. & Lionetti, V., 2014. Plant cell wall dynamics and wall-related susceptibility in plant-pathogen interactions. *Front. Plant Sci.*, 5, p.228.

Bennetzen, J.L. et al., 2000. The study and engineering of resistance to the parasitic weed Striga in rice, sorghum and maize. *Breeding for striga resistance in cereals. Ibadan, Nigeria: Margraf Verlag*, pp.197–205.

Bent, A.F. & Mackey, D., 2007. Elicitors, effectors, and R genes: the new paradigm and a lifetime supply of questions. *Annu. Rev. Phytopathol.*, 45, pp.399–436.

Bent, A.F. & Mackey, D., 2007. Elicitors, Effectors, and *R* Genes: The New Paradigm and a Lifetime Supply of Questions. *Annual Review of Phytopathology*, 45(1), pp.399–436. Available at: http://www.annualreviews.org/doi/10.1146/annurev.phyto.45.062806.094427.

Bethke, G. et al., 2009. Flg22 regulates the release of an ethylene response factor substrate from MAP kinase 6 in Arabidopsis thaliana via ethylene signaling. *Proc. Natl. Acad. Sci. U. S. A.*, 106(19), pp.8067–8072.

Biasini, M. et al., 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.*, 42(W1), pp.W252--W258.

Bleischwitz, M. et al., 2010. Significance of Cuscutain, a cysteine protease from Cuscuta reflexa, in host-parasite interactions. *BMC Plant Biol.*, 10, p.227.

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

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. *Proc. Natl. Acad. Sci. U. S. A.*, 107(21), pp.9909–9914.

Botanga, C. & Timko, M., 2006. Phenetic relationships among different races of Striga gesnerioides (Willd.) Vatke from West Africa. *Genome*, 1365(2006), pp.1351–1365. Available at: http://www.nrcresearchpress.com/doi/abs/10.1139/g06-086 [Accessed May 29, 2014].

Botanga, C.J. & Timko, M.P., 2005. Genetic Structure and Analysis of Host and Nonhost Interactions of Striga gesnerioides (Witchweed) from Central Florida. *Phytopathology*,

95(10), pp.1166–1173.

Bouwmeester, H.J. et al., 2003. Secondary metabolite signalling in host–parasitic plant interactions. *Curr. Opin. Plant Biol.*, 6(4), pp.358–364.

Bouwmeester, K. et al., 2011. The lectin receptor kinase LecRK-I.9 is a novel Phytophthora resistance component and a potential host target for a RXLR effector. *PLoS Pathog.*, 7(3), p.e1001327.

Bozkurt, T.O. et al., 2011. Phytophthora infestans effector AVRblb2 prevents secretion of a plant immune protease at the haustorial interface. *Proc. Natl. Acad. Sci. U. S. A.*, 108(51), pp.20832–20837.

Bungard, R.A., 2004. Photosynthetic evolution in parasitic plants: insight from the chloroplast genome. *Bioessays*, 26(3), pp.235–247.

van den Burg, H.A. et al., 2006. Cladosporium fulvum Avr4 protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection. *Mol. Plant. Microbe. Interact.*, 19(12), pp.1420–1430.

Bushnell, W.R., 1972. Physiology of Fungal Haustoria. *Annu. Rev. Phytopathol.*, 10(1), pp.151–176.

Campbell, E.J. et al., 2003. Pathogen-responsive expression of a putative ATP-binding cassette transporter gene conferring resistance to the diterpenoid sclareol is regulated by multiple defense signaling pathways in Arabidopsis. *Plant Physiol.*, 133(3), pp.1272–1284.

Capella-Gutiérrez, S., Silla-Martínez, J.M. & Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25(15), pp.1972–1973.

Chen, X. et al., 2014. An XA21-associated kinase (OsSERK2) regulates immunity mediated by the XA21 and XA3 immune receptors. *Mol. Plant*, 7(5), pp.874–892.

Cheng, W. et al., 2015. Host-induced gene silencing of an essential chitin synthase gene confers durable resistance to Fusarium head blight and seedling blight in wheat. *Plant Biotechnol. J.*, 13(9), pp.1335–1345.

Chinchilla, D. et al., 2007. A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature*, 448(7152), pp.497–500.

Choi, J. et al., 2010. Fungal secretome database: integrated platform for annotation of fungal secretomes. *BMC Genomics*, 11, p.105.

Colcombet, J. et al., 2005. Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR KINASES1 and 2 are essential for tapetum development and microspore maturation. *Plant Cell*, 17(12), pp.3350–3361.

Coll, N.S. et al., 2010. Arabidopsis type I metacaspases control cell death. *Science*, 330(6009), pp.1393–1397.

Couto, D. & Zipfel, C., 2016. Regulation of pattern recognition receptor signalling in plants. *Nat. Rev. Immunol.*, 16(9), pp.537–552.

Cui, H. et al., 2010. Pseudomonas syringae effector protein AvrB perturbs Arabidopsis hormone signaling by activating MAP kinase 4. *Cell Host Microbe*, 7(2), pp.164–175.

Cui, L. et al., 2012. Analysis of polymorphism and transcription of the effector gene Avr1b in Phytophthora sojae isolates from China virulent to Rps1b. *Mol. Plant Pathol.*, 13(2), pp.114–122.

Danchin, E.G.J. et al., 2010. Multiple lateral gene transfers and duplications have promoted plant parasitism ability in nematodes. *Proceedings of the National Academy of Sciences*, 107(41), pp.17651–17656.

Dangl, J.L., Horvath, D.M. & Staskawicz, B.J., 2013. Pivoting the plant immune system from dissection to deployment. *Science (New York, N.Y.)*, 341(6147), pp.746–51. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3869199&tool=pmcentrez&rendertype=abstract [Accessed May 25, 2014].

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

Day, B., Dahlbeck, D. & Staskawicz, B.J., 2006. NDR1 interaction with RIN4 mediates the differential activation of multiple disease resistance pathways in Arabidopsis. *Plant Cell*, 18(10), pp.2782–2791.

Der, J.P. et al., 2011. De novo characterization of the gametophyte transcriptome in bracken fern, Pteridium aquilinum. *BMC Genomics*, 12, p.99.

DeYong, B.J. & Innes, R.W., 2006. Plant NBS-LRR proteins in pathogen sensing and host defense. *Nat Immunol.*, 7(12), pp.1243–1249.

Ding, X. et al., 2009. A rice kinase-protein interaction map. *Plant Physiol.*, 149(3), pp.1478–1492.

Dixon, M.S. et al., 1996. The tomato Cf-2 disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. *Cell*, 84(3), pp.451–459.

Djamei, A. et al., 2007. Trojan horse strategy in Agrobacterium transformation: abusing MAPK defense signaling. *Science*, 318(5849), pp.453–456.

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.8888–93.

Available at: http://www.pnas.org/content/103/23/8888.

Dodds, P.N. et al., 2004. The Melampsora lini AvrL567 avirulence genes are expressed in haustoria and their products are recognized inside plant cells. *Plant Cell*, 16(3), pp.755–768.

Duplan, V. & Rivas, S., 2014. E3 ubiquitin-ligases and their target proteins during the regulation of plant innate immunity. *Front. Plant Sci.*, 5, p.42.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, 32(5), pp.1792–1797.

Ejeta, G., 2007. Breeding for Striga Resistance in Sorghum: Exploitation of an Intricate Host–Parasite Biology. *Crop Sci.*, 47, p.S--216--S--227.

Erbs, G. et al., 2008. Peptidoglycan and muropeptides from pathogens Agrobacterium and Xanthomonas elicit plant innate immunity: structure and activity. *Chem. Biol.*, 15(5), pp.438–448.

van Esse, H.P. et al., 2008. The Cladosporium fulvum Virulence Protein Avr2 Inhibits Host Proteases Required for Basal Defense. *the Plant Cell Online*, 20(7), pp.1948–1963. Available at: http://www.plantcell.org/cgi/doi/10.1105/tpc.108.059394.

Evangelisti, E. et al., 2013. The Phytophthora parasitica RXLR effector penetration-specific effector 1 favours Arabidopsis thaliana infection by interfering with auxin physiology. *New Phytol.*, 199(2), pp.476–489.

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.

Fatokun, C.A. et al., 2000. Challenges and opportunities for enhancing sustainable cowpea production. In *Proceedings of the World Cowpea Conference III held at the International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria*. pp. 4–8.

Flor, H.H., 1971. Current Status of the Gene-For-Gene Concept. *Annu. Rev. Phytopathol.*, 9(1), pp.275–296.

Friesen, T.L. et al., 2008. Host-specific toxins: effectors of necrotrophic pathogenicity. *Cell. Microbiol.*, 10(7), pp.1421–1428.

Fukushige, H., 1991. Propagation of Bursaphelenchus xylophilus(Nematoda:Aphelenchoididae)on Fungi Growing in Pine-Shoot Segments. *Appl. Entomol. Zool.*, 26(3), pp.371–376.

Fulton, T.M. et al., 2002. Identification, analysis, and utilization of conserved ortholog set markers for comparative genomics in higher plants. *Plant Cell*, 14(7), pp.1457–1467.

Gao, B. et al., 2001. Identification of putative parasitism genes expressed in the esophageal gland

cells of the soybean cyst nematode Heterodera glycines. *Mol. Plant. Microbe. Interact.*, 14(10), pp.1247–1254.

Gerber, K. & Giblin-Davis, R.M., 1990. Association of the Red Ring Nematode and Other Nematode Species with the Palm Weevil, Rhynchophorus palmarum. *J. Nematol.*, 22(2), pp.143–149.

Gimenez-Ibanez, S., Ntoukakis, V. & Rathjen, J.P., 2009. The LysM receptor kinase CERK1 mediates bacterial perception in Arabidopsis. *Plant Signal. Behav.*, 4(6), pp.539–541.

Gingerich, D.J. et al., 2007. Large-Scale, Lineage-Specific Expansion of a Bric-a-Brac/Tramtrack/Broad Complex Ubiquitin-Ligase Gene Family in Rice. *Plant Cell*, 19(8), pp.2329–2348.

Giraldo, M.C. & Valent, B., 2013. Filamentous plant pathogen effectors in action. *Nature reviews. Microbiology*, 11(11), pp.800–14. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24129511 [Accessed May 23, 2014].

Głowacki, S., Macioszek, V.K. & Kononowicz, A.K., 2011. R proteins as fundamentals of plant innate immunity. *Cellular & molecular biology letters*, 16(1), pp.1–24.

Goritschnig, S. et al., 2012. Computational prediction and molecular characterization of an oomycete effector and the cognate arabidopsis resistance gene. *PLoS Genetics*, 8(2).

Goujon, T. et al., 2003. AtBXL1 , a novel higher plant (Arabidopsis thaliana) putative beta-xylosidase gene , is involved in secondary cell wall metabolism and plant development. *The Plant Journal*, 33, pp.677–690.

Gowda, B.S., Riopel, J.L. & Timko, M.P., 1999. NRSA-1: a resistance gene homolog expressed in roots of non-host plants following parasitism by Striga asiatica (witchweed). *The Plant journal : for cell and molecular biology*, 20(2), pp.217–30. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10571881.

Grant, M.R. et al., 1995. Structure of the Arabidopsis RPM1 gene enabling dual specificity disease resistance. *Science*, 269(5225), pp.843–846.

Gurney, A.L. et al., 2006. A novel form of resistance in rice to the angiosperm parasite Striga hermonthica. *New Phytol.*, 169(1), pp.199–208.

Gurney, A.L. et al., 2003. Novel sources of resistance to *Striga hermonthica* in *Tripsacum dactyloides*, a wild relative of maize. *New Phytologist*, 160(3), pp.557–568.

Gust, A.A. & Felix, G., 2014. Receptor like proteins associate with SOBIR1-type of adaptors to form bimolecular receptor kinases. *Curr. Opin. Plant Biol.*, 21, pp.104–111.

Haas, B.J. et al., 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature protocols*, 8(8), pp.1494–512.

Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3875132&tool=pmcentrez&rendertype=abstract [Accessed May 23, 2014].

Hacquard, S. et al., 2012. A comprehensive analysis of genes encoding small secreted proteins identifies candidate effectors in Melampsora larici-populina (poplar leaf rust). *Molecular plant-microbe interactions : MPMI*, 25(3), pp.279–93. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22046958.

Haussmann, B.I.G. et al., 2004. Genomic regions influencing resistance to the parasitic weed Striga hermonthica in two recombinant inbred populations of sorghum. *Theor. Appl. Genet.*, 109(5), pp.1005–1016.

Haussmann, B.I.G. et al., 2001. Major and Minor Genes for Stimulation of Striga hermonthica Seed Germination in Sorghum, and Interaction with Different Striga Populations. *Crop Sci.*, 41, pp.1507–1512.

ten Have, A. et al., 2002. The Contribution of Cell Wall Degrading Enzymes to Pathogenesis of Fungal Plant Pathogens. In *Agricultural Applications*. The Mycota. Springer, Berlin, Heidelberg, pp. 341–358.

He, K. et al., 2007. BAK1 and BKK1 regulate brassinosteroid-dependent growth and brassinosteroid-independent cell-death pathways. *Curr. Biol.*, 17(13), pp.1109–1115.

Heese, A. et al., 2007. The receptor-like kinase SERK3/BAK1 is a central regulator of innate immunity in plants. *Proc. Natl. Acad. Sci. U. S. A.*, 104(29), pp.12217–12222.

Hewezi, T. et al., 2010. Arabidopsis spermidine synthase is targeted by an effector protein of the cyst nematode Heterodera schachtii. *Plant physiology*, 152(2), pp.968–84. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19965964%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2815906%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/19965964%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2815906.

Hewezi, T. et al., 2008. Cellulose binding protein from the parasitic nematode Heterodera schachtii interacts with Arabidopsis pectin methylesterase: cooperative cell wall modification during parasitism. *The Plant cell*, 20(11), pp.3080–3093. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19001564%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2613657.

Hewezi, T. & Baum, T.J., 2013. Manipulation of plant cells by cyst and root-knot nematode effectors. *Molecular plant-microbe interactions : MPMI*, 26(1), pp.9–16. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22809272.

Hofmann, K. et al., 1999. The PROSITE database, its status in 1999. *Nucleic Acids Res.*, 27(1),

pp.215–219.

van der Hoorn, R.A.L. & Kamoun, S., 2008. From Guard to Decoy: a new model for perception of plant pathogen effectors. *Plant Cell*, 20(8), pp.2009–2017.

Hu, G. et al., 2005. EDS1 in tomato is required for resistance mediated by TIR-class R genes and the receptor-like R gene Ve. *Plant Journal*, 42(3), pp.376–391.

Hu, H., Xiong, L. & Yang, Y., 2005. Rice SERK1 gene positively regulates somatic embryogenesis of cultured cell and host defense response against fungal infection. *Planta*, 222(1), pp.107–117.

Huang, G. et al., 2006. Engineering broad root-knot resistance in transgenic plants by RNAi silencing of a conserved and essential root-knot nematode parasitism gene. *Proc. Natl. Acad. Sci. U. S. A.*, 103(39), pp.14302–14306.

Huang, J. et al., 2014. Rapid evolution of avirulence genes in rice blast fungus Magnaporthe oryzae. *BMC Genet.*, 15(1), p.45.

Huang, K. et al., 2012a. Global changes in gene expression during compatible and incompatible interactions of cowpea (Vigna unguiculata L.) with the root parasitic angiosperm Striga gesnerioides. *BMC Genomics*, 13(1), p.402. Available at: BMC Genomics.

Huang, K. et al., 2012b. Global changes in gene expression during compatible and incompatible interactions of cowpea (Vigna unguiculata L.) with the root parasitic angiosperm Striga gesnerioides. *BMC genomics*, 13(1), p.402. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3505475&tool=pmcentrez&rendertype=abstract [Accessed May 29, 2014].

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. *J. Bacteriol.*, 175(15), pp.4859–4869.

Iseli, C., Jongeneel, C. V & Bucher, P., 1999. ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, pp.138–148.

Jasmer, D.P., Goverse, A. & Smant, G., 2003. Parasitic nematode interactions with mammals and plants. *Annu. Rev. Phytopathol.*, 41, pp.245–270.

Jia, Y. et al., 2000. Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *The EMBO Journal*, 19(15), pp.4004–4014.

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 of the United States of America*, 105(12), pp.4874–4879.

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

Kadota, Y. et al., 2014. Direct regulation of the NADPH oxidase RBOHD by the PRR-associated kinase BIK1 during plant immunity. *Mol. Cell*, 54(1), pp.43–55.

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*, 103(29), pp.11086–11091.

van Kan, J.A.L., 2006. Licensed to kill: the lifestyle of a necrotrophic plant pathogen. *Trends Plant Sci.*, 11(5), pp.247–253.

Karimi, M., Inzé, D. & Depicker, A., 2002. GATEWAY vectors for Agrobacterium-mediated plant transformation. *Trends Plant Sci.*, 7(5), pp.193–195.

Katoh, K. et al., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.*, 30(14), pp.3059–3066.

Kawarazaki, T. et al., 2013. A low temperature-inducible protein AtSRC2 enhances the ROS-producing activity of NADPH oxidase AtRbohF. *Biochim. Biophys. Acta*, 1833(12), pp.2775–2780.

Kay, S. et al., 2007. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science*, 318(5850), pp.648–651.

Kemen, E. et al., 2005. Identification of a protein from rust fungi transferred from haustoria into infected plant cells. *Mol. Plant. Microbe. Interact.*, 18(11), pp.1130–1139.

Kemmerling, B. et al., 2007. The BRI1-associated kinase 1, BAK1, has a brassinolide-independent role in plant cell-death control. *Curr. Biol.*, 17(13), pp.1116–1122.

Khang, C.H. et al., 2010. Translocation of Magnaporthe oryzae effectors into rice cells and their subsequent cell-to-cell movement. *Plant Cell*, 22(4), pp.1388–1403.

Kim, G. et al., 2014. Genomic-scale exchange of mRNA between a parasitic plant and its hosts. *Science*, 345(6198), pp.808–811.

Kim, G. & Westwood, J.H., 2015. Macromolecule exchange in Cuscuta-host plant interactions. *Curr. Opin. Plant Biol.*, 26, pp.20–25.

Kobayashi, M. et al., 2007. Calcium-dependent protein kinases regulate the production of reactive oxygen species by potato NADPH oxidase. *Plant Cell*, 19(3), pp.1065–1080.

Kong, G. et al., 2015. The Activation of Phytophthora Effector Avr3b by Plant Cyclophilin is Required for the Nudix Hydrolase Activity of Avr3b. *PLoS Pathog.*, 11(8), p.e1005139.

Krogh, A. et al., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.*, 305(3), pp.567–580.

Kruijt, M. et al., 2005. The Cf-4 and Cf-9 resistance genes against Cladosporium fulvum are conserved in wild tomato species. *Molecular plant-microbe interactions : MPMI*, 18(9), pp.1011–21. Available at: http://apsjournals.apsnet.org/doi/abs/10.1094/MPMI-18-1011.

Kudla, U. et al., 2007. Structural and functional characterization of a novel, host penetration-related pectate lyase from the potato cyst nematode Globodera rostochiensis. *Mol. Plant Pathol.*, 8(3), pp.293–305.

Kuijt, J., 1977. Haustoria of Phanerogamic Parasites. *Annu. Rev. Phytopathol.*, 15(1), pp.91–118.

Kunze, G. et al., 2004. The N terminus of bacterial elongation factor Tu elicits innate immunity in Arabidopsis plants. *Plant Cell*, 16(12), pp.3496–3507.

Lane, J.A. et al., 1996. Characterization of virulence and geographic distribution of Striga gesnerioides on cowpea in West Africa. *Plant Disease*, 80(3), pp.299–301. Available at: http://www.apsnet.org/publications/plantdisease/backissues/Documents/1996Articles/PlantDisease80n03_299.pdf.

Lane, J.A. et al., 1997. Phenotypic characterisation of resistance in Zea diploperennis to Striga hermonthica. *Maydica*, 42(1), pp.45–51.

Lane, J.A., Bailey, J.A., et al., 1993. Resistance of cowpea [Vigna unguiculata (L.) Walp.] to Striga gesnerioides (Willd.) Vatke, a parasitic angiosperm. *New Phytol.*, 125(2), pp.405–412.

Lane, J.A., Moore, T.H.M., et al., 1993. Virulence characteristics of a new race of the parasitic angiosperm, Striga gesnerioides, from southern Benin on cowpea (Vigna unguiculata). *Euphytica*, 72(3), pp.183–188.

Langmead, B. & Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, 9(4), pp.357–359.

Lannoo, N. & Van Damme, E.J.M., 2014. Lectin domains at the frontiers of plant defense. *Front. Plant Sci.*, 5, p.397.

Li, C.-J. et al., 2013. Cloning and characterization of an NBS-LRR resistance gene from peanuts (Arachis hypogaea L.). *Physiol. Mol. Plant Pathol.*, 84, pp.70–75.

Li, J. et al., 2002. BAK1, an Arabidopsis LRR receptor-like protein kinase, interacts with BRI1 and modulates brassinosteroid signaling. *Cell*, 110(2), pp.213–222.

Li, J., Lis, K.E. & Timko, M.P., 2009. Molecular genetics of race-specific resistance of cowpea to Striga gesnerioides (Willd.). *Pest management science*, 65(5), pp.520–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19222045 [Accessed May 29, 2014].

Li, J. & Timko, M.P., 2009. Gene-for-gene resistance in Striga-cowpea associations. *Science (New York, N.Y.)*, 325(5944), p.1094. Available at:

http://www.ncbi.nlm.nih.gov/pubmed/19713520.

Li, L. et al., 2014. The FLS2-associated kinase BIK1 directly phosphorylates the NADPH oxidase RbohD to control plant immunity. *Cell Host Microbe*, 15(3), pp.329–338.

Li, L., Stoeckert Jr, C.J. & Roos, D.S., 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.*, 13(9), pp.2178–2189.

Li, S., 2010. Characterization of soybean GmPUB1 proteins that interact with the Phytophthora sojae effector Avr1b protein.

Liebrand, T.W.H., van den Burg, H.A. & Joosten, M.H.A.J., 2014. Two for all: receptor-associated kinases SOBIR1 and BAK1. *Trends Plant Sci.*, 19(2), pp.123–132.

Lim, M.T.S. & Kunkel, B.N., 2004. Mutations in the Pseudomonas syringae avrRpt2 gene that dissociate its virulence and avirulence activities lead to decreased efficiency in AvrRpt2-induced disappearance of RIN4. *Mol. Plant. Microbe. Interact.*, 17(3), pp.313–321.

López-Solanilla, E. et al., 2004. HopPtoN is a Pseudomonas syringae Hrp (type III secretion system) cysteine protease effector that suppresses pathogen-induced necrosis associated with both compatible and incompatible plant interactions. *Mol. Microbiol.*, 54(2), pp.353–365.

Love, M., Anders, S. & Huber, W., 2013. Differential analysis of RNA-Seq data at the gene level using the DESeq2 package.

Lozano-Torres, J.L. et al., 2012. Dual disease resistance mediated by the immune receptor Cf-2 in tomato requires a common virulence target of a fungus and a nematode. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 109(25), pp.10119–24. Available at: http://www.pnas.org/content/109/25/10119.abstract.

Lu, D. et al., 2010. A receptor-like cytoplasmic kinase, BIK1, associates with a flagellin receptor complex to initiate plant innate immunity. *Proc. Natl. Acad. Sci. U. S. A.*, 107(1), pp.496–501.

Lu, D. et al., 2011. Direct ubiquitination of pattern recognition receptor FLS2 attenuates plant innate immunity. *Science*, 332(6036), pp.1439–1442.

Lum, G. & Min, X.J., 2011. FunSecKB: the Fungal Secretome KnowledgeBase. *Database*, 2011, p.bar001.

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.

Mackey, D. et al., 2002. RIN4 interacts with Pseudomonas syringae type III effector molecules and is required for RPM1-mediated resistance in Arabidopsis. *Cell*, 108(6), pp.743–754.

Maier, T.R. et al., 2013. Isolation of whole esophageal gland cells from plant-parasitic nematodes

for transcriptome analyses and effector identification. *Mol Plant Microbe Interactions*, 26(1), pp.31–35. Available at: https://www.ncbi.nlm.nih.gov/pubmed/22876962%5Cnhttp://apsjournals.apsnet.org/doi/pdfplus/10.1094/MPMI-05-12-0121-FI.

Maiti, R.K. et al., 1984. A Comparative Study of the Haustorial Development of Striga asiatica (L.) Kuntze on Sorghum Cultivars. *Ann. Bot.*, 54(4), pp.447–457.

Mao, G. et al., 2011. Phosphorylation of a WRKY transcription factor by two pathogen-responsive MAPKs drives phytoalexin biosynthesis in Arabidopsis. *Plant Cell*, 23(4), pp.1639–1653.

Mao, Y.-B. et al., 2007. Silencing a cotton bollworm P450 monooxygenase gene by plant-mediated RNAi impairs larval tolerance of gossypol. *Nat. Biotechnol.*, 25(11), pp.1307–1313.

Mateo, A. et al., 2004. LESION SIMULATING DISEASE 1 is required for acclimation to conditions that promote excess excitation energy. *Plant Physiol.*, 136(1), pp.2818–2830.

Mayer, A.M., 2006. Pathogenesis by fungi and by parasitic plants: Similarities and differences. *Phytoparasitica*, 34(1), pp.3–16.

Mazzucotelli, E. et al., 2006. The e3 ubiquitin ligase gene family in plants: regulation by degradation. *Curr. Genomics*, 7(8), pp.509–522.

Mbuvi, D.A. et al., 2017. Novel Sources of Witchweed (Striga) Resistance from Wild Sorghum Accessions. *Front. Plant Sci.*, 8, p.116.

McHale, L. et al., 2006. Plant NBS-LRR proteins: adaptable guards. *Genome Biol.*, 7(4), p.212.

Mellor, K.E., 2013. *Molecular Basis of Cowpea (Vigna Unguiculata) Resistance to the Root Parasitic Angiosperm, Striga Gesnerioides*. Charlottesville, VA: 2013 CN - Disk 2210-000; Diss.Biol.2013.M45. Available at: https://search.lib.virginia.edu/catalog/u6108837.

Mellor, K.E., Hoffman, A.M. & Timko, M.P., 2012. Use of ex vitro composite plants to study the interaction of cowpea (Vigna unguiculata L.) with the root parasitic angiosperm Striga gesnerioides. *Plant methods*, 8(1), p.22. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3441300&tool=pmcentrez&rendertype=abstract [Accessed May 29, 2014].

Meng, X. & Zhang, S., 2013. MAPK cascades in plant disease resistance signaling. *Annu. Rev. Phytopathol.*, 51, pp.245–266.

Mohamed, A. et al., 2003. Hypersensitive Response to Infection in. *Crop Sci.*, 43(4), pp.1320–1324.

Mohamed, A.H., Housley, T.L. & others, 2010. Inheritance of hyper sensitive response to Striga

parasitism in sorghum [Sorghum bicolor (L.) Moench]. *African Journal of*.

Monaghan, J. & Zipfel, C., 2012. Plant pattern recognition receptor complexes at the plasma membrane. *Current Opinion in Plant Biology*, 15(4), pp.349–357. Available at: http://dx.doi.org/10.1016/j.pbi.2012.05.006.

Montes, R.A.C. et al., 2008. Cell wall modifications in Arabidopsis plants with altered $α$-l-arabinofuranosidase activity. *Plant Physiol.*, 147(1), pp.63–77.

Mudgil, Y. et al., 2004. A large complement of the predicted Arabidopsis ARM repeat proteins are members of the U-box E3 ubiquitin ligase family. *Plant Physiol.*, 134(1), pp.59–66.

Mukhtar, M.S. et al., 2011. Independently evolved virulence effectors converge onto hubs in a plant immune system network. *Science (New York, N.Y.)*, 333(6042), pp.596–601. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3170753&tool=pmcentrez&rendertype=abstract [Accessed May 23, 2014].

Muñoz-Amatria, M. et al., 2017. Genome resources for climate-resilient cowpea, an essential crop for food security. *Plant J.*, 89(5), pp.1042–1054.

Musselman, L.J. & Parker, C., 1981. Studies on Indigo Witchweed, the American Strain of Striga gesnerioides (Scrophulariaceae). *Weed Sci.*, 29(5), pp.594–596.

Mutuku, J.M. et al., 2015. The WRKY45-Dependent Signaling Pathway Is Required For Resistance against Striga hermonthica Parasitism. *Plant physiology*, 168(3), pp.1152–63. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26025049.

Nam, K.H. & Li, J., 2002. BRI1/BAK1, a receptor kinase pair mediating brassinosteroid signaling. *Cell*, 110(2), pp.203–212.

Navarro, L. et al., 2004. The transcriptional innate immune response to flg22. Interplay and overlap with Avr gene-dependent defense responses and bacterial pathogenesis. *Plant Physiol.*, 135(2), pp.1113–1128.

Nimchuk, Z. et al., 2000. Eukaryotic fatty acylation drives plasma membrane targeting and enhances function of several type III effector proteins from Pseudomonas syringae. *Cell*, 101(4), pp.353–363.

Nowara, D. et al., 2010. HIGS: host-induced gene silencing in the obligate biotrophic fungal pathogen Blumeria graminis. *Plant Cell*, 22(9), pp.3130–3141.

Ogasawara, Y. et al., 2008. Synergistic activation of the Arabidopsis NADPH oxidase AtrbohD by Ca2+ and phosphorylation. *J. Biol. Chem.*, 283(14), pp.8885–8892.

Ogmen, U. et al., 2005. PRISM: protein interactions by structural matching. *Nucleic Acids Res.*, 33(Web Server issue), pp.W331--6.

Olivier, A., Benhamou, N. & Leroux, G.D., 1991. Cell surface interactions between sorghum roots and the parasitic weed Striga hermonthica: cytochemical aspects of cellulose distribution in resistant and susceptible host tissues. *Can. J. Bot.*, 69(8), pp.1679–1690.

Omoigui, L.O. et al., 2017. Breeding cowpea for resistance to Striga gesnerioides in the Nigerian dry savannas using marker-assisted selection. *Plant Breed.*, 136(3), pp.393–399.

Orosa, B. et al., 2017. BTB-BACK Domain Protein POB1 Suppresses Immune Cell Death by Targeting Ubiquitin E3 ligase PUB17 for Degradation. *PLOS Genetics*, 13(1), p.e1006540. Available at: http://dx.plos.org/10.1371/journal.pgen.1006540.

Pachter, L., 2011. Models for transcript quantification from RNA-Seq. *arXiv preprint arXiv:1104.3889*, pp.1–28. Available at: http://arxiv.org/abs/1104.3889 [Accessed May 30, 2014].

Park, C.-H. et al., 2012. The Magnaporthe oryzae effector AvrPiz-t targets the RING E3 ubiquitin ligase APIP6 to suppress pathogen-associated molecular pattern-triggered immunity in rice. *Plant Cell*, 24(11), pp.4748–4762.

Park, C.-J. et al., 2008. Rice XB15, a protein phosphatase 2C, negatively regulates cell death and XA21-mediated innate immunity. *PLoS Biol.*, 6(9), p.e231.

Parker, C., 2009. Observations on the current status of Orobanche and Striga problems worldwide. *Pest management science*, 65(5), pp.453–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19206075 [Accessed May 29, 2014].

Parker, C. & Polniaszek, T.I., 1990. Parasitism of cowpea by Striga gesnerioides: variation in virulence and discovery of a new source of host resistance. *Ann. Appl. Biol.*, 116(2), pp.305–311.

Parker, C. & Riches, C.R., 1993. *Parasitic weeds of the world: biology and control*, CAB International.

Peng, H.-C. & Kaloshian, I., 2014. The tomato leucine-rich repeat receptor-like kinases SlSERK3A and SlSERK3B have overlapping functions in bacterial and nematode innate immunity. *PLoS One*, 9(3), p.e93302.

Petersen, T.N. et al., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods*, 8(10), pp.785–786.

Postma, J. et al., 2016. Avr4 promotes Cf-4 receptor-like protein association with the BAK1/SERK3 receptor-like kinase to initiate receptor endocytosis and plant immunity. *New Phytol.*, 210(2), pp.627–642.

Rafiqi, M. et al., 2013. Effector candidates in the secretome of Piriformospora indica, a ubiquitous plant-associated fungus. *Frontiers in plant science*, 4(July), p.228. Available at:

https://apps.webofknowledge.com/full_record.do?product=UA&search_mode=GeneralSearch&qid=33&SID=U1jv6kHejfeRC3XMJXd&page=1&doc=1%5Cnhttps://apps.webofknowledge.com/full_record.do?product=UA&search_mode=GeneralSearch&qid=3&SID=U1jv6kHejfeRC3XMJXd&page=1&doc.

Rairdan, G.J. & Moffett, P., 2006. Distinct domains in the ARC region of the potato resistance protein Rx mediate LRR binding and inhibition of activation. *The Plant cell*, 18(August), pp.2082–2093.

Ranjan, A. et al., 2014. De novo assembly and characterization of the transcriptome of the parasitic weed dodder identifies genes associated with plant parasitism. *Plant Physiol.*, 166(3), pp.1186–1199.

Rehmany, A.P. et al., 2005. Differential recognition of highly divergent downy mildew avirulence gene alleles by RPP1 resistance genes from two Arabidopsis lines. *Plant Cell*, 17(6), pp.1839–1850.

Riopel, J.L. & Timko, M.P., 1995. Haustorial initiation and differentiation. *Parasitic plants*, pp.39–79.

Roberts, A. et al., 2011. Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol.*, 12(3), p.R22.

Rodenburg, J. et al., 2015. Do NERICA rice cultivars express resistance to Striga hermonthica (Del.) Benth. and Striga asiatica (L.) Kuntze under field conditions? *Field Crops Res.*, 170, pp.83–94.

Römer, P. et al., 2007. Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science*, 318(5850), pp.645–648.

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. *Plant Cell*, 23(6), pp.2440–2455.

Runo, S., 2011. Engineering host-derived resistance against plant parasites through RNA interference: challenges and opportunities. *Bioeng. Bugs*, 2(4), pp.208–213.

Rutter, W.B. et al., 2014. Mining novel effector proteins from the esophageal gland cells of Meloidogyne incognita. *Mol. Plant. Microbe. Interact.*, 27(9), pp.965–974.

Sarris, P.F. et al., 2015. A Plant Immune Receptor Detects Pathogen Effectors that Target WRKY Transcription Factors. *Cell*, 161(5), pp.1089–1100.

Schmittgen, T.D. & Livak, K.J., 2008. Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.*, 3(6), pp.1101–1108.

Schommer, C. et al., 2008. Control of jasmonate biosynthesis and senescence by miR319 targets.

PLoS Biol., 6(9), p.e230.

Schornack, S. et al., 2004. The tomato resistance protein Bs4 is a predicted non-nuclear TIR-NB-LRR protein that mediates defense responses to severely truncated derivatives of AvrBs4 and overexpressed AvrBs3. *Plant J.*, 37(1), pp.46–60.

Schulze, B. et al., 2010. Rapid heteromerization and phosphorylation of ligand-activated plant transmembrane receptors and their associated kinase BAK1. *J. Biol. Chem.*, 285(13), pp.9444–9451.

Seybold, H. et al., 2014. Ca2+ signalling in plant immune response: from pattern recognition receptors to Ca2+ decoding mechanisms. *New Phytol.*, 204(4), pp.782–790.

Shabab, M. et al., 2008. Fungal Effector Protein AVR2 Targets Diversifying Defense-Related Cys Proteases of Tomato. *Plant Cell*, 20(4), pp.1169–1183. Available at: http://www.plantcell.org/cgi/content/abstract/20/4/1169.

Shahid, S. et al., 2017. *MicroRNAs from the parasitic plant Cuscuta campestris target host messenger RNAs*,

Shao, F. et al., 2002. A Yersinia effector and a Pseudomonas avirulence protein define a family of cysteine proteases functioning in bacterial pathogenesis. *Cell*, 109(5), pp.575–588.

Shen, Q.-H. et al., 2007. Nuclear activity of MLA immune receptors links isolate-specific and basal disease-resistance responses. *Science*, 315(5815), pp.1098–1103.

Shimizu, T. et al., 2010. Two LysM receptor molecules, CEBiP and OsCERK1, cooperatively regulate chitin elicitor signaling in rice. *Plant J.*, 64(2), pp.204–214.

Singh, B.B. & Emechebe, A.M., 1990. Inheritance of Striga resistance in cowpea genotype B301. *Crop Sci.*, 30.

Sonah, H., Deshmukh, R.K. & Bélanger, R.R., 2016. Computational Prediction of Effector Proteins in Fungi: Opportunities and Challenges. *Front. Plant Sci.*, 7, p.126.

Spallek, T., Mutuku, M. & Shirasu, K., 2013. The genus Striga: a witch profile. *Molecular plant pathology*, 14(9), pp.861–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23841683 [Accessed May 29, 2014].

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22(21), pp.2688–2690.

Staskawicz, B. et al., 1987. Molecular characterization of cloned avirulence genes from race 0 and race 1 of Pseudomonas syringae pv. glycinea. *J. Bacteriol.*, 169(12), pp.5789–5794.

Swarbrick, P.J. et al., 2009. A major QTL for resistance of rice to the parasitic plant Striga hermonthica is not dependent on genetic background. *Pest Manag. Sci.*, 65(5), pp.528–532.

Swarbrick, P.J. et al., 2008. Global patterns of gene expression in rice cultivars undergoing a

susceptible or resistant interaction with the parasitic plant Striga hermonthica. *New Phytol.*, 179(2), pp.515–529.

Szabo, L.J. & Bushnell, W.R., 2001. Hidden robbers: the role of fungal haustoria in parasitism of plants. *Proc. Natl. Acad. Sci. U. S. A.*, 98(14), pp.7654–7655.

Tang, D., Wang, G. & Zhou, J.-M., 2017. Receptor Kinases in Plant-Pathogen Interactions: More Than Pattern Recognition. *Plant Cell*, 29(4), pp.618–637.

Terauchi, R. & Yoshida, K., 2010. Towards population genomics of effector – effector target interactions. *New Phytologist*, 187, pp.929–939.

Tian, M. et al., 2007. A Phytophthora infestans Cystatin-Like Protein Targets a Novel Tomato Papain-Like Apoplastic Protease. *Plant Physiology*, 143(January), pp.364–377.

Timko, M.P., Huang, K. & Lis, K.E., 2012. Host Resistance and Parasite Virulence in Striga–Host Plant Interactions: A Shifting Balance of Power. *Weed Sci.*, 60(2), pp.307–315.

Tomilov, A.A. et al., 2008. Trans-specific gene silencing between host and parasitic plants. *Plant J.*, 56(3), pp.389–397.

Tomilov, A., Tomilova, N. & Yoder, J.I., 2007. Agrobacterium tumefaciens and Agrobacterium rhizogenes transformed roots of the parasitic plant Triphysaria versicolor retain parasitic competence. *Planta*, 225(5), pp.1059–1071.

Torres, M.A., Jones, J.D.G. & Dangl, J.L., 2006. Reactive oxygen species signaling in response to pathogens. *Plant Physiol.*, 141(2), pp.373–378.

Touré, M. et al., 1997. Inheritance of resistance to Striga gesnerioides biotypes from Mali and Niger in cowpea (Vigna unguiculata (L.) Walp.). *Euphytica*, 94(3), pp.273–278.

Trujillo, M. et al., 2008. Negative regulation of PAMP-triggered immunity by an E3 ubiquitin ligase triplet in Arabidopsis. *Curr. Biol.*, 18(18), pp.1396–1401.

Tsuda, K. et al., 2009. Network properties of robust immunity in plants. *PLoS Genet.*, 5(12), p.e1000772.

Tsuda, K. & Katagiri, F., 2010. Comparing signaling mechanisms engaged in pattern-triggered and effector-triggered immunity. *Curr. Opin. Plant Biol.*, 13(4), pp.459–465.

Vogler, R.K., Ejeta, G. & Butler, L.G., 1996. Inheritance of Low Production of Striga Germination Stimulant in Sorghum. *Crop Sci.*, 36, pp.1185–1191.

Wall, P.K. et al., 2008. PlantTribes: a gene and gene family resource for comparative genomics in plants. *Nucleic Acids Res.*, 36(Database issue), pp.D970--6.

Wan, J. et al., 2008. A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in Arabidopsis. *Plant Cell*, 20(2), pp.471–481.

Wang, J. et al., 2010. Dual roles for the variable domain in protein trafficking and host-specific

recognition of Heterodera glycines CLE effector proteins. *The New phytologist*, 187(4), pp.1003–17. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20497349.

Wang, J. et al., 2011. Identification of potential host plant mimics of CLAVATA3 / ESR (CLE) -like peptides from the plant-parasitic nematode Heterodera schachtii. *Molecular plant pathology*, 12, pp.177–186.

Wang, Y. et al., 2010. A *Pseudomonas syringae* ADP-Ribosyltransferase Inhibits *Arabidopsis* Mitogen-Activated Protein Kinase Kinases. *The Plant Cell*, 22(6), pp.2033–2044. Available at: http://www.plantcell.org/lookup/doi/10.1105/tpc.110.075697.

Wang, Y.-S. et al., 2006. Rice XA21 binding protein 3 is a ubiquitin ligase required for full Xa21-mediated disease resistance. *Plant Cell*, 18(12), pp.3635–3646.

Weiberg, A. et al., 2013. Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science (New York, N.Y.)*, 342(6154), pp.118–23. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24092744 [Accessed May 24, 2014].

Weßling, R. et al., 2014. Convergent targeting of a common host protein-network by pathogen effectors from three kingdoms of life. *Cell Host and Microbe*, 16(3), pp.364–375.

Westwood, J. & Yoder, J., 2010. The evolution of parasitism in plants. *Trends in plant ...*, 15(4), pp.227–235. Available at: http://dx.doi.org/10.1016/j.tplants.2010.01.004 [Accessed May 30, 2014].

Westwood, J.H. et al., 2010. The evolution of parasitism in plants. *Trends in plant science*, 15(4), pp.227–35. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20153240 [Accessed May 29, 2014].

Whisson, S.C. et al., 2007. A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature*, 450(7166), pp.115–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17914356 [Accessed May 29, 2014].

Whitehead, A. & Crawford, D.L., 2006. Neutral and adaptive variation in gene expression. *Proc. Natl. Acad. Sci. U. S. A.*, 103(14), pp.5425–5430.

Wickett, N.J. et al., 2011. Transcriptomes of the parasitic plant family Orobanchaceae reveal surprising conservation of chlorophyll synthesis. *Current biology : CB*, 21(24), pp.2098–104. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22169535 [Accessed May 29, 2014].

Williams, J.S. et al., 2014. Transcriptome analysis reveals the same 17 S-locus F-box genes in two haplotypes of the self-incompatibility locus of Petunia inflata. *Plant Cell*, 26(7), pp.2873–2888.

Win, J. et al., 2007. Adaptive evolution has targeted the C-terminal domain of the RXLR

effectors of plant pathogenic oomycetes. *Plant Cell*, 19(8), pp.2349–2369.

Win, J. et al., 2012. Effector biology of plant-associated organisms: concepts and perspectives. *Cold Spring Harb. Symp. Quant. Biol.*, 77, pp.235–247.

Wu, F. et al., 2006. Combining bioinformatics and phylogenetics to identify large sets of single-copy orthologous genes (COSII) for comparative, evolutionary and systematic studies: a test case in the euasterid plant clade. *Genetics*, 174(3), pp.1407–1420.

Wydro, M., Kozubek, E. & Lehmann, P., 2006. Optimization of transient Agrobacterium-mediated gene expression system in leaves of Nicotiana benthamiana. *Acta Biochim. Pol.*, 53(2), pp.289–298.

Xiang, T. et al., 2008. Pseudomonas syringae effector AvrPto blocks innate immunity by targeting receptor kinases. *Curr. Biol.*, 18(1), pp.74–80.

Yang, C.-W. et al., 2006. The E3 ubiquitin ligase activity of arabidopsis PLANT U-BOX17 and its functional tobacco homolog ACRE276 are required for cell death and defense. *Plant Cell*, 18(4), pp.1084–1098.

Yang, S. et al., 2008. Alfalfa benefits from Medicago truncatula: the RCT1 gene from M. truncatula confers broad-spectrum resistance to anthracnose in alfalfa. *Proc. Natl. Acad. Sci. U. S. A.*, 105(34), pp.12164–12169.

Yang, Z. et al., 2015. Comparative Transcriptome Analyses Reveal Core Parasitism Genes and Suggest Gene Duplication and Repurposing as Sources of Structural Novelty. *Molecular Biology and Evolution*, 32(3), pp.767–790. Available at: http://mbe.oxfordjournals.org/cgi/doi/10.1093/molbev/msu343.

Yang, Z. et al., 2016. Horizontal gene transfer is more frequent with increased heterotrophy and contributes to parasite adaptation. *Proc. Natl. Acad. Sci. U. S. A.*

Yoder, J.I. & Scholes, J.D., 2010. Host plant resistance to parasitic weeds; recent progress and bottlenecks. *Current opinion in plant biology*, 13(4), pp.478–84. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20627804 [Accessed May 29, 2014].

Yoshida, S. et al., 2010. Horizontal gene transfer by the parasitic plant Striga hermonthica. *Science*, 201, p.1121. Available at: http://www.sciencemag.org/content/328/5982/1128.short [Accessed May 30, 2014].

Yoshida, S. & Shirasu, K., 2009. Multiple layers of incompatibility to the parasitic witchweed, Striga hermonthica. *New Phytologist*, 183(1), pp.180–189.

Zhang, J. et al., 2007. A Pseudomonas syringae Effector Inactivates MAPKs to Suppress PAMP-Induced Immunity in Plants. *Cell Host and Microbe*, 1(3), pp.175–185.

Zhang, J. et al., 2010. Receptor-like cytoplasmic kinases integrate signaling from multiple plant

immune receptors and are targeted by a Pseudomonas syringae effector. *Cell Host Microbe*, 7(4), pp.290–301.

Zhang, T. et al., 2016. Cotton plants export microRNAs to inhibit virulence gene expression in a fungal pathogen. *Nat Plants*, 2(10), p.16153.

Zhou, J. et al., 2014. The Pseudomonas syringae effector HopF2 suppresses Arabidopsis immunity by targeting BAK1. *Plant J.*, 77(2), pp.235–245.

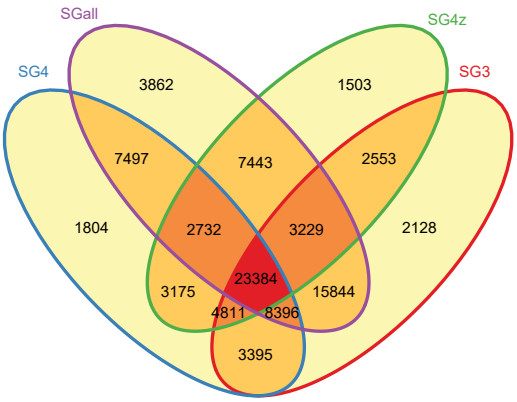
Zhu, Z. et al., 2010. Arabidopsis resistance protein SNC1 activates immune responses through association with a transcriptional corepressor. *Proc. Natl. Acad. Sci. U. S. A.*, 107(31), pp.13960–13965.

Zipfel, C. et al., 2004. Bacterial disease resistance in Arabidopsis through flagellin perception. *Nature*, 428(6984), pp.764–767.

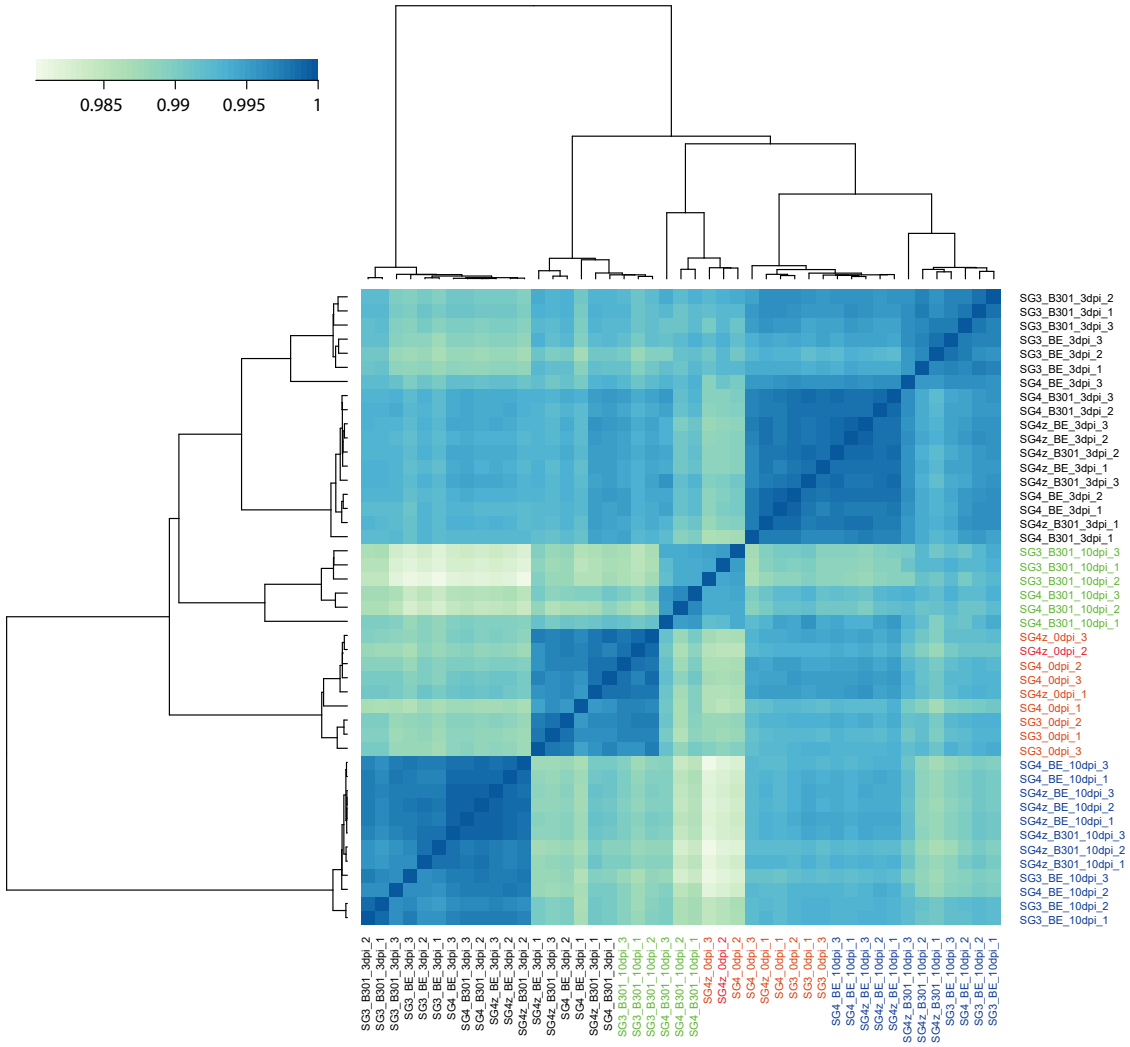
Zipfel, C. et al., 2006. Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts Agrobacterium-mediated transformation. *Cell*, 125(4), pp.749–760.

Zipfel, C. & Rathjen, J.P., 2008. Plant immunity: AvrPto targets the frontline. *Curr. Biol.*, 18(5), pp.R218--20.

Appendix Figures

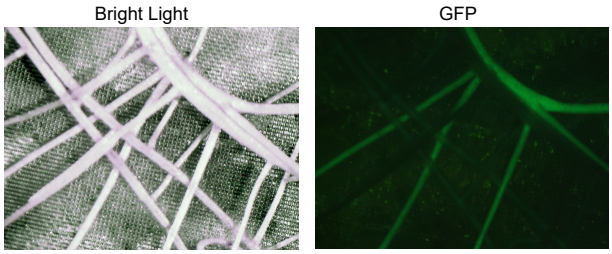


Appendix Figure 1. **The number of orthogroups shared by race-specific assembly and SGall hybrid assembly.** Predicted peptides of all assemblies were combined and subject to orthomcl orthogroup clustering. The number of orthogroups shared by assemblies were listed in venn diagram.

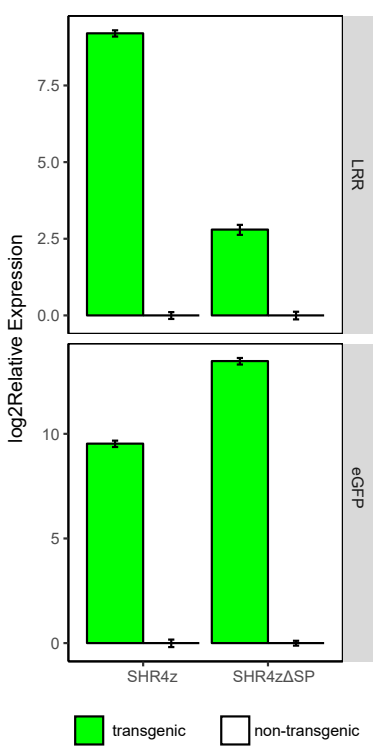

0.985
0.99
0.995
1
SG3_B301_3dpi_2
SG3_B301_3dpi_1
SG3_B301_3dpi_3
SG3_BE_3dpi_3
SG3_BE_3dpi_2
SG3_BE_3dpi_1
SG4_BE_3dpi_3
SG4_B301_3dpi_3
SG4_B301_3dpi_2
SG4z_BE_3dpi_3
SG4z_BE_3dpi_2
SG4z_B301_3dpi_2
SG4z_BE_3dpi_1
SG4z_B301_3dpi_3
SG4_BE_3dpi_2
SG4_BE_3dpi_1
SG4z_B301_3dpi_1
SG4_B301_3dpi_1
SG3_B301_10dpi_3
SG3_B301_10dpi_1
SG3_B301_10dpi_2
SG4_B301_10dpi_3
SG4_B301_10dpi_2
SG4_B301_10dpi_1
SG4z_0dpi_3
SG4z_0dpi_2
SG4_0dpi_2
SG4_0dpi_3
SG4z_0dpi_1
SG4_0dpi_1
SG3_0dpi_2
SG3_0dpi_1
SG3_0dpi_3
SG4_BE_10dpi_3
SG4_BE_10dpi_1
SG4z_BE_10dpi_3
SG4z_BE_10dpi_2
SG4z_BE_10dpi_1
SG4z_B301_10dpi_3
SG4z_B301_10dpi_2
SG4z_B301_10dpi_1
SG3_BE_10dpi_3
SG4_BE_10dpi_2
SG3_BE_10dpi_2
SG3_BE_10dpi_1

Appendix Figure 2. Genome-wide similarity of transcriptional profile of three race at interaction with cowpea during the development. Two cowpea cultivars: B301 and Blackeye (BE); three *Striga* races: SG3, SG4 and SG4z correlation. Three developmental stages: germination (0 dpi), early attachment (3 dpi), late attachment (10 dpi). The heatmap colors represents pairwise correlation of the mean value of regularized log transformation of read count between two conditions. Hierarchical clustering was performed based on pairwise correlation and supported by multiscale bootstrap resampling as described in the methods.

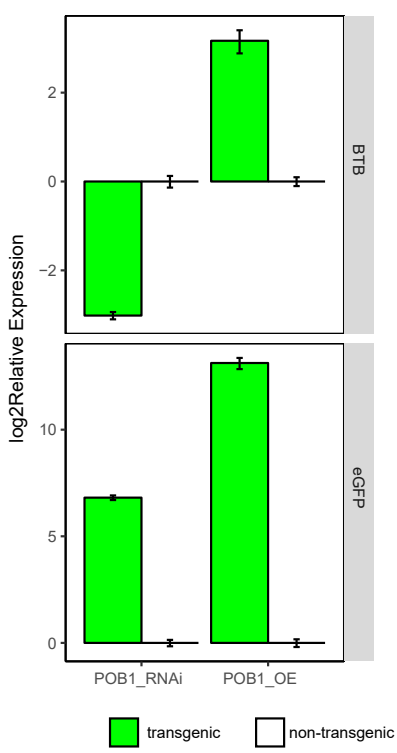
A.


Bright Light
GFP

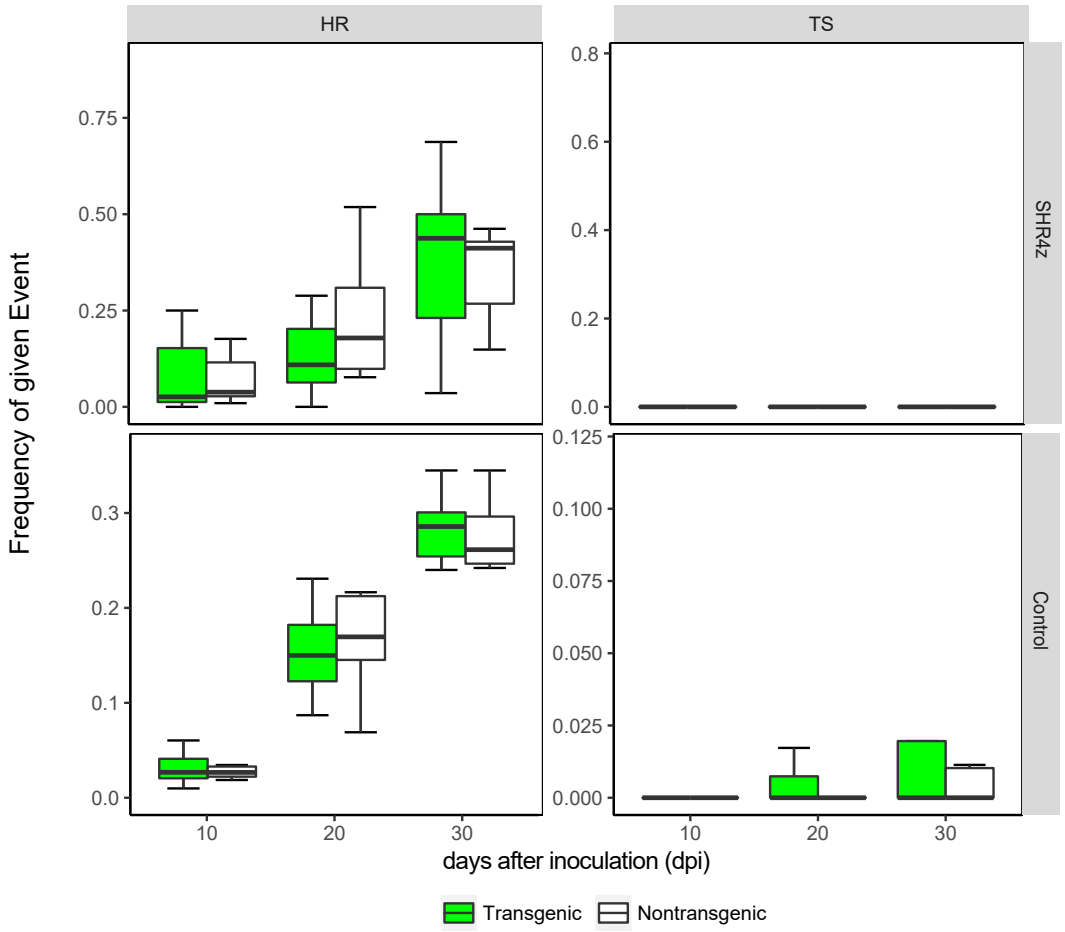
B.


log2Relative Expression
7.5
5.0
2.5
0.0
LRR
10
5
0
eGFP
SHR4z
SHR4zΔSP
transgenic
non-transgenic

C.


log2Relative Expression
2
0
−2
BTB
10
5
0
eGFP
POB1_RNAi
POB1_OE
transgenic
non-transgenic

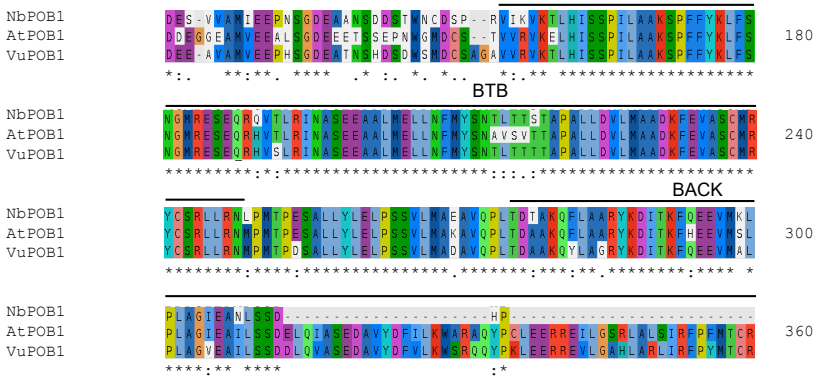
Appendix Figure 3. qPCR to verifying overexpression of LRR and POB1 and silencing POB1 in transgenic roots vs. non-transgenic roots. (A) Representative images of transformed roots and non-transformed roots on composite plants. **(B-C)** qRT-PCR analysis of gene transcript levels in transgenic roots of cowpea composite plants. Total RNA was isolated from transgenic (GFP expressing) and nontransgenic (no visible GFP) roots of composite B301 plants transformed with constructs (**B:** pK7WG2D-SHR4z, pK7WG2D-SHR4z**Δ**SP; **C:** pK7WG2D-POB1 and pK7GWIWG2D-VuPOB1-RNAi). Relative transcript levels were obtained by qRT-PCR by calibrating its threshold cycles relative to control 18S transcripts. Means and standard errors (SE) were given based on 3 replicates.



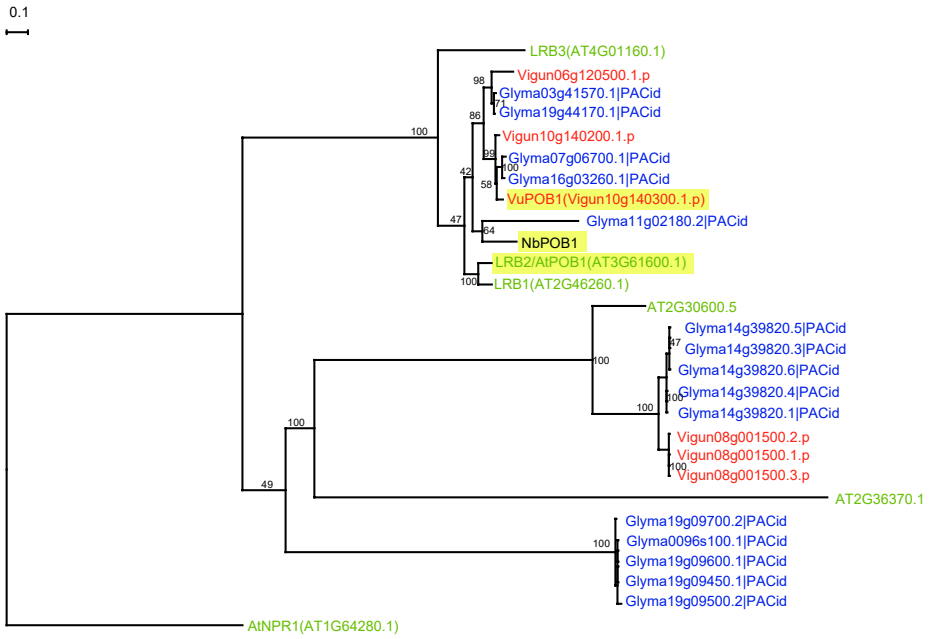
Appendix Figure 4. B301 roots expressing the full length SHR4z and control did not show difference between transgenic and non-transgenic roots. Cowpea composite plants (B301) were generated by transforming full length SHR4z constructs (pK7WG2D-SHR4z) and control

(pK7WG2D-FLAG). Transgenic and non-transgenic roots were inoculated with two-day germinated SG4 seedling, and at 10 dpi, 20 dpi, and 30 dpi the phenotypic responses of the roots scored. The abbreviation of the phenotypic event categories are as follows: HR, hypersentive response; TS, tubercle swelling; and CE, cotyledon expansion. The interaction event ratio for each category was obtained by counting the number of each event category and dividing by the total number of phenotypic events occurring on each host plant.

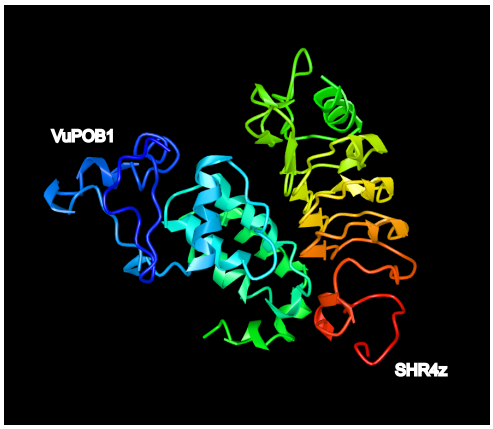
A.


NbPOB1
AtPOB1
VuPOB1
180
BTB
240
BACK
300
360

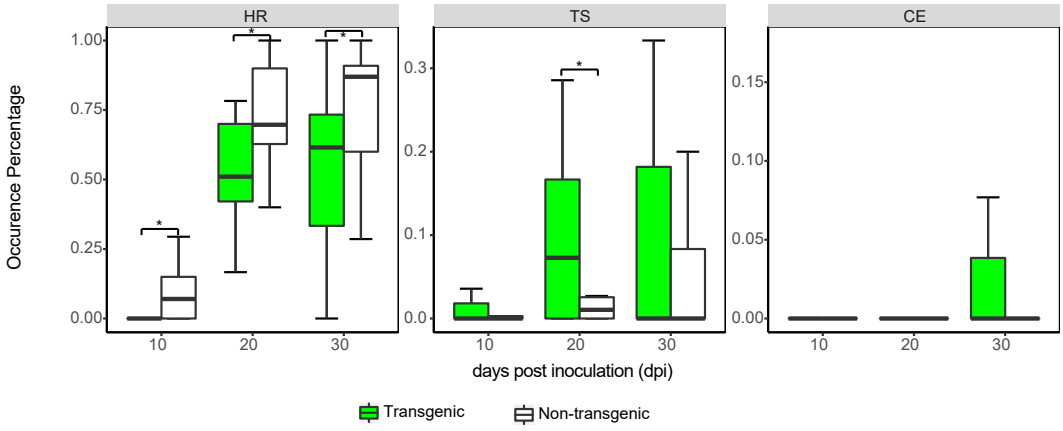
B.


0.1
LRB3(AT4G01160.1)
Vigun06g120500.1.p
Glyma03g41570.1|PACid
Glyma19g44170.1|PACid
Vigun10g140200.1.p
Glyma07g06700.1|PACid
Glyma16g03260.1|PACid
VuPOB1(Vigun10g140300.1.p)
Glyma11g02180.2|PACid
NbPOB1
LRB2/AtPOB1(AT3G61600.1)
LRB1(AT2G46260.1)
AT2G30600.5
Glyma14g39820.5|PACid
Glyma14g39820.3|PACid
Glyma14g39820.6|PACid
Glyma14g39820.4|PACid
Glyma14g39820.1|PACid
Vigun08g001500.2.p
Vigun08g001500.1.p
Vigun08g001500.3.p
AT2G36370.1
Glyma19g09700.2|PACid
Glyma0096s100.1|PACid
Glyma19g09600.1|PACid
Glyma19g09450.1|PACid
Glyma19g09500.2|PACid
AtNPR1(AT1G64280.1)

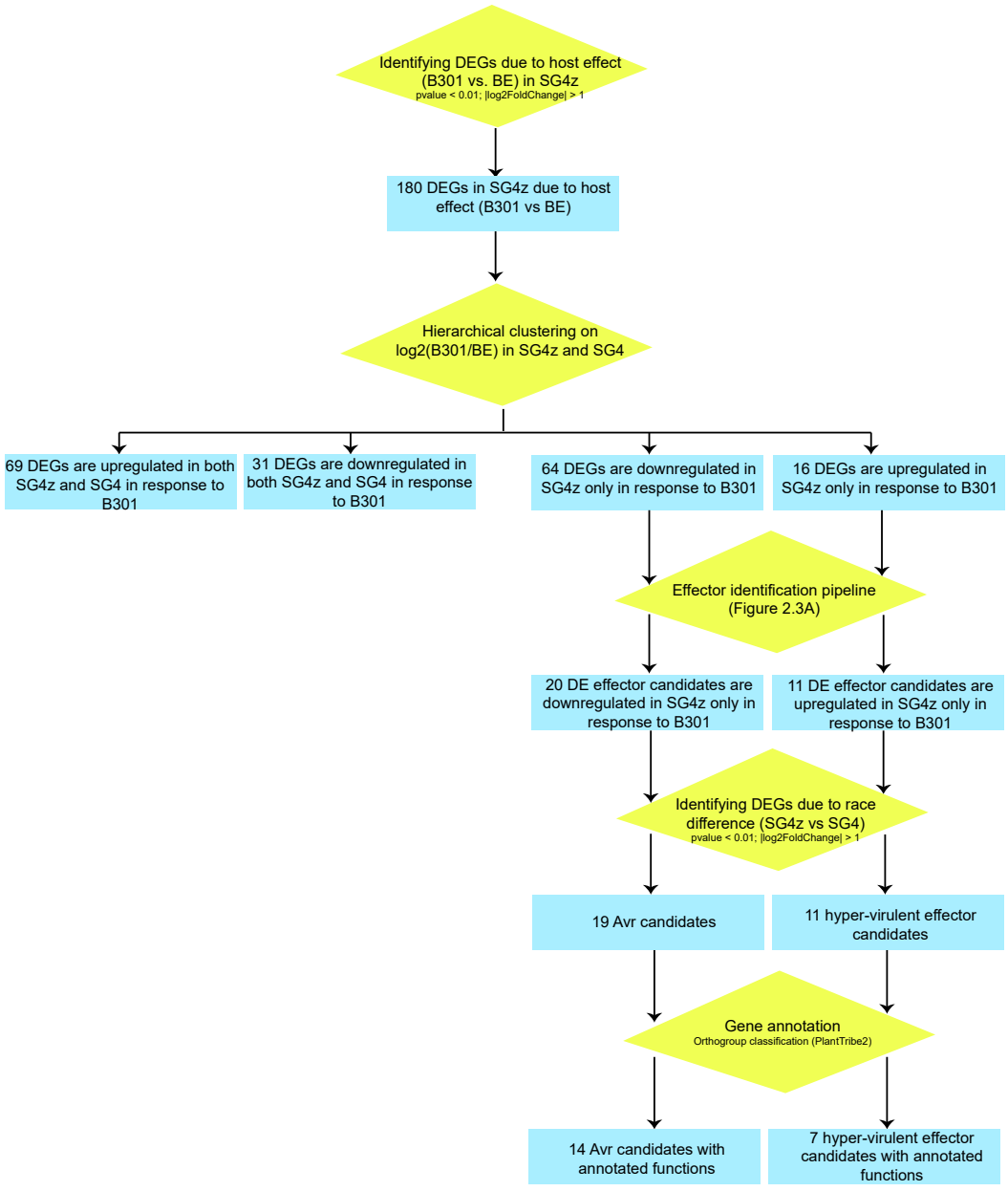
C.


VuPOB1
SHR4z

Appendix Figure 5. Protein Structure of VuPOB1. **(A)** Putative protein sequence alignment of Yeast Two Hybrid screening (Y2H) hit and AtPOB1. The alignment were generated by MUSCLE 3.8 (Edgar 2004). The BTB and BACK domains are labeled with black line. **(B)** Phylogenetic relationship reveals that the BTB-BACK protein is homologous to AtPOB1. The protein sequences containing both the BTB and BACK domains were identified by searching genome of cowpea (Muñoz-Amatria et al. 2017), closest legume soybean (Glycine max Wm82.a2.v1) and Arabidopsis (TAIR10). A maximum likelihood (ML) phylogenetic trees of protein alignments was generated using RAxML with PROTGAMMAWAG model (Stamatakis 2006). The scale bar indicates discrepancy distance. Bootstrap values were shown to evaluate the reliability of the branches on the tree. (C) Predicted interaction surface between SHR4 and VuPOB1. SWISS-MODEL was used to predict macromolecular structures of SHR4z and POB1 based on sequence homology (Biasini et al. 2014). Model 4z64.1.B annotated as SERK1 and Model 4j8z.1.A annotated as Speckle-type POZ protein were predicted as most evolutionarily related structure to SHR4z (QMEAN = -2.34) and VuPOB1 (QMEAN= -2.65), respectively. Macromolecular structures of both proteins were used as input for Protein Interactions By Structural Matching (PRISM (Ogmen et al. 2005)) to predict interaction surface between SHR4z and POB1 and view by Web-based 3D structure viewer iCn3D (https://www.ncbi.nlm.nih.gov/Structure/icn3d/docs/icn3d_about.html).



Appendix Figure 6. Effect of RNAi-silencing of VuPOB1 on SG4z-triggered host responses. *Ex vitro* composite B301 plants were generated by expressing constructs capable of RNAi silencing of VuPOB1 expression. Transgenic and non-transgenic roots were inoculated with two-day germinated SG4z seedling, and at 10 dpi, 20 dpi, and 30 dpi the phenotypic responses of the roots scored. The abbreviation of the phenotypic event categories are as follows: HR, hypersentive response; TS, tubercle swelling; and CE, cotyledon expansion. The interaction event ratio for each category was obtained by counting the number of each event category and dividing by the total number of phenotypic events occurring on each host plant. An asterisk (*) indicates a p value < 0.05.



Appendix Figure 7. The flowchart to identify Avr and effector candidates in SG4 and SG4z using differential expression analysis based on host effect. Differentially expressed genes (DEGs) due to host effect are identified in SG4z. Hierarchical clustering reveals those DEGs are

clustered to 4 groups (Figure 4.2) with 80 DEGs differentially regulated in SG4 and SG4z. Effector identification pipeline (Figure 2.3A) is applied to identify effector candidates that are differentially regulated due to host effect in SG4 and SG4z. Finally differentially expression analysis based on race difference discovered the Avr candidates in SG4 that may trigger HR in B301 and the hyper-virulent effector candidates in SG4z that may suppress B301 resistance.

Appendix Table A1. Summary of RNA-seq coverage for each library

Appendix Table A2. The number of DE contigs among three races of S. gesnerioides at compatible and incompatible interactions with cowpea.

Appendix Table A3. The number of effector candidates among three races of S. gesnerioides at compatible and incompatible interactions with cowpea.

Appendix Table A4. Annotation of effector candidates that are differentially expressed in SG4 and SG4z.

Appendix Table A5. GO enrichment of contigs shared by SG3 and SG4 during interaction with B301.

Appendix Table A6. Annotation of Avirulence factor candidates that are specifically expressed in SG3.

Appendix Table A7. Sequences of primers that are used in this thesis work.

Appendix Table A8. The expression changes of NBS-LRR genes in SG3 and SG4 in response to host resistance

Appendix Table A9. Annotation of contigs that are differentially expression in SG4z in response to host difference.