Transcriptome analysis of the parasitic plant *Phtheirospermum japonicum*

（寄生植物 *Phtheirospermum japonicum* のトランスクリプトーム解析）

ジュリアニ　カリニ　イシダ

Transcriptome analysis of the parasitic plant *Phtheirospermum japonicum*

（寄生植物 *Phtheirospermum japonicum* のトランスクリプトーム解析）

ジュリアニ　カリニ　イシダ

論 文 の 内 容 の 要 旨

環境生物学専攻

平成０３年度博士課程　入学（進学）

ｼﾞｭﾘｱﾆ ｶﾘﾆ ｲｼﾀﾞ

指導教員名：難波成任

寄生植物 *Phtheirospermum japonicum* のトランスクリプトーム解析

(Transcriptome analysis of the parasitic plant *Phtheirospermum japonicum)*

Angiosperms have acquired a parasitic lifestyle at least 12 times during evolution. A common feature of all parasitic plant is the presence of haustorium, a root feeding structure responsible for connecting the parasite to the host. Parasitic plants in Orobanchaceae cause serious agricultural problems in worldwide. *Phtheirospermum japonicum*, a member of Orobanchaceae, is a facultative parasite native in East Asia which infects a broad range of hosts. *P. japonicum* represented a good model for parasitism studies since it is suitable for genetic and reverse-genetic analyses, such as crossing and transformation.

In this study I established a large-scale transcriptome platform of *P. japonicum* to identify the genes involved in haustorium development. First I used next-generation sequencing technologies to *de novo* assemble the transcriptome of *P. japonicum* with enrichment of transcripts expressed in the haustoria tissue. The results were discussed in Chapter 2. In this experiment the whole root transcriptome of *P. japonicum* was newly assembled using two platforms, Roche 454 with longer reads (~300 bp) and Illumina Hi-Seq with short reads (~90 bp). Combined sequencing strategies helped increasing the quality of *de novo* assembly. At the end, 58,137 sequences were assembled, showing an average size of 811 bp. Out of them, 45,323 (78%) had predicted coding proteins and 29,767 (51.2%) blast hits against *Arabidopsis* genome. Evaluation of the coverage rate showed that *P. japonicum* assembly covered all the 357 highly conserved single-copy orthologs found in eukaryotic genomes, showing that the assembly presented a reasonable coverage of *P. japonicum* transcriptome. Gene Ontology analysis revealed that genes in the category of “structural molecules activity” and “ribosome” cell compartment were overrepresented in the haustoria

tissues, suggesting extensive protein synthesis for parasitism. In addition, my results showed that the expression profile of quinone reductase 1 (*QR1*), a gene described as necessary for haustorium development in the facultative parasitic plant *Triphysaria vesicolor*, was unaltered by the contact of host or haustorium-inducing chemical. Instead, the related quinone reductase 2 (*QR2*) was up-regulated. Similar results were found in *S. hermonthica*, where *QR2* expression was higher in haustoria tissues compared with *QR1*, which was more expressed in reproductive structures.

Interestingly, among the transcripts enriched from the parasitic stage, many of them were annotated to encode subtilases (SBT). Thus I decided to further investigate the role of these enzymes in plant parasitism and the results are discussed in Chapter 3. The subtilase family comprehended one of largest protein families in plants. It has been suggested that *SBT* appeared in the genome of land plants through a single event of horizontal gene transfer from a bacterial, followed by rapid duplication events. SBT is composed of a signal peptide, prodomain, catalytic, PA and Fn-III domains. It is translated as pre-pro-protein with a signal peptide at the N-terminal end targeting to the apoplastic region. To be translocated to outside of plant cell, a maturation step, in which the prodomain is cleaved, is required. I monitored the *SBT* expression pattern by qRT-PCR during the interaction of the parasite *P. japonicum* with a host (*O. sativa*) or a nonhost (*L. japonicus*) plant for 1, 2, 3 and 7 days of infection. The expression of five *SBT* genes (*PjSBT2,-4,-7,-8 and -11*) were detected only at 7 days after the interaction with the host, while no expression of these genes was detected in contact with nonhost *L. japonicus*. Morphological studies with of a haustorium stained by *Safranin-O* revealed that the formation of a vascular bridge, which connects xylem vessels between a host and a parasite, occurs after 7 days of interaction. Promoter analysis of a *PjSBT* gene showed that its expression was localized at the interacting site in haustoria with the host at penetrating stage. The parasitism-induced *PjSBTs* are the homologs of *Arabidopsis AtSBT* group 5, including symbiosis-induced *L. japonicus SbtS,* and *Arabidopsis AIR3* (At2g04160) involved with lateral root formation, indicating a possible shared mechanism in haustorium, nodulation and lateral root formation. In addition, the parasitism-induced *ShSBT1*, an *S. hermonthica* gene phylogenetically distant from *PjSBTs* of group 5 suggested that this gene may have had acquired a distinct function during the development of terminal haustoria.

In order to understand early molecular events associated with host perception on the first 48 h of haustorium development. I designed a custom microarray based on the assembled sequences. The gene expression was analyzed in 8 different time points upon the treatment

with DMBQ (2, 6-dimethoxy-p-benzoquinone), a natural compound which induces haustorium *in vitro*. Comparison of gene expression profiles between DMBQ-treated and non-treated roots identified 1577 differentially expressed genes divided into three clusters according to Self-Organizing Maps algorithm. Cluster 1 contains 706 genes negatively modulated along the time course. Cluster 2 has 396 positively regulated genes with a peak of expression before 3 h of treatment, designated as early responsive genes. Cluster 3 contains 475 positively regulated genes with a peak of expression after 3 h of treatment, designated as late responsive genes. Based on their best BLAST-hit annotations, the gene ontology (GO) terms were assigned for each gene, aiming to investigate functional gene populations in each cluster. Blast alignment indicated that large amount of parasite sequences did not have any similar sequence in available database, indicating that haustorium development is a rich material for discovery and further characterization of novel gene function.

Cluster 2 has the GO term “transcription factor activity” overrepresented compared with all the entities in the microarray, indicating that massive transcriptional reprograming occurs in the first 3 h of haustorium development. Based on this analysis I found that members of WRKY family, which are often involved in disease resistance in plants, were up-regulated suggesting that plant immunity system is activated on the first hours of haustorium development. In contrast, Cluster 1 showed higher percentage of sequences assigned as “other molecular functions”, 139 sequences out of 701. This cluster includes downregulated of genes encoding proteins responsible for lignification and strengthening of cell wall. Similarly, Cluster 3 is composed of late responsive up-regulated genes involved in cell-wall-modification, indicating that from 3 h to 48 h of induction of haustorium formation the parasite cells undergo intense restructuring, mirroring the morphological changes observed in the organ.

Transcription of plant hormone-related enzymes such as YUCCA, CKX3 and GA2ox8 was induced during the haustorium development. To further understand the involvement of auxin in plant parasitism I focused on the homologs of *Arabidopsis YUCCA* genes which encode key auxin biosynthesis enzymes. Four YUCCA homologs were identified from the *P. japonicum* transcriptome but only *PjYUC3,* which belongs to Cluster 3, was specifically up-regulated after the haustorium-inducing treatments. *P. japonicum* hairy roots transformed with the *PjYUC3* overexpression construct resulted in a typical auxin-overproducing phenotype, indicating that *PjYUC3* encodes a functional YUCCA enzyme. Promoter analysis showed that *PjYUC3* expression occurs in two different locations, at the root apical meristem tip and at the haustorium initiation site. *PjYUC3* is responsible for accumulation of newly-synthesized auxin

specifically in epidermal and cortical cell of the emerging haustorium. Silencing *PjYUC3* provoked reduction of haustorium numbers, emphasizing its relevance for the haustorium development. These results suggested that the *de novo* auxin biosynthesis by *PjYUC3* at the root tip and/or haustorium initiation region is essential for the haustorium development.

Finally, in the chapter 6 I described the transcriptome analysis that I have done in the tissues of the parasitic plant *P. japonicum* interacting with its host rice. I identified regulated genes in the parasitic as well as in the host tissues after 1 or 7 days of the host-parasite interaction. At the first step I generated Illumina Hi-Seq sequences and *de novo* assembled the *P. japonicum* transcriptome. The transcript levels were determined by mapping those reads against the assembled transcriptome. The rice transcripts were mapped against available rice cDNAs and expression levels were accessed. Finally, statistical analysis identified 2917 parasite and 1155 rice genes regulated during parasitism. These genes were divided into nine clusters according to their expression profiles. Co-expression analysis identified the hub genes which are highly co-expressed with other genes. Among them, a hub gene encoding the transcription factor Cytokinin Response Factor (CRF) was highly up-regulated. CRF is a member of AP2/ERF transcriptional factor family. In both parasitic and host tissues, up-regulation of cell wall-modifying enzymes were observed, inferring intense modification of cell shape during the establishment of plant parasitism. To identify the rice genes similarly regulated during its interaction with *P. japonicum* and *S. hermonthica*, I compared the RNA-Seq data described in this chapter with published microarray data from *S. hermonthica*-interacting rice roots. This analysis revealed that the rice homeobox-leucine zipper transcription factor, Oshox16, is up-regulated after interaction with both parasitic plants. Taken together, I identified host and parasite genes relevant for host-parasite infections, which includes two transcription factor encoding genes.

In summary, I have established the whole genome-scale transcriptome of the model parasitic plant *P. japonicum*. Expression analyses using custom microarray and RNA-seq technologies revealed important genes involved in the haustorium development. Thus this work has enabled to obtain the first insight of molecular mechanism underlying the infection strategy by a parasitic Orobanchaceae plant. Future studies using this data set will help to combat against the other Orobanchaceae plants such as *Striga* and *Orobanche* that cause devastating agriculture losses in the world.

Transcriptome analysis of the parasitic plant *Phtheirospermum japonicum*

by

Juliane Karine Ishida

Thesis submitted for the degree of Doctor of Philosophy.

Department of Agricultural and Environmental Biology,

Graduate School of Agricultural and Life Science,

Faculty of Agriculture,

2014

To Nobuhiro Kawai

ACKNOWLEDGEMENTS

First I would like to extend my gratitude to Professors Ken Shirasu and Shigetou Namba for accepting me in their laboratories, for their warm encouragement and thoughtful guidance. It has been a great privilege and an enriching experience being working with them. I hope that I can in turn pass on the research values that have given to me.

I would like to express my sincere gratitude to Satoko Yoshida for continuous support of my research. Without her patience, motivation and knowledge this thesis would have not been possible.

I need to thank my collaborators Prof. Dr. Masaki Ito from Nagoya University and Prof. Dr. John Yoder from UC Davis for providing me the plasmids used in this thesis. Prof. Claude dePamphilis and Eric Wafula from Penn State University for their help with transcriptome assembly. I am deeply grateful to present and past members of Plant Immunity Research Group to make my stay in Japan such a pleasant time. I would like to thank specially Kohki Yoshimoto, Takashi Yaeno, Musembi Mutuku, Satoko Yoshida and Ken Shirasu for collecting parasite seeds or for taking the pictures of *Phtheirospermum* and *Striga* in the field used in this thesis.

I owe deepest gratitude to Yoko Nagai, Hiromi Arai and Kyoko Sugawara for their indispensable assistance in dealing with bureaucratic matters. I must thank all staffs from International Student Center and Division for Counseling and Support, I am especially grateful for Tae Ito, Neelam Ramaiah and Isao Fujioka.

I also would like to thank the Japanese Government for providing me the financial support through Mombukagakusho (MEXT) scholarship.

Lastly, I would like to thank my family for their unconditional encouragement, love and affection. I am truly thankful to Nobuhiro Kawai, who I dedicated this thesis, for staying on my side every moment that I needed. And the last, but not the least I have to thank God.

LIST OF ABBREVIATIONS

4BD	4-Br-dedranone
ARR	Arabidopsis Response Regulator
AM	Arbuscular Mycorrhiza
CCD	Carotenoid Cleavage Dioxygenase
CRF	Cytokinin Response Factor
DMBQ	2,6-dimethoxy-p-benzoquinone
HGT	Horizontal gene transference
HIF	Haustorium-inducing factor
LRR RLK	Leucine-rich repeat protein kinase family protein
NGS	Next Generation Sequencing
PME	Pectin Methyl Esterase
PPGP	Parasitic Plant Genome Project
PR	Pathogenesis-related protein
QTL	Quantitative trait loci
RNS	Root Nodule Symbiosis
SCL	SCARECROW-like transcription factors
SA	Salicylic Acid
SL	Strigolactone
SSP	Small Secreted Peptide
ROS	Reactive oxygen species
XET	Xyloglucan Endotransglucosylase

TABLE OF CONTENTS

1

CHAPTER

PARASITIC PLANTS

1.1 INTRODUCTION

Parasitism occurs when an organism (parasite) benefits from another individual (host), causing harm to the host without its immediate exclusion. Plants face wide range of parasitism caused by fungi, oomycetes, bacteria, viruses, animals, and parasitic plants. During evolution, some angiosperms lost partially or totally their photosynthetic ability, requiring the association with other plants to ensure their survival in nature. The parasitism in angiosperms evolved independently at least 12 times and successfully established in different plant communities, while the exception occurs in aquatic environment where water is not a limited resource (Joel et al., 2013). As consequence, parasitic angiosperms show high morphological diversity, ranging from tiny flower of *Apodanthes caseariae* (Apodanthaceae) to one of the largest flowers in the world, of *Rafflesia* genus (Rafflesiaceae). The common characteristic organ that connects a parasite to the host is called haustorium (*pl.* haustoria). The appearance of the haustorium varies across the different parasites, which corroborates with the hypothesis of multiple origins of parasitic life style in the eudicotyledonous group. Haustorium is a multi-functional structure. At the beginning of parasitization, haustorium serves as an attachment organ to connect a parasite to the host tissue and, at posterior stage, its vascular cells are developed its inside to form a duct between the host vascular vessels and the conduction system of the parasite to draw nutrients and water from the host. Depending on host's tissue where the haustorium penetrates, parasitic plants can be divided into stem or root parasites. Misodendraceae and Apodanthaceae families have only stem parasites, while Orobanchaceae has only the root type and Santalaceae comprises both stem and root parasites (Joel et al., 2013).

Parasitism in plants offers intriguing biological questions since they have accepted drastic changes in their life styles, from autotrophs to heterotrophs. The degree of host dependency and efficiency of photosynthesis is variable across different species. The

"facultative" parasites maintain their ability of free living and opportunistically parasitize if a host is nearby, while "obligate" parasites cannot complete their life cycle without hosts at least in natural condition. Other classification depends on their photosynthetic activity; some species lost partially (hemiparasites) or completely (holoparasites) their photosynthetic ability. In general, the hemiparasites maintain a reduced photosynthetic competency, requiring additional nutrients supplied by host. The non-photosynthetic holoparasite relies completely on host interaction.

1.2 OROBANCHACEAE

Orobanchaceae contains species widely distributed in all continents of the globe, except for Antarctica (Rubiales and Heide-Jørgensen, 2011). It covers plants with all levels of host-dependency ranged from free-living non-parasites to facultative or obligate root parasites. Phylogenetic tree of Orobanchaceae built based on sequences of photoreceptor phytochrome A (PHYA) (Bennett and Mathews, 2006) showed that the nonparasitic *Lindenbergia* genus is placed apart from parasitic species. In addition, facultative parasites *Triphysaria versicolor* and *Phtheirospermum japonicum* are phylogenetically close to obligate parasites from the genera *Striga*, *Orobanche* and *Phelipanche*, which have evolved to parasitize economically importance crop species. Obligate hemiparasite *Striga hermonthica*, popularly known as witchweed, parasitize Gramineae species, such as sorghum, maize, millet and rice. Fields infected by *S. hermonthica* reduce their yields between 20 – 80%. The problem is more severe in the Africa where 25 countries have reported huge yield losses due to *Striga* infestations, and the impact on their economies reaches the sum of 1 billion USD per year (Pennisi, 2010). The droughts, high temperature and infertile soils together with social fragility contribute to the rapid appearance of fields severely infected by *Striga* (Spallek et al., 2013). Out of Africa, the holoparasite broomrapes (*Orobanche* and *Phelipanche*) represent a risk for agriculture in Middle East and in Mediterranean region, causing significant economic losses for farmers of sunflowers, legumes, potatoes, carrots, tomatoes and *Brassica* crops (Parker, 2009; Rubiales and Fernández-Aparicio, 2011; Moreno and Rubiales, 2008).

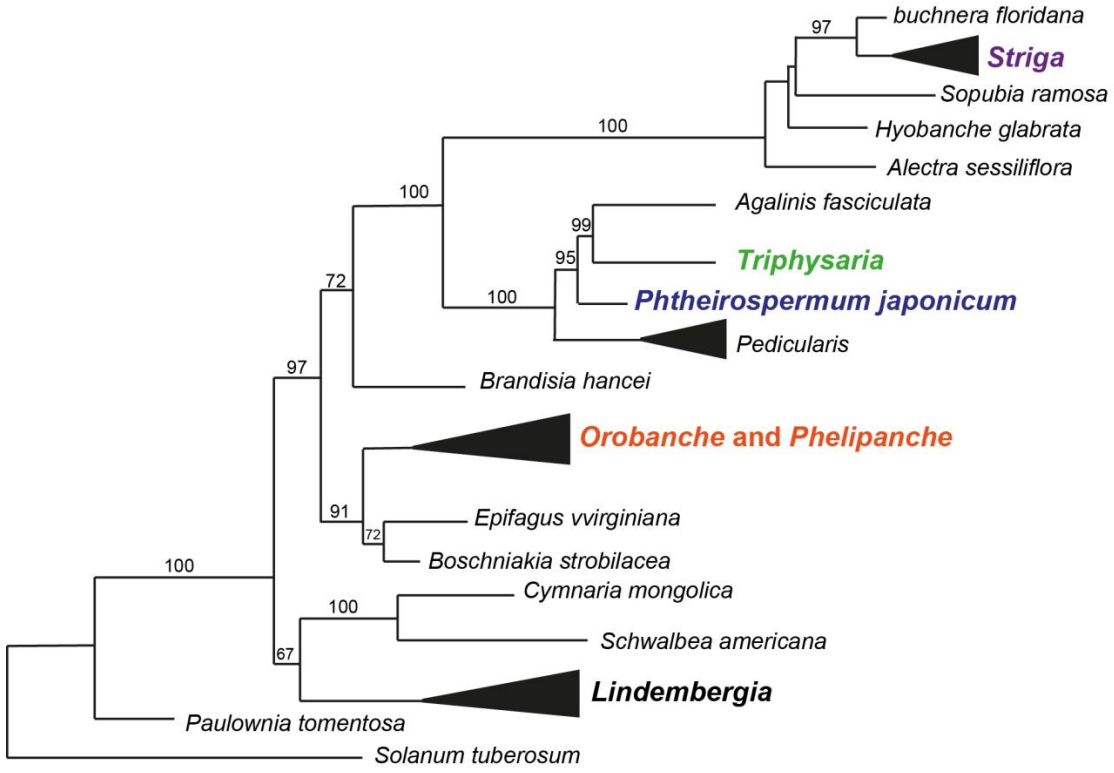


Figure 1.1 Phylogenetic tree of Orobanchaceae. Highlighted in bold the obligate parasites *Striga*, *Orobanche* and *Phelipanche*; facultative parasitic plants *Triphysaria* and *Phtheirospermum japonicum*; non-parasitic plant *Lindenbergia*. The tree was modified from Bennett and Mathews, 2006.

1.2.1 PARASITIC WEED - AGRICULTURAL PEST

The world population is expected to reach the number of 9.7 billion by 2050. Major part of this population will come from developing countries especially those in Sub-Saharan region, which is rising faster than the rest of the world. By 2050 is expected that 2.2 billion people will live in the Sub-Saharan Africa compare with the current population of 0.9 billion (data by Population Reference Bureau - www.prb.org). Along this expectation, there will be serious concerns about the worldwide food production. Industrialization and economic activities may have affected continuous increase of the global temperature, causing desertification of arable lands and water shortage in agriculture. Moreover, 20 % to 40% of world potential food production is lost every year due to effects of weeds, pest and disease (Food and Agriculture Organization of the United Nations or FAO). In Sub-Saharan Africa, the most serious damage for cereal production is caused by a parasitic plant, *S. hermonthica*.

The parasite *Striga* infests fields of sorghum, maize, millet, rice and sugarcane, which represent the main source of carbohydrates of subsistence farmers in Sub-Saharan Africa (reviewed in Spallek et al., 2013). The crop field infected by *Striga* registered losses ranged from 20% to 80% and many of farmers who had infested fields ended up with abandonment of their fields (reviewed in Atera et al., 2011 and Waruru, 2013). Currently, the area affected by *Striga* is estimated around 50 million hectares (Westwood et al., 2010), affecting the lives of 100 million people (Pennisi, 2010).

Other economically important parasitic plants are broomrapes, chlorophyll-lacking obligate parasites belonging to *Phelipanche* and *Orobanche* genera. They attack various crops and vegetables, including legumes, tomatoes, sunflowers and carrots (Parker, 2009; Rubiales and Fernández-Aparicio, 2011; Moreno and Rubiales, 2008). Broomrapes are mainly found in agricultural areas in Mediterranean Europe, Northern Africa, Middle East and parts of Asia. *Orobanche crenata* is estimated to cause average 30% yield losses in faba beans in Egypt, and over 50% in Turkey, while *Orobanche cumana* reduces the sunflower yield around 50% in Greece and China. Recently, *O. cernua* also became a serious problem for tomato plantation in Southern Europe and Northern Africa and in tobacco in India. In *Phelipanche* genus, *Phelipanche ramose* is likely to cause the most serious damage in agriculture. In Italy, it was reported losses up to 80% of tomato, cauliflower and eggplant fields. High infection levels by *Phelipanche aegyptiaca* resulted in 30% and 70% reduction of carrot and potato production in Israel and Iran, respectively. The impact of broomrape in the world agriculture is carefully reviewed by Parker (2013, 2009). Recently, it was reported for the first time that *P. aegyptiaca*

infected on kenaf (Lati et al., 2013), a Malvaceae crop used as a basis for fiber production in Middle East. This fact suggests that continuous evolution pressure results in adaptation to new hosts and warns the possibility of other crops which are not suffered by parasitic plants are at a risk of future infestation.

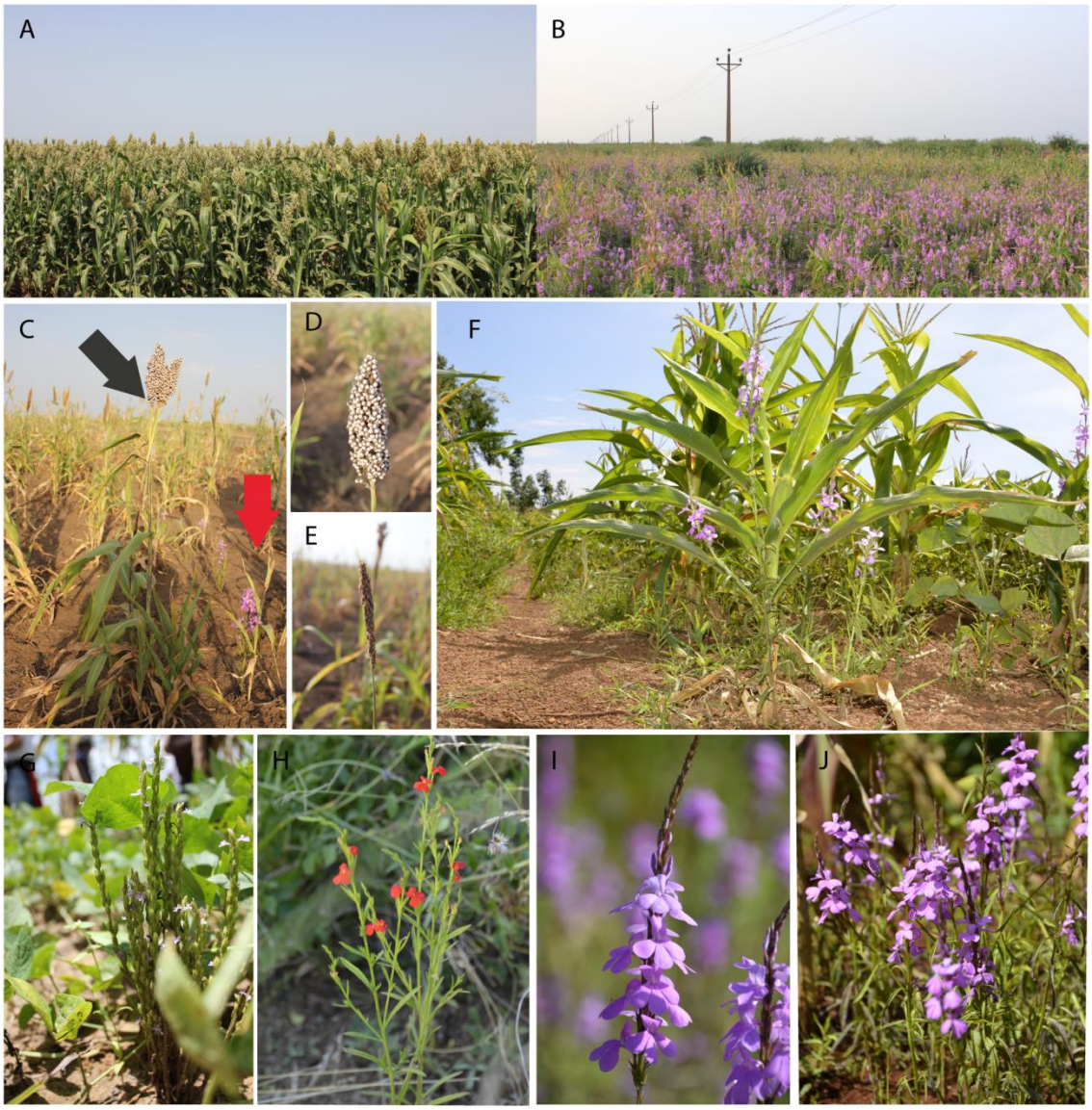


Figure 1.2 Photographs of *Striga* species. **A-** A sorghum field without *Striga* infestation. **B-** A sorghum field infested by *Striga hermonthica.* **C-** A sorghum plant growing without attack of *S. hermonthica* (left, black arrow) and a Striga-parasitized sorghum plant (right, red arrow). **D-E** Sorghum ears from a non-parasitized (D) and a parasitized plant (E). **F-** Maize infected by *S. hermonthica*. **G-** *Striga gesnerioides* infecting its host cowpea. **H-** *Striga asiatica*. **I-J** *Striga hermonthica* flowers. Photographs were courtesy of Satoko Yoshida and Musembi Mutuku.

1.3 LIFE CYCLE - OBLIGATE AND FACULTATIVE PARASITE

Some of weedy parasitic plants produce a large amount of tiny seeds. This characteristic is linked with their survival strategy to unsure seed dispersal to increase the chance to attach a compatible host. The obligate parasitic plants in Orobanchaceae also acquired unique features of seed dormancy, which limit the germination under unfavorable conditions. For example, the *Striga* and *Orobanche* seeds need to go through a period of wet and high (~30 ^{o}C) temperature (conditioning) (Mohamed and Ejeta, 1998), which in laboratory is mimicked by keeping the seeds on moisten filter paper at 26-30 ^{o}C for period of 7 days. During the conditioning, high gene expression and protein synthesis occur and activate physiological and metabolic changes in viable seeds (Nun and Mayer, 1993). However, the conditioned seeds enter into as the status of secondary dormancy. To break the secondary dormancy, the seeds need to be exposed to host-released germination stimulants. Among several known germination stimulants, strigolactones (SL) are the most potent stimulants (reviewed in Xie et al., 2010). SLs have other important functions and the details are described in the next subsection. Close association between parasite and host is essential to ensure parasite's survival, since nutrient-poor seeds and photosynthetic deficiency of obligate parasites are not able to survive after the hypocotyl stage in the absence of host connection.

In contrast, the life cycle of a facultative parasite is marked by the ability to induce the germination of its seeds without stimulus from other plants, and thus they can live without a host. However, the facultative parasite leaves the free living stage when host-derived molecules are detected. The critical step for shifting its lifestyle is haustorium development. The formation of this feeding structure is induced in both obligate and facultative parasite upon reception of chemical stimuli (HIF - haustorium-inducing factors) provided by the host. Interestingly, HIFs are perceived differently depending on parasites. For example, the obligate parasites, such as *Striga* and *Orobanche,* initially form a haustorium at the tip of a radicle (terminal haustorium), and once the haustorium succeeds to infect a host root, adventitious roots emerge and secondary haustoria form at the lateral side of roots (lateral haustorium). In contrast, the facultative parasite, such as *P. japonicum* and *T. versicolor* develops only the lateral haustorium but not the terminal haustorium. Thus, it has been speculated that the lateral haustorium is the ancient characteristic of parasitic plants in Orobanchaceae, while the terminal haustorium is a derivate feature acquired only in the obligate parasites (Westwood et al., 2010). The terminal and lateral haustoria are structurally different; in the terminal haustorium it is apparent that the reorganization of the meristematic cells at terminal portion

of the root, compromises the continuous root growth, whereas the lateral haustorium appears at the upper part of the meristematic zone without affecting the root growth.

1.3.1 STRIGOLACTONES – GERMINATION STIMULANT FOR OBLIGATE PARASITES

Strigolactones were initially described as host-released molecules, which are able to induce seed germination of obligate parasitic plants. Decades later, however, it was proved that SLs are the new class of plant hormones controlling the shoot branching. Furthermore, researchers found that SL action is not restricted to shoot branching. SLs regulate different aspects of plant development, including lateral root formation, root hair elongation (Kapulnik et al., 2011), extension of internodes (de Saint Germain, Ligerot, et al., 2013), senescence and nodulation in *Lotus japonicus* (Liu et al., 2013). In addition, SLs also associate with auxin signaling in the regulation of secondary growth (Agusti et al., 2011). The SL production is derived from the carotenoid pathway. The all-trans-β carotene is converted by an isomerase called D27 to 9-cis-β-carotene, which is the substrate of Carotenoid Cleavage Dioxygenase 7 (CCD7), encoded by *MAX3* in *Arabidopsis* (Booker et al., 2004), *RMS5* in pea , *D17* in rice and *DAD3* in petunia (Napoli and Ruehle, 1996). CCD7 catalyzes the reaction converting it to β-apo-10'-carotenal. Next, carlactone is produced by the action of CCD8 (*MAX4* in *Arabidopsis*, *RMS1* in pea, *D10* in rice and *DAD1* in petunia). Carlactone is supposed to be converted to a strigolactone by a putative cytochrome P450 enzyme, probably the homolog of enzyme MAX1 in *Arabidopsis*, but this step is still not fully understood (reviewed in Saint Germain et al., 2013). Interestingly, SLs are detected in species that lack known SL-biosynthesis genes, such as liverwort *Marchantia polymorpha* which lacks *CCD8* gene in its genome or green algae Charales that miss *CCD8* and *CCD7*, (Delaux et al., 2012) . Thus, it is predicted that addition to the known pathway mediated by CCD8, an alternative CCD8-independent SL biosynthesis pathway is likely to exist.

As mentioned previously, SL was originally discovered as a germination stimulant for parasitic plants. However, from an ecological point of view there is no benefit for a host plant to report its own localization to destructive parasitic partners. Thus it was hypothesized that SLs may have alternative biological roles. In 2005, SL was found to be an essential mediator between plants and symbiotic arbuscular mycorrhizal (AM) fungi (Akiyama et al., 2005). This mutualistic association provides organic compounds to fungi and in return its highly branched hyphae increase the area for the plant nutrition uptake, extending it to region beyond the root system can reach. Under limited nutrient conditions, such as low phosphorus or nitrogen, SL

biosynthesis increases (Yoneyama, et al., 2007). As consequence, the plant reduces the shoot branching and stimulates AM fungi colonization. The parasitic plants adopted the signal of this beneficial interaction to recognize their potential hosts by stimulating of seed germination. This survival strategy increases the chances of parasites' small seeds to connect to a potential host.

Is the original role of SL in evolution as AM-attractor molecule or a hormonal role such as a shoot branching inhibitor? The presence of SL in green algae Charales favors the idea of primitive function of SL was mainly hormonal rather than for fungal attraction. In Charales, SLs are associated with the control of rhizoid elongation (Delaux et al., 2012). In a similar way, in *Physcomitrella patens* SLs are detected in the exudate and moss tissues, being related with the regulation of filament branching. However, the function of SL as a communication signal seems also evident, as in the bryophyte the external released SL is used to regulate the extension of moss community in response to population densities (Proust et al., 2011).

Few molecules are known to be involved in SL signaling. One of them is D14 α-β hydrolase from rice (Arite et al., 2009), the ortholog of DAD2 in petunia (Hamiaux et al., 2012). The enzyme is formed by a lip with four α-helices and a core composed by seven α-helices and seven β-sheets, forming a high hydrophobic pocket where SL coordinates with the catalytic triad Ser-His-Asp (Zhao et al., 2013) (Hamiaux et al., 2012). The D14 binds directly to SL to catalyze its hydrolysis, and the SL-D14 complex becomes stable after the association with MAX2, an F-box protein, which is a part of the SCF complex (Hamiaux et al., 2012). The typical structure of SL is formed by four rings (A-D), and the D-ring is credited to have an active function for inhibition of shoot branching (Fukui et al., 2013; Boyer et al., 2012). The synthetic compound 4-Br-debranone (4BD) shares the butenolide moiety with SLs. Although it is able to mimic most of its hormonal functions, 4BD was shown to have less activity in inducing germination in parasitic plant seeds or in promoting AM hyphal branching (Fukui et al., 2013). Similarly, natural and synthetic SL compounds showed a range of SL activities. Structure-activity relationship studies suggest that the structural requirements for SL response in AM fungal and parasitic plants seems to be associated with an enol ether bridge in connection between C and D rings (Akiyama et al., 2010). These observations imply the existence of multiple mechanisms to perceive SL molecules to activate various physiological responses. This hypothesis might explain how the parasitic plant distinguishes host-released SLs from self-produced hormone. It is expected that at least the obligate parasites *Striga spp.* are likely to be able to produce SLs since known SL-biosynthesis homologs are present in the recently

sequenced *S. asiatica* genome (Yoshida et al. in preparation), though biochemical evidence is still missing.

The smoke-derived compound Karrikins shares the butenolide structure with SL. Karrikins stimulate the germination of certain plants after field fire. Ecologically, Karrikins play an essential role as the germinated plants can grow in fire-destroyed field with plenty of light virtually without competition with other plants. Interestingly, similarities between Karrikins and SLs are beyond the structure. Similarly to SL, Karrikin also binds to α-β hydrolase KAI2 (Janssen and Snowden, 2012). In addition, genetic evidences suggest a physical interaction of KAI-Karrikin with MAX2 (Waters et al., 2012), the same F-box protein that is the substrate of the SCF ubiquitin E3 ligase involved in the SL signaling. Although such interaction still needs to be proved by biochemical approaches, it strongly suggests that the Karrikin- and SL-mediated responses converge to MAX2. Such finding may provide a clue how the perception of SL as a germination stimulant in parasitic plants evolved.

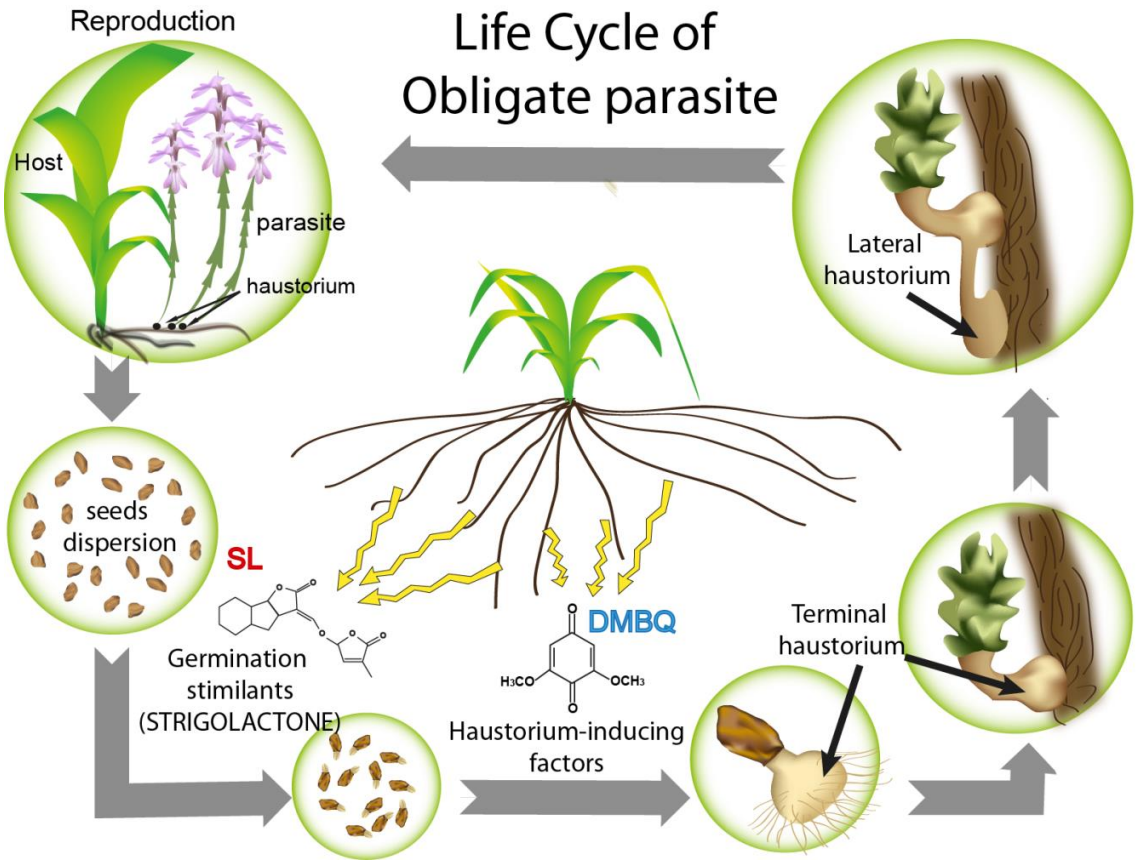


Figure 1.3 - Life cycle of obligate parasitic plant.

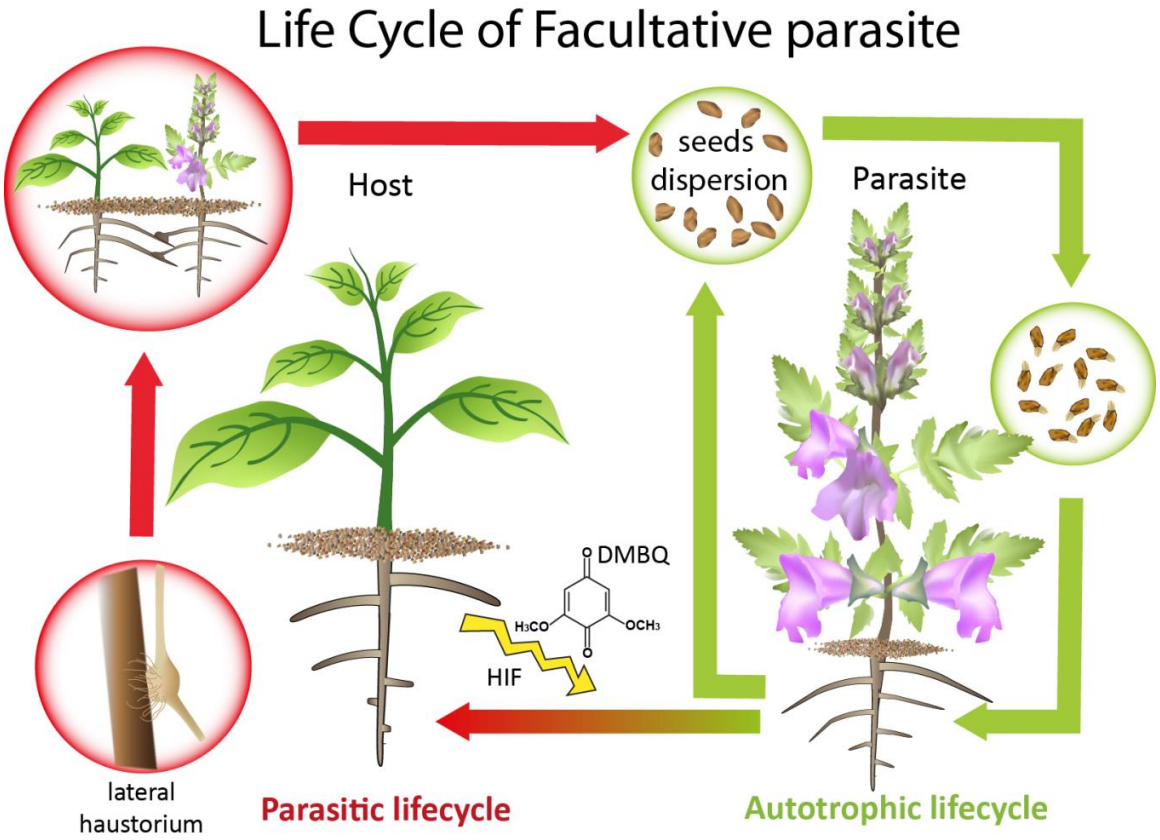


Figure 1.4 - Life cycle of facultative parasitic plant.

1.4 HAUSTORIUM

The word haustorium came from the Latin *haustus* that means drain or drink, and, because of its Latin origin, the plural is haustoria. The term is largely utilized to refer for diverse structures that are related with absorption and nutrition of a variety of unrelated organisms, as occurs in fungal hyphae, sporophyte of mosses, and in parasitic plants. In this last group, especially in Orobanchaceae, the haustorium acquires different roles according to its developmental stage. At the beginning, it is an attachment organ that physically connect the parasite to the host. Next the parasite penetrates into host tissue. In a mature stage, the haustorium turns to a feeding structure, transferring water and nutrients through vascular connection between the host and the parasite. The next section focuses on the haustorium development in facultative parasitic plants, although some features are also extended to obligate parasites.

1.4.1 MECHANISM OF HAUSTORIUM INDUCTION

For haustorium initiation is closely linked with the host presence. The obligate and facultative parasites sense their potential hosts through physical contact and the host-provided chemical stimuli such as HIF. The HIFs are normally found in host root exudates, and the parasite roots exposed to HIF induce haustoria *in vitro*. Active HIF compounds include phenolic derivatives, anthocyanins, flavonoids and quinones, although degrees of haustorium inducing activities vary among compounds, and optimization of the conditions and concentration is required to achieve the higher percentage of haustorium induction (Albrecht, 1999). The natural HIF, 2,6-dimethoxy-p-benzoquinone (DMBQ) was originally identified from cell wall fragments extracted from Sorghum root surface, and known to activate morphological changes in parasite roots resembling haustorium development (Chang and Lynn, 1986). Although haustoria can be observed at different parts of roots, the area distal to the elongation zone is most sensitive to HIF. Typically, this area becomes swollen and forms haustorium at 24 - 48 h of DMBQ treatment.

Histological studies in a facultative parasitic plant *Agalinis purpurea* revealed morphological changes during haustorium development in detail (Riopel and Musselman, 1979; Wm. Vance Baird and James L. Riopel, 1984). One of earliest events is radial enlargement of cortex, spreading from inner to outer layers. Around 10 h after HIF treatment, it was observed that a haustorial apex forms a group of epidermal cells with dense cytoplasm without anticlinal divisions, and concomitantly other epidermal cells start differentiating into

haustorial hairs. After 20 h, the periclinal division at cortical and pericyclic layers occurs. From 24 h to 36 h the increased volume of haustorium is mainly caused by enlargement of cortical cells, and the cell division is continuously observed in the inner cortex and pericycle. Similar morphological events were observed in the facultative parasitic plants *T. versicolor* (Bandaranayake and Yoder, 2013a) and *P. japonicum* (Ishida et al., 2011), suggesting a conserved mechanism for haustorium development.

Molecular mechanism to trigger initiation of haustorium is appear to involve redox signaling. Quinones and flavonoids with redox potential are able to induce haustorium formation in Orobanchaceae plants (Matvienko et al., 2001; Albrecht, 1999). Current model for haustorium inducing signal suggests involvement of the catalyzation of quinone compounds by single-electron reducing quinone reductase (QR1). Previous studies in *T. versicolor* described that TvQR1 enzyme generates the unstable radical semiquinones and reactive oxygen species (ROS), as sub-product. In contrast, TvQR2 catalyzes two electron reductions and generate no-harmful hydroquinone, preventing the formation of potentially toxic semiquinone intermediates. *T. versicolor* genes that encode these enzymes are highly up-regulated after the DMBQ treatment. Interestingly, RNAi silencing of *TvQR1* led to reduction of the haustoria number, while the knock-down transgenic roots targeting for *TvQR2* show normal haustorium development (Bandaranayake et al., 2010). Thus, it was hypothesized that *TvQR1* catalyzes univalent reduction of quinone derived from the host, activating the signal transduction pathway of haustorium development via a redox mechanism. The role of *TvQR2* is speculated to be involved in the cell protection against the generation of harmful semiquinones and ROS.

One of the early regulated genes essential for haustorium induction in *T. versicolor* encodes Pirin (TvPirin), a nuclear protein associated with transcription factor activity. *TvPirin* has homologs in parasitic and non-parasitic plants and was originally identified in haustoria-specific cDNA library. The importance of *TvPirin* for haustorium development is reinforced by experiment with transgenic roots with lower *TvPirin* transcripts levels, in which the percentage of haustoria were compromised (Bandaranayake et al., 2012). Further analysis is necessary to link the participation of Pirin and QR1 in the haustorium development pathway. For instance, it is not yet shown if Pirin-mediated and QR1-mediated pathway converge at some point.

1.4.2 PENETRATION INTO HOST TISSUES

The haustorium development triggered by HIF treatments *in vitro* occurs in a similar way to that induced by host root contact. It is driven by coordinated morphological events,

including haustorial hair prolongation, cell expansion and cell division. The chemical or host-root-exudate-induced haustorium develops enlargement of lateral parts of a root, frequently near to the root tip. However, *in vitro* haustorium development is not completely same as the natural host interaction; it lacks differentiation of xylem cells inside the haustorium unless attaching host roots. It is noteworthy that the chemical-induced haustorium maintains its ability to penetrate to host roots when it is placed nearby host roots (Smith et al., 1990). Therefore, it suggests the existence of unknown signals probably derived from host that are necessary for the subsequent haustorium developmental stages. Host-specific differential expression pattern of parasite genes reinforces the existence of those factors that are recognized by the parasite. Micro-sectioned tissue of the hemiparasite *T. versicolor* collected at the interface with a monocot (*Zea mays*) or dicot host (*Medicago truncatula*) showed that monocot-parasitized parasite preferentially liberates expansin-β at the apoplastic region rather than expansin-α (Honaas et al., 2013). The expansin proteins function in the cell wall loosening in a non-enzymatic manner, promoting plant cell enlargement via low extracellular pH (so called "acid growth"). The major target of expasin-β proteins are β-(1$\rightarrow$3), (1$\rightarrow$4)-D-glucan and glucuronoarabinoxylan, which are the component of cell wall in grass species, whereas the expasin-α target xyloglucan and pectin, enriched in dicots primary cell wall (reviewed in Sampedro and Cosgrove, 2005). Hence, parasites are likely to be able to sense the host difference and release molecules aiming to loosening the cell wall according to the host type.

At the penetration stage, the host defense response represents a barrier for successful establishment of parasitism. The resistant mechanism remains elusive in several perspectives, but it is known that at least in the case of obligate hemiparasite *S. hermonthica* it occurs at different levels depending on plants. In the nonhost *L. japonicus, S. hermonthica* stops at the cortical region without reaching the vasculature (Yoshida and Shirasu, 2009). Chromatography and mass spectroscopy analysis revealed that the legume-specific phytoalexin vestitol contributes, at least partially, to the resistant process since it reduces significantly the parasite seedling growth (Eda and Sugimoto, 2010). *Arabidopsis* and cowpea do not support the complete lifecycle of *S. hermonthica*, although the parasite is able to establish with them a certain level of vascular connection (Yoshida and Shirasu, 2009).

Once the host barriers are overcome, the parasite still needs to open space into host tissues. Although mechanisms would differ depends on species, the members in Orobanchaceae chose combination of enzymatic and mechanical way to penetrate host tissues

(Alejandro, 2013). One relevant enzyme in this process is Pectin Methyl Esterase (PME) detected in intrusive cells of Orobanche by cytochemical techniques using specific antibodies. This pectolytic enzyme acts at infection zone of host cortex by changing cell wall composition in favor of low esterified pectin, debilitating the host cells (Losner-Goshen, 1998). In the stem parasitic plant *Cuscuta reflexa* during the invasion, the expression of cysteine proteinases are induced. Those enzymes are transcribed as an inactive propeptide form and are activated by releasing the inhibitor domain. It is assumed that the role of this proteinase is related with weakening of host structures by protein degradation (Bleischwitz et al., 2010). Interestingly, expression of a similar enzyme is also shown to be up-regulated in phylogenetically distant Orobanche species, suggesting its relevance for parasite penetration (Rehker et al., 2012). The mechanical driven force relies on successive cell division and elongation of the parasite intrusive cells toward the host stele (Alejandro, 2013).

1.4.3 ESTABLISHMENT OF VASCULAR PARASITE-HOST ASSOCIATION

At later stages, a haustorium becomes a physiological bridge that allows the parasite to withdraw water and nutrients from the conductive system of host plants. Differentiation of intrusive cells into tracheary elements is observed in a haustorium when host and parasite are connecting and the differentiation seems to occur along the auxin flow in the haustorium. Experiments with *Phelipanche aegyptiaca* provided an evidence for association of successful establishment of terminal haustorium with the flow of auxin as treatment of IAA antagonist results in drastically reduction of haustorium number (Bar-Nun et al., 2008). It is not clear in the case of haustorium development if auxin accumulates at haustorium region is due to *de novo* auxin production or being transported.

Materials exchanged between parasite and host are not limited to organic substances; the mature haustorium allows bidirectional movement of wide range of other substances including protein (Aly et al., 2011) virus (Gal-On et al., 2009) and nucleic acids (Westwood et al., 2009). Protein exchange between host and parasite was visually demonstrated by Aly et al (2011) using transgenic host expressing green fluorescent protein (GFP) parasitized by non-transgenic parasitic plants *P. aegyptiaca*. After infection, a dense accumulation of GFP was detected in the parasitic tissues, suggesting the transference of GFP, a cytosolic protein, from host to parasites probably though plasmodesmata (Aliy et al., 2011). Similarly nucleic acid exchange was detected in the facultative *T. versicolor* (Tomilov et al., 2008). In *T. versicolor*, the host harboring GUS-targeted RNA hairpin construct is able to

suppress *GUS* activity at the intersection between host and parasite genetically modified to express this reporter gene, suggesting translocation of smalls RNAs from host to parasite.

The trafficking of molecules in host-parasite interaction raised the possibility to adopt a novel strategy to combat parasitic plants: using a host plant as vehicle to deliver into parasite tissue the double stranded hairpin RNA (hpRNA), which targets a parasitic-specific gene (Yoder et al., 2009). Engineering host resistance against parasitic plants with RNA interference was demonstrated with success in the parasite *P. aegyptiaca* (Aly et al., 2009) and *T. versicolor* (Bandaranayake and Yoder, 2013b). In the former case, the authors transformed host plants with silencing construct targeting the parasite Mannose 6-phosphate reductase (M6PR), a key enzyme in mannitol biosynthesis that is important for *P. aegyptiaca* development. In the latter case, *M. truncatula* was genetically transformed with a hairpin construct targeting to acetyl-CoaA carboxylase (ACCase) gene of *T. versicolor*. In both cases the interspecific silencing of parasite genes resulted in a drastic reduction of parasite infection. Similarly, transgenic tobacco plants harboring a RNAi construct targeting for KNOTTED1-like HOMEOBOX1 (KNOX1), a developmental gene known to promote cytokinin biosynthesis in the meristem of stem parasite dodder. Dodder-parasitized transgenic tobacco plant increased its fitness through reduction of the dodder infection and by compromising the parasite fertility (Alakonya et al., 2012).

1.5 METHODS TO CONTROL PARASITIC PLANTS

Parasitic weeds are the serious threat for food security worldwide and the methods to control them have a limited effectiveness especially because of the complex nature of the parasites. The parasites produce small and long-living seeds. *Striga* seeds can keep its dormancy for over ten years in soil and it is difficult to detect their presence until starting to damage crops (reviewed in Spallek et al., 2013). The application of herbicides faces a difficulty because both the host and the parasite are plants. Development of resistant cultivars shows in general better results for pest control if good genetic sources of resistant cultivars are available in natural crop population. However, the complete resistant against parasitic plants seems does not exist and only a partial tolerant phenotype has been described so far.

1.5.1 CONTROL THROUGH ENHANCE GERMINATION

Suicidal germination

Suicidal germination involves the application of germination stimulants in the soil to induce the germination of obligate parasite seeds in the absence of host crops to reduce the soil seed bank. This approach leads parasites to death since the small seeds of parasitic plants do not possess enough nutrient sources to support the parasite growth, and therefore it is necessary for the parasites to associate with a compatible host. The SLs are natural germination stimulants for obligate parasites. However, SLs are not suitable for field application since their structures are too complex and the large scale synthesis is prohibitive expensive (Sugimoto et al., 1998). To overcome this barrier artificial germination stimulant GR with simpler structure was projected, including GR-24 (Johnson et al., 1981). However, the high humidity necessary to overcome the seed dormancy (conditioning) reduce the germination induction efficiency of GR24 (BABIKER et al., 1987). Currently, there is a great effort from the scientific community to synthesize a germination stimulant compound showing stability in soil and reducing its synthesis cost suitable for field application (Zwanenburg et al., 2011) (Prandi and Franca, 2011) (KGOSI et al., 2012). The most successful application of suicidal germination approach to combat parasitic plants was carried out in United Stated in the states of South and North Caroline. *Striga asiatica* was first identified in 1955 by an Indian graduate student who knew the pest from his country. The origin and the exact time that the parasite arrived in US has still been obscure (APHIS, 2011). To control *S. asiatica* the the US government launched a Federal-State Programme for applying a large scale of ethylene gas, which also induces its germination,(Langston M. and English TJ., 1990). The strategy was

succeeded in US, but it still remains an expensive alternative for the poor countries in Africa. Other reasonable alternative is the use of the biocontrol agent *Pseudomonas syringae* pathovar *glycinea* (Psg), an ethylene-producing rhizobacterium, which can induce the suicidal germination of *Striga* seeds (Ahonsi et al., 2003) and the studies using this bacterium are currently ongoing.

Trap and catch cropping

The trap and catch croppings are alternative approaches that aim to control the parasite by enhancing germination. In the trap crop strategy, suicidal germination of parasite seeds is induced by rotating or intercropping with species not recognized by the parasite as a host, but still able to release germination stimulants into rhizosphere. Soybean (*Glycine max*), cotton (*Cicer arietinum*), and sesame (*Sesamum indicum*) are among the trap crops for *Striga*. Other intercrop species with successful field-tests is *Desmodium*, a legume crop, which increase the soil fertility while induce the germination of the parasites (reviewed in Goldwasser and Rodenburg, 2013). In contrast, in the catch crop approach, host crops that support the parasitism are planted in the field. The parasite germinates and attaches to the compatible host, which is removed together with the parasites before the flowering and seed dispersion (reviewed in Goldwasser and Rodenburg, 2013).

1.5.2 CONTROL THROUGH REDUCED GERMINATION

Increase the soil fertility

Application of fertilizers can also compromise the parasitic plant proliferation. Parasitic plant infestation is more accentuated in soil poor in nutrients. This fact is associated with the finding that host plants exudate more SLs under phosphate (Yoneyama, Takeuchi, et al., 2007) and nitrogen starving conditions (Yoneyama, Xie, et al., 2007). Thus fertilized soil prevents the crops to release SLs into rhizosphere and reduces the germination and consequently infestation rates of parasitic weeds (Smaling et al., 1991).

Chemical control

In general, herbicides show a limited efficiency to control parasitic plants. The critical steps of the pest lifecycle occur in the underground stages when the herbicide application is not effective. The administration after the parasite stem emerges from the soil and it is already too late to prevent yields losses (reviewed in Eizenberg et al., 2013). One alternative to

circumvent the problem is coating crop seeds with pesticides. The pesticide-coating seeds of imidazolinone-resistant maize showed effective for fields with moderate *Striga* infection. The practice is more powerful if it is combined with continuous farmer vigilance in removing flowering *Striga* plants, which can eventually be resistant to herbicide and fertilizer applications (Ransom et al., 2012). The ample adoption of this technology by the farmers is limited and depending on implementation of the logistic system by the governors and the private sector (De Groote et al., 2008). Although soil fumigation may also be effective for reducing the witchweed seed bank, its usage is rather limited due to its toxicity for humans, requirement of special equipment for application, high costs, and environmental damages (Goldwasser and Rodenburg, 2013).

1.5.3 RESISTANT CROP

Compare with other methods for protecting crop yield against parasitic weeds, breeding of resistant lines is expected to be the most effective, particularly because the other alternatives are not often applicable for poor subsistence farmers in Sub-Saharan Africa. However, to improve cultivars, it is necessary to have a good source of resistance lines and an efficient screening procedure to provide a sufficient selection pressure (Mower et al., 2010). Unfortunately, this is rarely the case for the parasitic weeds. Breeding for resistance against parasites is difficult to assess, especially due to its complex nature and low heritability (Haussmann et al., 2000) (Moreno and Rubiales, 2008). So far only a few agronomically important species showed resistance or tolerance to the parasite. Resistant is assumed as the ability of the host to prevent the parasite growth and development, while the tolerant plant are those that show eventually less damages caused by the parasite infection. It is noteworthy that the resistance against the parasitic plants is associated with the different stages of parasite lifecycle; it may occur pre-attachment or post-attachment to host root.

Genetic of Resistant - Monogenic or Multigenic resistant to parasitic weed

The genetics of parasitic weed resistant is complex and ruled by co-existence of different mechanisms depending on type of host-parasite interaction. *S. hermonthica* showed host specificity for monocots and its interaction seems to be associated with multiple levels of resistance (Yoshida and Shirasu, 2009), while the *S. gesnerioides* is able to infect the dicot cowpea controlled by a single dominant gene (Li and Timko, 2009a). The evidence of genetic resistance controlled by a gene that encodes a resistance (R) protein involved in plant defense response was revealed in the studies of race-specificity of the *S. gesnerioides*-cowpea

interaction. *S. gesnerioides* race SG3 succeeded to parasitize diverse varieties of cowpea, but failed to infect the cultivar B301 (Li and Timko, 2009a)

In contrast, the polygenic resistant is also largely reported in the *Striga* genus. The 9-11 quantitative trait loci (QTL) are responsible for 77-82% of the post-attachment resistant phenotype in Sorghum lines against *S. hermonthica* (Haussmann et al., 2004). The rice cultivar NERICA (NEw RICe for Africa) exhibited the post-attachment resistant against to *Striga* species. It was initially developed by Africa Rice Center from the hybrid between *Oryza sativa* and *Oryza glaberrima* (Jones et al., 1997). Quantitative and histological studies showed that the resistant occurs at different levels ranging from strong resistant to tolerance phenotypes, and even in the most resistant rice cultivars, NERICA 1 and NERICA 10, the damages caused are significant compared with non-infected rice (Cissoko et al., 2011).

Similarly, the broomrapes interactions with their hosts are as complex as those observed in *Striga* species. The five single dominant genes *Or1* - *Or5* seem to be associated with race-specific resistance after parasite penetration between sunflower and *O. cumana*. The *Or5* can confer resistant to *O. cumana* race E, but is not effective for the parasite race F (Alonso et al., 1996). In contrast, the interactions of broomrape species with legumes are controlled by multi-components. Three QTLs in Faba beans were identified to be responsible for over 70% of resistant phenotype against to *O. crenata* (Román et al., 2002). In the same way, the resistance described in legume pea (*Pisum sativum*) is reflection of the expression of 2 QTLs responsible for 9.6% and 11.4 of phenotype, respectively (Valderrama et al., 2004).

1.6 PARASITIC PLANT GENOME AND TRANSCRIPTOME

The negative effects of parasitic plants are aggravated by the lack of knowledge of their biology, evolution and molecular strategies to infect the host. The advent of new sequencing technologies and the dramatical reduction of sequencing cost have been contributing to change this scenario. A large number of non-model organisms have been targeted to the genome wide sequencing, including parasitic plants. An expressive step forward to the elucidation of the plant parasitism is the public availability of full nuclear genetic information of the parasitic plant *Striga asiatica* that is currently being carried out by the group of Ken Shirasu and Satoko Yoshida. *S. asiatica* was chosen due to technical advantages over the other *Striga* species. *S. asiatica* has a relatively small genome size, about 600 Mb, while *S. hermonthica* shows a genome with a size around 1.6 Mb. Besides, it is a self-crossing plant, whereas *S. hermonthica* is an obligate out-crossing one. Thus it is expected that

the comparative analysis of non-parasitic plant may bring new findings about the biological and evolutional aspects of parasitic angiosperms.

To increase the genetic information for the parasitic plant field a consortium formed by American research group called Parasitic Plant Genome Project generated large-scale transcriptome data from different stages across the lifecycle of 4 species in Orobanchaceae, *S. hermonthica*, *P. aegyptiaca*, *T versicolor* and the non-parasitic *L.indenbergia philippensis*. The raw reads and *de novo* assembled sequence contigs are available for community through the website http://ppgp.huck.psu.edu/. One interesting finding generated by PPGP is observed in the holoparasite *P. aegyptiaca.* This parasite is a green-less plant that lost its capacity to make photsynthesis, but its transcriptome analysis revealed the enzyme responsible for chlorophyll synthesis (chlorophyll a/b synthesis and chlorophyllide a oxydase) are still detectable. However, the chlorophyll binding proteins of the photosynthetic apparatus are not expressed in above-ground tissues of these parasitic plants. It raised a surprising hypothesis that the detectable chlorophylls might be serving an alternative role in the parasite besides photosynthesis (Wickett et al., 2011).

Transcriptome of *S. hermonthica* showed that a haustorium can serve as bridge for horizontal gene transfer (HGT) between host and parasite. Yoshida et. al. (2010) described the first evidence for a nuclear gene found in *S. hermonthica* and *S. gesnerioides* which is closer to its homologs in Poaceae but not present in other dicot genomes (Yoshida, Maruyama, et al., 2010). Other evidence of HGT is described in plastid genome of *Phelipanche* and *Orobanche* genera. The presence of a gene with high homology with each other in those species indicates that HGT occurs between both parasite species probably using a common attacked host as a vector for RNA transference (Park et al., 2007). Outside Orobanchaceae, material genetic exchanging is found in the interaction between the parasite dodder (*Cuscuta* ssp*)* and the host *Plantago* ssp; it was shown that three mitochondrial genes were recently transferred from the parasite to the host genome (Mower et al., 2010). In the parasitic *Rafflesiaceae*, transcriptome analysis revealed that at least dozen actively transcribed genes were acquired from host genome (Xi et al., 2012). Therefore, the large functional genomic data reveals an interesting aspect of the parasitic plant biology, that is, the exchange of genetic material through the physical link between host and parasitic plants. However, other important questions remain elusive such as the molecular mechanisms behind the gene transfers and whether these transferred genes have a function in the parasite-host communication.

1.7 PHTHEIROSPERMUM JAPONICUM

P. japonicum is a facultative hemiparasite belonging to Orobanchaceae, which is the same family as the agricultural pest *Striga* ssp and *Orobanche* ssp. It is an autumn-flowering herb native in East Asia without economic relevance, and known as its common name Koshiogama in Japan. *P. japonicum* is associated with herbaceous plants, and there is no scientific report describing the host range of *P. japonicum* in nature. In laboratory, *P. japonicum* is able to form connection in dicot *Arabidopsis,* cowpea and in the monocot *Oryza sativa* and maize, and thus being considered as generalist. Interestingly, this hemiparasite is reported to avoid infection by *S. hermonthica* (Yoshida and Shirasu, 2009), raising the possibility that *P. japonicum* might release haustorium inhibitor(s) into rhizosphere.

Among the advantages of *P. japonicum* is a short life cycle that can be 3-4 month at short-day condition (8 h of light and 16 h of darkness). In addition, it is a self-crossing plant and it is also suitable for artificial crossing, thus amenable for genetic analysis. Other important point is the possibility to perform reverse genetic analysis. I have developed an efficient transformation protocol for *P. japonicum* using *Agrobacterium rhizogenes* to mediated the T-DNA delivery (Ishida et al., 2011).

Figure 1.5. *P. japonicum.* **A-B –** *P. japonicum* in natural environment. **B-** A potential natural host of *P. japonicum*. Plants were collected around Kusatsu area in Gunma prefecture by Ken Shirasu. **C-** *P. japonicum* growing in a green house. **D-** Fruit. **E-** *Arabidopsis* seeds Columbia-0 (left) and *P. japonicum* seeds (right). **F-G** Natural variability of *P. japonicum* collected at different locations in Japan. F- 4 month-old *P. japonicum* collected at Mount Takao (right) by Kohki Yoshimoto with pink flowers, which is taller than the white-flower variety (left) collected at Izu peninsula in Shizuoka prefecture by Takashi Yaeno (G-H). I- 1-month-old plants whose seeds were collected from indicated location in Japan.

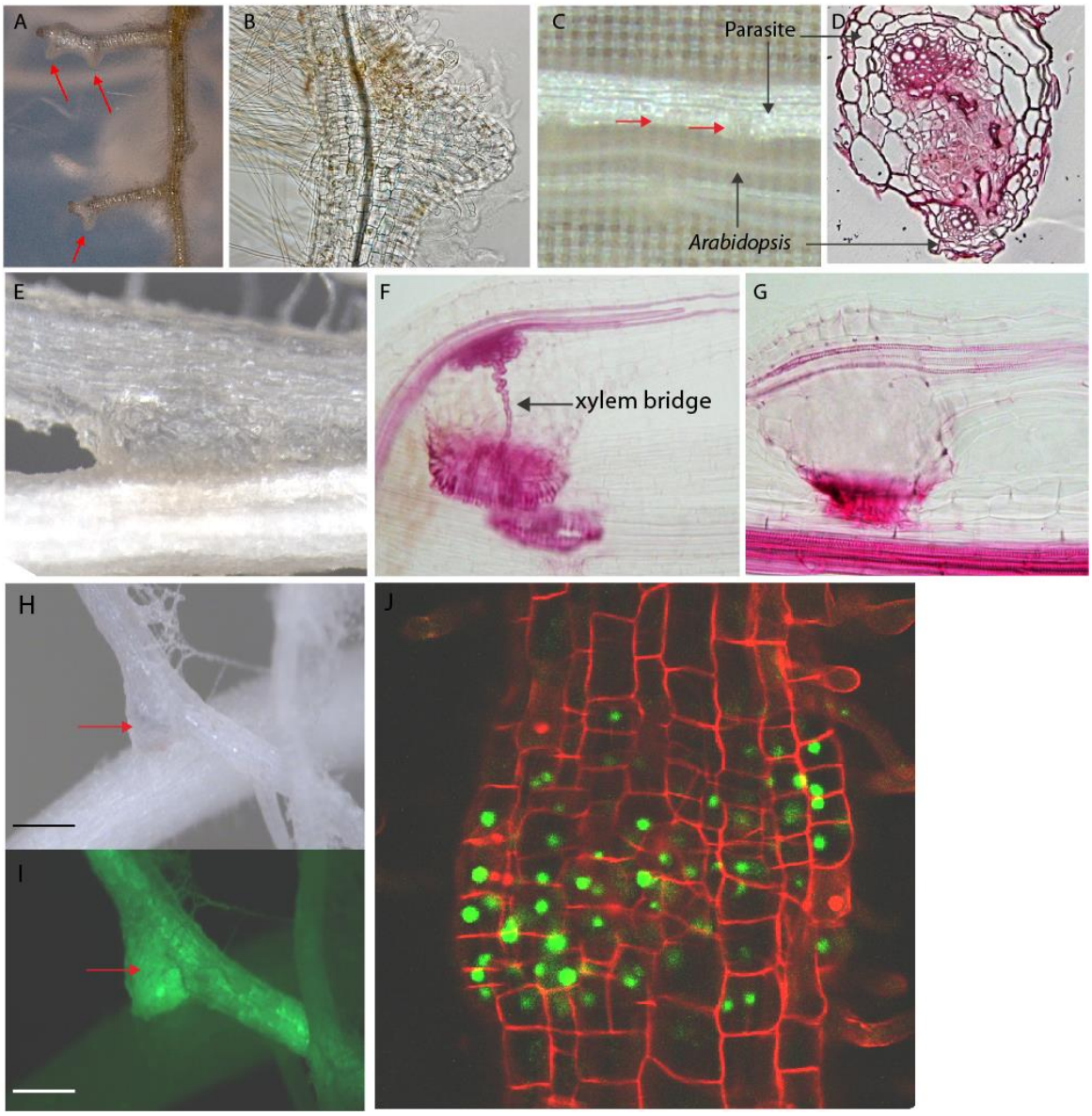


Figure 1.6. Haustoria of *P. japonicum.* A-B – haustoria induced by host exudate. C-D- *P. japonicum* attached to Arabidopsis root. D- Anatomical section stained with safrarin. E-F – Haustoria connected to rice root. F- Safrarin stained root highlighting the xylem bridge. G- Safrarin stained root with a haustorium formed on non-host *L. japonicus* root. No xylem bridge is observed. H-J transgenic root of *P. japonicum*. H-I root transformed with constitutive promoter 35S upstream to *green fluorescent protein* gene. The haustorium formed at rice root under bright field (H) and fluorescent stereoscope (I). J- *P. japonicum* root transformed with construct harboring the promoter of cell cycle marker cycline B1 upstream to *yellow fluorescent protein* (YFP) gene

1.8 OBJECTIVES AND EXPERIMENTAL APPROACHES

The goal of this work is to identify and functionally characterize essential genes for the establishment of plant parasitism. For this purpose, I used the facultative hemiparasitic plant *P. japonicum* since it offers advantages over other parasites, including the suitability for forward and reverse genetic approaches. In addition, *P. japonicum* is a member of Orobanchaceae. This group harbors non-parasitic and parasitic plants with all the degree of host dependency, representing an interesting model to answer biological questions regarding the evolution and specialization of parasites. Moreover, members of Orobanchaceae are most harmful agricultural pests in parasitic plant groups, affecting the food production and crop yields especially in developing countries. In this work, I focused on the critical step to trigger parasitism, the haustorium development. Firstly, a whole root transcriptome of *P. japonicum* was sequenced and *de novo* assembled. It was used as a reference to detect the genes overexpressed in the haustorial tissues at different developmental stages, as described in the chapter 2. Furthermore, *de novo* assembled transcriptome was used as a base to design a custom microarray slide as described in chapter 4, aiming to identify the parasite genes expressed at initial steps of haustorium formation. For this purpose, the *P. japonicum* roots were treated with a chemical able to induce haustorium *in vitro* and microarray analysis was performed. Among the differentially expressed genes, I identified *YUCCA* flavin monooxygenase homolog (*PjYUC3*). *PjYUC3* revealed to be the unique member of *YUCCA* family to have a role in the parasitism. To further characterize it, transgenic hairy roots were made harboring construct of PjYUC3 promoter region upstream to reporter genes. Additionally, I generated overexpression and knock-down lines targeting to *PjYUC3* and their phenotypes were analyzed.

2

CHAPTER

DE NOVO CHARACTERIZATION OF *P. JAPONICUM* TRANSCRIPTOME

2.1 SUMMARY

To find the genes essential for parasitism the transcriptome of *P. japonicum* on parasitic or non-parasitic stages were sequenced. To maximize the gene discovery and aid the *de novo* assembly the sequencing analysis was carried out using two platforms, Roche 454 and Illumina Hi-Seq. The hybrid sequencing strategy resulted in 58,137 unigenes with an average size of 811 bp and N50 of 1,088 bp. Gene Ontology analysis revealed that genes in the category of "structural molecules activity" and "ribosome" cell compartment were overrepresented in the haustoria tissues, suggesting extensive protein synthesis for parasitism. Among differentially regulated 6,143 sequences (fold change with cut-off 2), the SCARESCROW-like (SCL) transcription factor showed higher expression in parasitic stage. In addition, my results showed that the expression profile of quinone reductase 1 (*QR1*), a gene described as necessary for haustorium development in the facultative parasitic plant *Triphysaria vesicolor*, is unaltered by the contact of host or haustorium-inducing chemical. Instead, the related quinone reductase 2 (*QR2*) is up-regulated. Similar results were found in *S. hermonthica*, where *QR2* expression is higher in haustoria tissues compared with *QR1*, which is more expressed in reproductive structures. Taking together, I concluded that the strategy based on hybrid-sequencing plataform increased the quality of transcriptome assembly, enabling the identification of up-regulation of SCL transcription factor and QR2 enzyme in parasitic stage of *P. japonicum*.

2.2 INTRODUCTION

The development of Next Generation Sequencing (NGS) enabled the access of genetic information under a reasonable price. A thousand of sequence bases are generated popularizing genome era for wide range of non-model organisms which was unthinkable few years ago with solely Sanger-based technique. Among NGS are Roche 454 based on Pyrosequencing which the main advantage relies on producing relatively long reads (300-400 bp), the weak point is the few reads and low coverage of sequencing. The Illumina/Solexa use the sequencing based on synthesis which provide a deep coverage and accurate measurements of gene expression, whereas the short reads represents a serious obstacle to assemble sequences from those organisms which the reference genome is not yet available. The other popular techniques are ABI SOLID using sequencing by ligation method and Helicos tSMS with single-molecule sequencing (reviewed in Ekblom and Galindo, 2011). Most recently, the PacBio strategy was developed, producing long reads with good coverage. However, it remains prohibitive expensive for most researchers and produces high frequency of sequencing errors which requires a combined sequencing strategy with Illumina/Solexa reads to correct them (Koren et al., 2012). In this study, I aimed to identify the genes involved in parasitism of *Phtheirospermum japonicum*. For this purpose transcript from haustoria tissues of parasitizing *P. japonicum* and of roots at autotrophic stage were sequenced by the Illumine HiSeq2000 and the Roche 454 pyrosequencer. The *P. japonicum* transcriptome was *de novo* assembled and annotated.

2.3 RESULTS AND DISCUSSION

2.3.1 HIGH-THROUGHPUT SEQUENCING AND *DE NOVO* ASSEMBLY OF *P. JAPONICUM* ROOT TRANSCRIPTOME

To generate the first overview of *P. japonicum* genes involved in the regulation of haustorium development, the whole root transcriptome of *P. japonicum* in the parasitic and non-parasitic stages were sequenced. *P. japonicum* and its host rice were germinated and grown separately for 1 week, and *P. japonicum* seedlings were transferred to rhizotron chambers with (parasitic stage) or without rice (autotrophic stage) for 4 to 6 weeks (Fig. 2.1 A-B). During the first 4-6 weeks various haustoria at the lateral side of *P. japonicum* roots at different developmental stages can be observed (Fig 2.1 C). The newly formed *P.japonicum* roots frequently show immature haustoria. In this case the swollen root is clearly visible in the parasite roots, but the connection to host vascular system is not yet formed. Consequently, the attached haustorium is easily loosened from the host root (Fig 2.1 D). In contrast, a haustorium in older roots firmly attach to the host tissues (Fig 2.1 E). To sequence samples from distinct parasitic stages, a pool of haustoria tissues at different stages was carefully collected. When a haustorium was tightly attached to the host root, haustorium region with both host and parasite tissues were sampled with care to avoid host tissues as much as possible (Fig 2.1 F).

To improve *de novo* assembly of *P. japonicum* transcriptone two distinct sequencing platforms were adopted, Hiseq2000 (Illumina) and GS-FLX pyrosequencer with FLX titanium reagents (454 Life Sciences). Total RNAs were extracted from autotrophic and parasitic stages of *P. japonicum* and Illumina libraries were generated respectively. For the 454 pyrosequencing platform, a normalized library was constructed with mRNAs isolated from each sample and mixed in an equal ratio. The 454 pyrosequencer produces longer reads (up to 400 bp in average) but with less read numbers comparing to Hiseq2000. Thus, a normalized library covers as much variety of transcripts as possible to aid *de novo* assembly of *P. japonicum* transcriptome. Approximately 24 million paired-end reads were generated by Hiseq2000, while 888,638 reads with average length of 333.8 bp were generated by the 454 sequencer. In total about 4.5 Gbp were sequenced by Hiseq2000 and 296 Mbp by 454. The reads from parasitic tissues were mapped against the available rice full length transcripts to remove reads derived from host tissues. The unmapped reads and the reads from autotrophic stage libraries were assembled using CLC Genomics Workbench software (ver. 4.8) followed by CAP3 assembler (Huang and Madan, 1999) to reduce redundancy. From the resulting non-

redundant sequences (unigenes), the sequences shorter than 300 bp were removed as possible unreliable sequences.

Assembled yielded 42 Mbp residues distributed in 58,137 unigenes with length ranged from 300 bp to 12,710 bp, with an average size of 811 bp, median of 522 bp and N50 of 1,088. Among unigenes, 45,323 (78%) include a putative coding region detected by ESTScan (ver. 2.0) and 19,036 (38.75%) have the predicted start codon. The local alignment using the BLAST algorithm shows that 29,767 (51.2%) of *P. japonicum* sequences have homologous sequences in *Arabidopsis* genome (Table S2). These numbers were higher than other published *de novo* dicot transcriptome which adopted a single sequencing platform. For example, the chickpea *de novo* assembly based solely on 454 reads gene yielded 44.7% of the unigenes with a valid BLAST hits. Using a high-throughput Illumina platform, 43.3% of 55,088 unigenes from *Camellia sinnensis* (Shi et al., 2011) and 59.65% of 57,050 sequences of *Litchi chinensis* (Li et al., 2013) associated with known genes in public databases, respectively. To investigate if the rice sequences were efficiently removed, the parasite transcriptome was BLAST-aligned against the rice genome, applying strict parameters (e-value less than $1e^{-30}$ and 100% of alignment score) I could still identified 155 rice sequences among 58,137 unigenes, suggesting that in the newly assembled *P. japonicum* transcriptome still remains around 2.6% of host sequence contamination.

Next I assessed the coverage rate of the *P. japonicum* assembly in the entire transcriptome. The ultraconserved orthologs (UCOs) are 357 highly conserved single-copy genes detected in representative eukaryotes, *Arabidopsis thaliana*, humans, mice, yeast, fruit flies and *Caenorhabditis elegans* (http://compgenomics.ucdavis.edu/compositae_reference). All UCOs were detected by BLAST in the *P. japonicum* transcriptome, but only 24 (6.72%) showed a complete coding sequence. Next, I assessed the assembly with the APVO (*Arabidopsis thaliana*, *Populus trichocarpa*, *Vitis vinifera* and *Oryza sativa*) sequences, which are 959 single copy genes shared with the four plant genomes (Duarte et al., 2010). The PlantTribes 2.0 program (Wall et al., 2008) detected 918 (95.7%) APVO homologs among *P. japonicum* unigenes. Based on these analyses, we estimated that the transcriptome has covered more than 95% of the whole *P. japonicum* transcripts. Thus the *de novo* assembled transcriptome presents a reasonable coverage of *P. japonicum* genes with enrichment of transcripts expressed at different stages of haustorium development, representing thereby an excellent basis to mine essential genes for the haustorium development.

2.3.2 FUNCTIONAL ANNOTATION OF *P. JAPONICUM* TRANSCRIPTOME

The transcriptome data was analyzed using BLASTX against non-redundant databases to annotate with functional homologs. Out of 58,137 sequences, 36,151 (62.18%) matched to known genes (Table 1) and, among them, 9,830 were assigned as expressed sequences with unavailable function. In addition, the annotation analysis showed that 21,986 remained as orphan sequences, which means that there were no transcript in the searched database that matched to them under the adopted parameters described in Material and Methods chapter (Chapter 7). To verify whether these orphans were misassembled sequences I analyzed their transcript levels, 5,388 showed an expression value (measured as RPKM *R*ead *P*er *K*ilo base *M*illion) higher than 5.0 and 3,328 were at least 2 times higher expressed in haustoria tissues compare with non-interacting *P. japonicum* roots, used as a control. This result indicated that these orphans found in *P. japonicum* transcriptome most likely corresponded to expressed sequences and at least 3,328 of them had their expression activated during the parasitism, suggesting the involvement of large amount of unknown genes recruited by the parasite.

The predicted gene product of each unigene was also classified into Tribes and Orthogroup based on cluster analyses of the inferred proteomes of ten sequenced angiosperms (*Arabidospsis thaliana*, *Carica papaya, Populus trichocarpa, Medicago truncatula, Oryza sativa, Sorghum bicolor,* the fern *Selaginella moellendorfii,* the moss *Physcomitrella patens* and the green alga *Chlamydomonas reinhardtii*) as described in (Wickett et al., 2011). Tribes are similar to gene families, while the orthogroups are narrower than Tribes and corresponding to genes with monophyletic origin in the studied species. The 36,151 sequences were assigned into 8,461 different tribes, with 5,166 Tribes containing more than one associated sequences, while 3,295 tribes were represented by single sequence. 12,663 orthogroups were captured by this analysis.

To identify the genes relevant for establishment of parasitism, the sequences extracted from haustoria tissues were compared with those from roots growing in autotrophic mode. A total of 4,809 unigenes were exclusively expressed in haustoria, whereas 1,426 were expressed solely in autotrophic stage. From all the sequences, 50,730 were expressed in both libraries (Fig. 2.2). For comprehensive understanding of the gene functions, which are modulated during the parasitism, gene ontology (GO) analysis was performed. Among all the unigenes, 33,159 (57.0%) were assigned in at least one GO Slim terms in three GO domains; Molecular Function, Cellular Component and Biological Process. The GO analysis revealed that

the terms “structural molecule activity” in Molecular function category and “ribosome” in the Cellular Component were over represented in genes differentially regulated in haustoria tissues (Fig. 2.3). About 81.2% of sequences classified as “structural molecule activity” encode ribosomal proteins, which are structural compounds of ribosomes. Riopel and Timko (1995) observed that haustorium formation requires continues protein synthesis; when protein synthesis is blocked by inhibitors or amino acid analogues, haustorium development stops at intermediate stages in *Agalinis purpurea*. The protein synthesis inhibitor also affects the parasitic plant response to DMBQ. Therefore, our results together with previous observations suggest that the protein turnover is important for the completion of haustorium development.

It is worth noting that there are considerable numbers of the transcripts without assignment to GO category, and the number is higher in genes up-regulated in parasitic stages. In the Molecular function category around 3,670 (59.7%) of differentially expressed genes in the haustoria tissues and 25,907 (44.5%) of total unigenes were not assigned to any GO term. The over-representation of no-assigned genes were also described in the host-parasite interface between the parasitic plant *T. vesicolor* and their hosts (Honaas et al., 2013). Thus the large-scale transcriptome of *P. japonicum* (this study) and *T. vesicolor* (Honaas et al., 2013) revealed that the function of marjority of genes recruited by the parasite for the host invasion still remain elusive.

2.4 IDENTIFICATION OF TRANSCRIPTIONAL FACTORS FAMILIES DIFFERENTIALLY EXPRESSED GENES DURING PARASITISM

Investigation of *P. japonicum* transcriptome reveals that exchanging from autotrophy to dependency on host plants (parasitism) correlates with up regulated expression of certain set of transcriptional factors. The list of transcription factors with a fold change higher than 3.0 in parasitic stage library are shown on Table 3.2. In total there were ten transcription factor families positively regulated in the haustoria tissues (GRAS protein, MYB, MADS-box, Aux/IAA, bHLH, bZIP, C2H2 HSF, NAC and WRKY). Interestingly, the group in GRAS protein family with higher expression belongs to SCARECROW-like transcription factors (SCL). Previous results showed that SCARECROW and SHORT-ROOT, another GRAS-type transcription factor, play central role for the establishment and maintenance of root meristem architectures through regulation of division in initial cells and around the quiescent center (Sabatini et al., 2003) (Koizumi et al., 2012). Our data suggest that SCL might have a role during the parasitism. Its biological function in *A. thaliana* in controlling the cell division might imply similar function in parasitic plants. Consistently, haustorium development is characterized by the extensive cell division in the root cortex and epidermis, which is evidenced in transformed *P. japonicum* root carrying the Cyclin B1 promoter fused to YFP with nuclear signal (Ishida et al., 2011). Thus, I speculated that the SCARECROW-like transcription factors is up regulated in haustoria tissues as it may be required to stimulate the localized cell division in cortical and epidermal cells after the haustorium initiation being provoked by host root.

2.5 *QUINONE REDUCTASE 2* GENE RESPONDS TO HOST-PARASITE INTERACTION IN *P. JAPONICUM* AND *S. HERMONTHICA*

Gene encoding the NAD(P)-H-dependent quinone reductase in *T. vesicolor* (*TvQR1*) was reported to play an essential role for haustorium formation (Bandaranayake et al., 2010). To investigate whether this gene is similarly important in *P. japonicum* we searched its homologous genes of *TvQR1* using BLAST program from the *P. japonicum* transcriptome. For comparison, we also identified the *P. japonicum* homolog of *TvQR2*. The *TvQR1* gene is highly expressed in parasitic plants after treatment with haustoria-inducing factors (HIF) or host roots, whereas expressions of TvQR2 is induced after exposure to haustorium-inducer chemical DMBQ, but not host roots. In addition, Bandaranayake et al. (2010) showed that the silencing of *TvQR1* provokes reduction of haustorium initiation in *T. vesicolor*, while the *TvQR2*-silenced roots had similar haustorium development compared with control. Thus it is suggested that

TvQR2 is not involved in the parasitism, instead it was associated with plant tissue protection by detoxifying quinones in the rhizosphere.

The predicted protein of PjQR1 (329 aa) and PjQR2 (205 aa) shows identical length compare with TvQR1 (329 aa) and TvQR2 (205 aa) proteins, respectively. The phylogenetic tree based on amino acid alignment of quinone reductase from *P. japonicum*, *T. vesicolor*, *S. hermonthica* and *S. asiatica* shows the sequence similarities between single-electron (QR1) and two-electron reducing enzyme (QR2) (Fig. 2.4). Surprisingly, the similarities found in protein sequences of PjQR1 and TvQR1 contrast with their differences in gene expression patterns during the parasitism. The contig corresponding to *PjQR2* is 3 times higher in parasitic stage compared with autotrophic stage, whereas the expression level of *PjQR1* kept unaltered in both growth conditions (Fig. 2.5A). The expression profile of *PjQR1* and *PjQR2* were also measured in roots treated with DMBQ by qRT-PCR (Fig. 2.5B). Similar to RNA-Seq data the expression profile of *PjQR1* gene is not differentially regulated in roots treated with DMBQ. In contrast *PjQR2* gene is highly expressed specially at the first 24 hours of the chemical treatment (Fig. 2.5B). Therefore our results show that the *PjQR2* is highly expressed in parasite roots challenged directly by host or under DMBQ treatment in contrast to unresponsive expression profile of PjQR1. Thus, our data suggest a distinct modulation of *QR2* and *QR1* might reflect an altered role of these genes during the haustorium initiation in *P. japonicum* roots compare with *T. vesicolor*.

To further investigate the role of quinone oxidoreductase in the root parasitism in Orobanchaceae we also searched for *QR1* and *QR2* homologs in *Striga hermonthica*, one of the most dangerous parasitic plants in Orobanchaceae family. First *ShQR1* and *ShQR2* were identified using the blast tool available in Parasitic Plant Genome Project webpage (http://ppgp.huck.psu.edu). The full length of ShQR1 and ShQR2 proteins show identities of 87% and 92% in amino acid level compared with TvQR1 and TvQR2, respectively. To estimate their expression values we used the reads from various libraries available at Parasitic Plant Genome Project website constructed based on different development stages of *S. hermonthica*. The reads of each library were mapped onto coding sequence of *ShQR2* and *ShQR1*. Similar to what was observed in *P. japonicum ShQR2* gene displayed higher expression levels compared with *ShQR1* in libraries derived from tissues where the haustoria were present. In the pre-attachment stage which contains reads from haustoria induced by HIF, the average of RPKM values for *ShQR2* gene was 0.8 million, while for *ShQR1* was 0.1 million (Fig. 2.6). In the early post-attachment (48h after attach to the host) this difference was higher, 0.88 and 0.087 for

ShQR2 and *ShQR1*, respectively. The unique situation where *ShQR1* shown higher expression compared with *ShQR2* was in reproductive structures, the transcript of both genes were not detected in the vegetative organs in late post-attachment stages or in the seed germination (Fig. 3.11). Thus, the lack of regulation in transcript levels of *QR1* in *P. japonicum* and in *S. hermonthica* either in contact with HIF or in presence of host suggests that this gene may not play an essential role to induce the haustorium in all parasitic plants in Orobanchaceae as was initially proposed.

2.6 CONCLUSION

In this work we provided the first genetic information of the emerging model plant in parasitic plant studies *P. japonicum*. The main conclusions are:

(1) The hybrid sequencing platform resulted in representative *de novo* assembled transcriptome of *P. japonicum.*

(2) Gene ontology enrichment analysis revealed that genes related with protein synthesis are preferentially up-regulated in the haustoria tissues and the data corroborates with previous studies showing the importance of continue protein synthesis to complete the haustorium development.

(3) A member of SCARECROW-like transcription factor family is activated during the parasitism

(4) Parasitism in *P. japonicum* and *S. hermonthica* roots recruited the gene encoding Quinone Reductase 2, which differs from *T. vesicolor* where the Quinone Reductase 1 is essential for haustorium development.

2.7 TABLE AND FIGURES

Table 2.1 – Summary of *de novo* assembly of *P. japonicum* transcriptome

	Numbers
Number of contigs	58,137
Residue counts	47,149,654
Min length	300 bp
Max length	12,710 bp
Average of length	811.01 bp
N50	1,088
Protein coding	45,324 (78%)
Blast hit	36,151 (62.18%)
Blast hits to Arabidopsis*	29,767 (51.2%)
Rice**	155

*** e-value less than $1e^{-10}$**
*** e-value less than $1e^{-30}$ and 100% identity in local alignments**

Table 2.2 – Transcription factor families differentially expressed during parasitism

Transcription Factor Family	Feature ID	log_2 (Fold Change)	Best_Blast_hit	e-value*
GRAS (SCARECROW-like proteins)	Pj3contig_35655	20.30	gnl|Carpa1|supercontig_81.122	1E-62
	Pj3contig_1611	19.32	gnl|Vitvi1|GSVIVP00021332001	8E-24
	Pj3contig_45282	14.77	gnl|Vitvi1|GSVIVP00021332001	2E-41
	Pj3contig_6135	6.79	gnl|Vitvi1|GSVIVP00021332001	3E-30
	Pj3contig_24348	5.75	gnl|Poptr1|417914	3E-61
	Pj3contig_24728	5.10	gnl|Poptr1|203692	4E-17
	Pj3contig_55282	4.57	gnl|Poptr1|172616	2E-97
	Pj3contig_7034	4.08	gnl|Vitvi1|GSVIVP00024294001	3E-55
	Pj3contig_33862	3.04	gnl|Vitvi1|GSVIVP00018466001	3E-16
MYB	Pj3contig_22975	7.62	gnl|Carpa1|supercontig_798.1	1E-56
	Pj3contig_53222	6.58	gnl|Orysa5|Os01g74410	2E-40
	Pj3contig_27606	3.55	gnl|Vitvi1|GSVIVP00020267001	1E-20
MADS-box	Pj3contig_19957	5.77	gnl|Vitvi1|GSVIVP00008566001	3E-77
	Pj3contig_49486	5.52	gnl|Poptr1|660583	8E-95
AUX/IAA	Pj3contig_16742	3.08	gnl|Poptr1|643213	3E-80
bHLH	Pj3contig_9151	3.96	gnl|Vitvi1|GSVIVP00029322001	4E-54
	Pj3contig_27317	3.89	gnl|Poptr1|771962	2E-22
	Pj3contig_50777	3.37	gnl|Poptr1|754916	4E-41
bZIP	Pj3contig_19048	3.40	gnl|Carpa1|supercontig_129.74	1E-24
C2H2	Pj3contig_28192	7.56	gnl|Vitvi1|GSVIVP00025561001	3E-24
HSF	Pj3contig_27751	16.27	gnl|Poptr1|717386	2E-57
NAC	Pj3contig_17979	3.14	gnl|Poptr1|760326	3E-81
WRKY	Pj3contig_27251	17.89	gnl|Poptr1|816178	7E-40

*Cut-off of e-value < $1e^{-15}$; Log_2(fold change) > 3.0

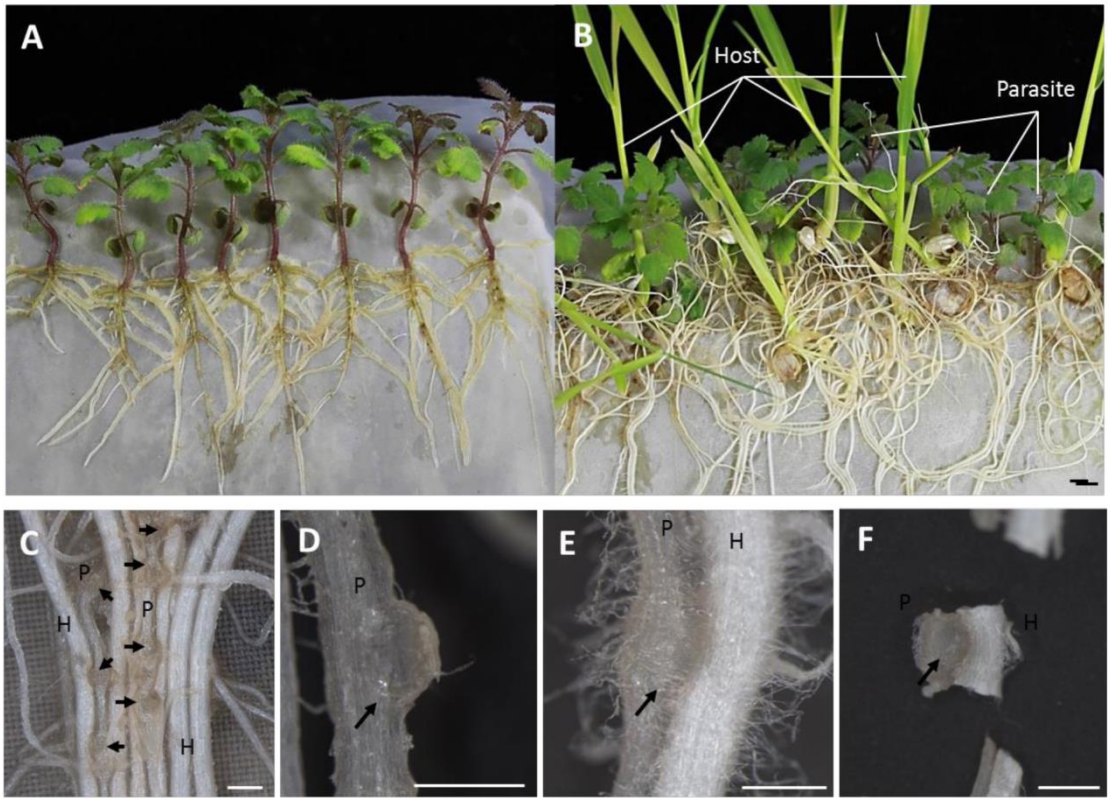


Figure 2.1 – Root tissue of 4 week-old *P. japonicum* used for the construction of RNA-seq libraries. (**A**) Plant material used for autotrophic stage library. Whole parasite root growing in rhizotron chambers without present of the host. (**B** - **F**) Haustoria formed in parasite root in contact with rice plant were collected in different developmental phases for parasitic stage library. (B) Parasite growing together with host. (C) Haustoria formed along the parasite root. (D) Example of sampled haustorium in immature phase where it was easily disconnected from host. (E - F) Haustorium in the later stage, when parasites are already tightly connected to the host root, in this case both parasite and host roots were sampled together. Bars correspond to 0.5 mm.

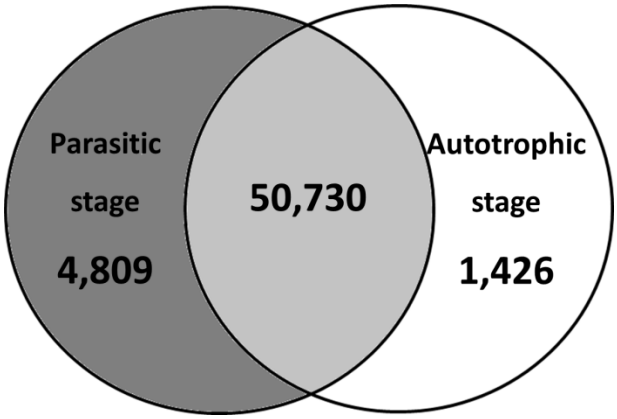


Figure 2.2 – Venn diagram showing the number of sequences expressed in parasitic and autotrophic stages. Expression levels of transcript in *P. japonicum* transcriptome were calculated based on RPKM (reads per kilobase per million). The number of sequences exclusively expressed in the haustorial tissues is colored in gray.

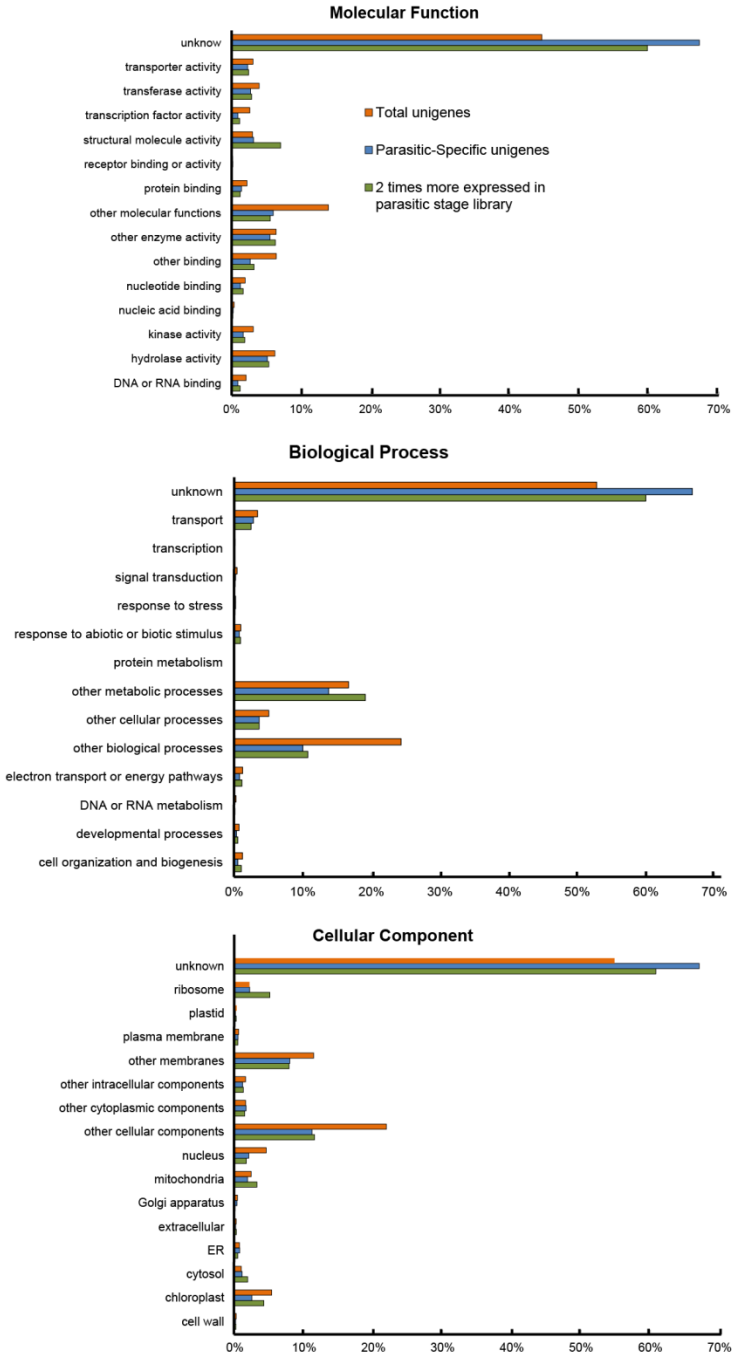


Figure 2.3 – Summary of Gene ontology analysis. The *P. japonicum* sequences were assigned into GO Slim categories according to their putative molecular function (A), cellular component (B) or predicted biological process (C). The graph compared sequences two times more expressed (green) in the haustoria tissues with total assignments in *P. japonicum* transcriptome (orange). The fold change for each gene was calculated through the ratio between $\log_2^{(1+RPKM)}$ expression values in parasitic stage and $\log_2^{(1+RPKM)}$ from autotrophic stage library.

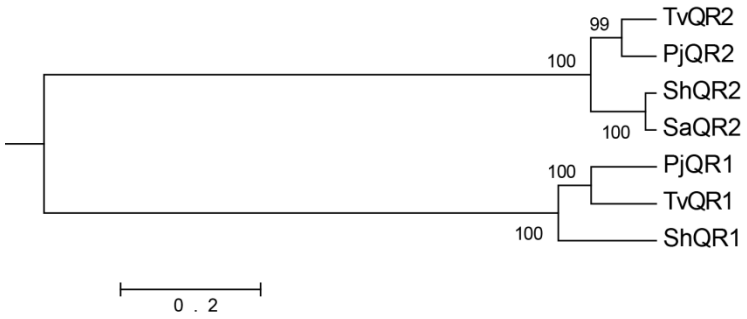


Figure 2.4 –Phylogenetic relationship of quinone oxidoreductase. UPGMA tree based on full length protein alignment of quinone oxidoreductase from *P. japonicum* (Pj), *Striga hermonthica* (Sh), *Striga asiatica* (Sa) and *Triphysaria vesicolor* (Tv). Quinone Oxidoreductase from *Striga asiatica* is available at Genbank under the code DQ442405.1

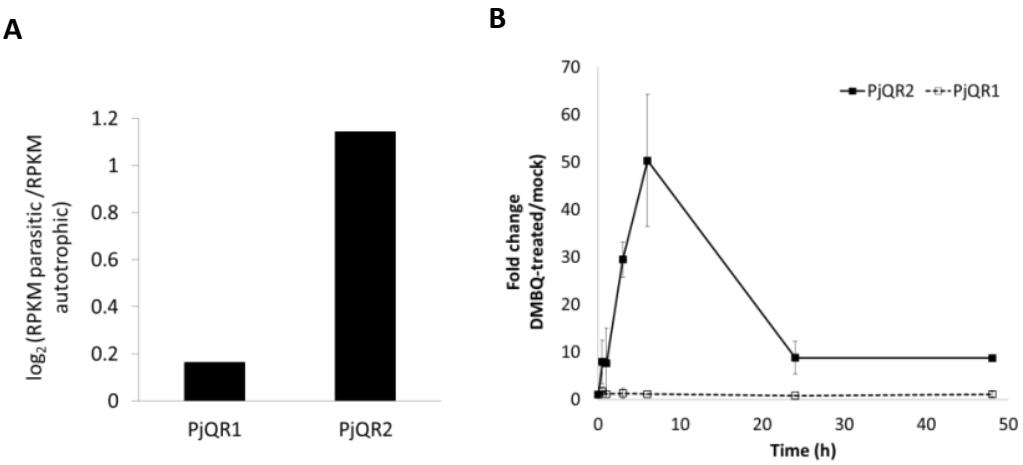


Figure 2.5 – Expression of *PjQR2* and *PjQR1* genes in parasite roots during parasitism. (**A**) Transcript levels determined as RPKM from RNA-Seq analysis represented as $\log_2{}^{(\text{Fold change})}$. Fold change means the difference between RPKM of parasitic stage land RPKM of autotrophic stage library. (**B**) qRT-PCR analysis comparing the expression of *PjQR2* and *PjQR1* during the haustoria development induced by the chemical DMBQ. Treated roots were collected at 0, 0.5, 1, 3, 6, 24, 48 h after the addition of chemical and the expression levels were compared with mock-treated roots collected at same time. The error bars correspond to two biological replicates with three technical replications each.

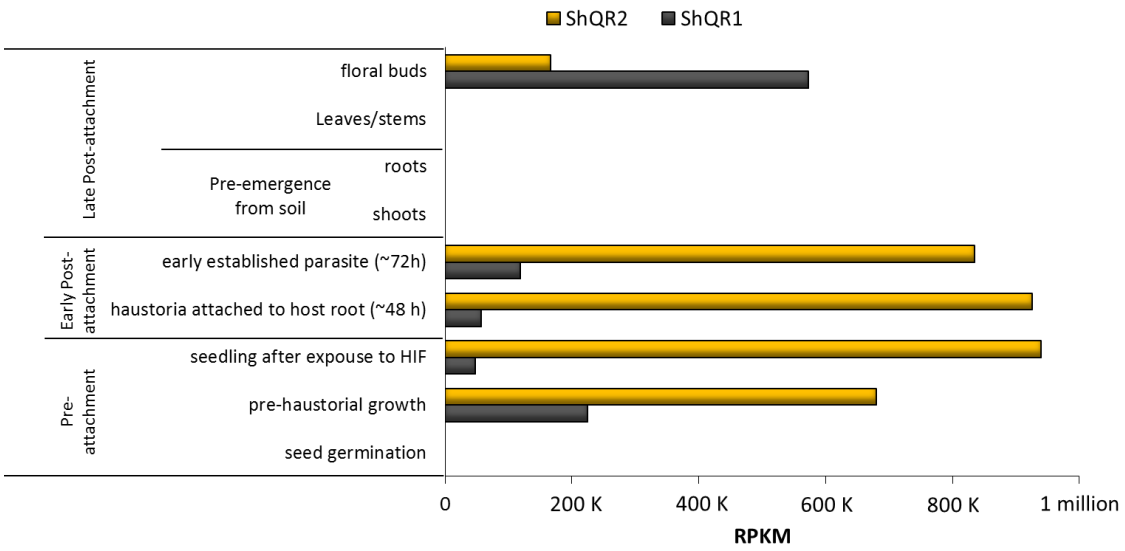


Figure 2.6 – Expression of *ShQR2* and *ShQR1* genes in different development stages in *S. hermonthica* life cycle. Transcript levels were calculated based on reads of parasitic plant genome project mapped onto full length transcript of *ShQR1* and *ShQR2.* The expression levels are shown as $Log_2^{(1+RPKM)}$. The reads derived from plant tissues collected at different developmental stages of *S. hermonthica* cell cycle

3

CHAPTER

REQUIREMENT OF SUBTILASES FOR INTERACTION BETWEEN PARASITIC PLANTS AND A COMPATIBLE HOST

3.1 SUMMARY

RNA-Seq analysis indicated that transcriptd assigned as as subtilases (SBT) were highly expressed in parasitic stage. To investigate the role of these proteases in the plant parasitism, their expression patterns were monitored by qRT-PCR during the interaction of the parasite *P. japonicum* with a host (*O. sativa*) or a nonhost (*L. japonicus*) plant for 1, 2, 3 and 7 days of infection. The expression of five *SBT* genes (*PjSBT2,-4,-7,-8 and -11*) were detected only at 7 days after the interaction with the host, while no expression of these genes was detected in contact with nonhost *L. japonicus*. Morphological studies with of a haustorium stained by *Safranin-O* revealed that the formation of a vascular bridge, which connects xylem vessels between a host and a parasite, occurs after 7 days of interaction. Promoter analysis of one of these *PjSBT* genes showed that its expression was localized at the interacting site in haustoria with the host at penetrating stage. These parasitism-induced *PjSBTs* are homologs of *Arabidopsis AtSBT* group 5, symbiosis-induced *L. japonicus SbtS,* and *Arabidopsis AIR3* (At2g04160) involved with lateral root formation, indicating a possible shared mechanism in haustorium, nodulation and lateral root formation. In addition, the parasitism-induced *ShSBT1*, an *S. hermonthica* gene phylogenetically distant from *PjSBTs* of group 5 suggested that this subtilase may have had acquired a distinct function during the development of terminal haustoria.

3.2 INTRODUCTION

Subtilisin-like serine proteases (Subtilases) contains three conserved amino acids aspartate, histidine and serine, in the catalytic site (Dodson and Wlodawer, 1998). According to the MEROPS database (Rawlings et al., 2012), Subtilases, or often called S8 family peptidases, are divided into two subfamilies, S8A and S8B (Table 3.1). The S8B, also known as Kexin, comprises eukaryotic enzymes common in fungi and mammals. In contrast, within the S8A group there are members in prokaryote and eukaryotic organisms, including pyrolysin subtilase, the unique type of subtilase found in plant genomes (Schaller et al., 2012). In view of simplification, the plant pyrolysin-related subtilase is referred in this thesis just as subtilase (SBT). SBT is composed of signal peptide, prodomain, catalytic, PA and Fn-III domains. It is translated as pre-pro-protein with a signal peptide at the N-terminal end targeting to the apoplastic region. To be translocated to outside of plant cell, a maturation step, in which the prodomain is cleaved, is required. The structure of a SBT from tomato (SlSBT3) revealed that this enzyme is highly stable even at high temperature or in alkaline conditions. Based on the structural information, it was also shown that the PA domain is essential for a Calcium-independent homodimerization (Ottmann et al., 2009).

SBT proteins comprise one of the largest families in plants. The *Arabidopsis* genome contains 56, rice 64 and grape 113 genes (Table 3.3) encoding SBTs. Analysis of *P. patens* genome revealed that *SBT* appeared in the genome of land plants through a single event of horizontal gene transfer from a bacterial subtilase (Yue et al., 2012), followed by rapid duplication events. This finding corroborates with the lack of other types of subtilases in plants and the fact that SBT is closer to bacterial subtilases than those found in animals or yeast (Schaller et al., 2012). In addition, the physiological functions of SBTs have been associated with morphology, physiology, response to drought, stomatal density or tolerance to biotic stress such as defense against pathogens (Rautengarten et al., 2005; Budič et al., 2013; Tanaka et al., 2001; von Groll, 2002; Ramírez et al., 2013), suggesting the acquisition of novel roles that allowed plants to leave the aquatic life to conquer the terrestrial environment (Yue et al., 2012).

Arabidopsis subtilases have been classified into six groups based on their amino acid sequences. The well characterized members of the group 1 are At1g04110 (SDD1) which negatively controls the stomatal pattern (von Groll, 2002), and SbtM1/3 as well as P69B/C which are required for interaction with other organisms. *SbtM1/3* are upregulated during the invasion of arbuscular mycorrhizal fungi (Takeda et al., 2009), while P69B/C upon pathogen

attack (Jordá and Vera, 2000). Another interesting enzyme in the group 1 is Phytaspases which is involved in the programmed cell death and assigned as a plant version of mammal caspase (Vartapetian et al., 2011). A well-studied member in the group 2 is At1g04110 (ALE1) required for development of cuticle and epidermal layers in *Arabidopsis* embryo. In the group 3, the overexpression of *Arabidopsis SBT3.3* intensifies the innate immune response mediated by SA and chromatin remodeling (Ramírez et al., 2013). Other SBT involved with plant defense response and closer to group 5 is Gm-1. This enzyme described in *Glycine max* suffered an internal processing step, releasing a 12 aa peptide (GmSubPep) that activates the expression of defense-related genes (Pearce et al., 2010). The other *Arabidopsis* SBT, At2g04160 (AIR3) from this group is associated with promotion of lateral roots and is expressed in xylem vessels (Zhao et al., 2005). In the group 4, Cucumisin is a representative essential for fruit maturation in melon (Cucumis melo L) (Nakagawa et al., 2010). The group 6 compromised only two pyrolysins subtilases (At4g20850 and At5g19660) which the sequences are phylogenetic distant from the others subtilase groups, their sequences are closer to the mammalian than any other bacterial homologs (Schaller et al., 2012). The gene At5g19660 function in *Arabidopsis* is related with accumulation of unfolded protein and it is responsive to salt stress. The gene At4g20850 encodes the only exopeptidase among subtilases and it acts in degradation of oxidized proteins caused by metal exposure (Schaller et al., 2012).

In this chapter the role of *P. japonicum* subtilases during the parasitism was investigated. The functional analyses indicate that these enzymes are recruited for establishment of compatible interaction between the parasite and its host. Promoter analysis of one of haustoria-induced *SBTs* showed specific expression potentially at interface between the parasite and the host. The phylogenetic tree places the haustoria-induced SBTs closer to the *Arabidopsis* group 5 SBTs. Interestingly, this clade includes AIR3 involved in lateral root promotion and SbtS expressed during the arbuscular mycorrhizal and rhizobia-mediated nodulation, suggesting similarities among these developmental processes. To verify if the role of subtilases is conserved in other parasitic plants, three genes from *S. hermonthica* were also isolated. Homologs of parasitic-specific *PjSBT* showed expression profile not associated with parasitism, indicating that their functions were not conserved among parasitic plants. Interestingly, *ShSBT1* a subtilase phylogenetically distant from those enzymes from group 5 was specifically upregulated during the interaction with a compatible host which suggested that this subtilase may have had acquired a distinct function compared with those *PjSBTs* of group 5 during the development of terminal haustoria.

3.3 RESULTS AND DISCUSSION

To characterize genes recruited for establishment of plant parasitism, we analyzed the 4,809 genes exclusively expressed in the parasitic stage. The list of the highly expressed genes and their expression values are shown on Table 3.3. Interestingly, seven out of ten highest expressed genes encode subtilisin-like proteases or subtilases (PjSBT). Sequence analysis of these PjSBTs showed that they have a large number of homologs in land plants, but not in the green algae *Chalamydomonas reinhardtii*, potentially hinting functional roles of SBTs during the plant adaptation in terrestrial environments. To investigate if SBTs acquired a function associated with plant parasitism, the full length sequence of seven genes from *P. japonicum* were cloned and coding region confirmed by Sanger sequencing. *PjSBTs* show in average a a putative amino acid region around 700 aa, with the exception of *PjSBT6,* which a has coding sequence of 3096 bp (1032 aa). The RNA-Seq contigs assigned for *PjSBT2,-4,-7,-8 and -11* showed specific expression in haustorial tissues with a high fold change compared with the control (Table 3.4). *PjSBT3* and *PjSBT6* were not differentially expressed in the parasitic tissues. Therefore, RNA-seq data suggested a possible involvement of subtilases in plant parasitism.

To further analyze the involvement of *PjSBTs* for parasitic infection, the expression of *PjSBTs* were quantified in *P. japonicum* roots infecting a host (rice) or a non-host plant (*L. japonicus*). *Lotus* is a non-host for *P. japonicum* and the parasite fails to penetrate into its stele (Fig. 3.1 A and C). Resistant phenotype of *L. japonicus* against *P. japonicum* is featured by the deposition of dark materials at the interface of interaction (Fig. 3.1 A). In contrast, the interaction of *P. japonicum* with rice is marked by the absence of this barrier and the parasite successfully reaches the rice xylem vessels around 1 week of interaction (Fig. 3.1B and D). Expression of *PjSBT2,-4,-7,-8 and -11* occurred exclusively at 7 days of the rice - *P. japonicum* interaction. However, when the parasite is associated with the non-host *L. japnonicus*, of their expression were under the detection limit (Fig. 3.2). The transcript of *PjSBT3* was mainly observed in non-interacting roots, and *PjSBT3* and *PjSBT6* were not specifically expressed neither in lotus-parasitizing nor in rice-parasitizing tissues (Fig. 3.2). To confirm that the expression of *PjSBT2,-4,-7,-8 and -11* are not from unspecific amplification of rice cDNAs, cDNAs originated from non-infecting root tissues and genomic DNAs of rice and lotus were used as template in a PCR reaction with the *PjSBT* primer pairs. No DNA amplification was observed in rice and lotus cDNAs nor genomic DNAs (Fig 3.3). Expression analysis in host and non-host suggested that at least five *SBT* genes of *P. japonicum* are potentially involved for compatible interaction between the parasite and its host.

To verify the specific localization of *SBTs*, the promoter of *PjSBT8* was isolated and cloned upstream to *Green Fluorescent Protein* (*GFP*) and *β–glucuronidase* (*GUS*) (*PjSBT8$_{pro}$::GFP/GUS*) into a modified version of R4L1pGW532 plasmid (Nakamura et al., 2009). I introduced into the original R4L1pGW532 the *35S* promoter fused to *Red Fluorescent Protein* (*RFP*) (R4L1-35SRFP) as a transformation marker. The modified R4L1-35SRFP has an advantage to allow the visual selection of transformed tissues without antibiotic selection the map R4L1-35SRFP is shown in Appendix section (Chapter 9). The transgenic *P. japonicum* roots carrying *PjSBT8$_{pro}$::GFP/GUS* parasitizing Arabidopsis showed that expression of *GFP* is detected apparently at the most external parasite cells which were already penetrating into host tissue (Fig. 3.4). In contrast, in the immature haustoria, where the parasite invasion was yet not completed, the expression of *GFP* was not detected. These observation at the tissue level were consistent with the quantification of *PjSBT8* transcripts only 7 days after the host-parasite interaction. In this stage, the formation of the bridge that connects the parasite to host xylem was observed as shown in Fig. 3.1-D. Note that such bridge structure is not developed when the parasite is interacting with nonhost such as *L. lotus* (Fig. 3.1-C). The results raised a possibility of SBT involvement in the successful establishment of xylem connection between the parasite and the host.

Phylogenetic analysis with putative sequence proteins encoded by seven *PjSBT* genes (*PjSBT2,-4,-7,-8,-11,-3 and -6*) from *P. japonicum* and three genes from the parasitic plant *S. hermonthica* (*ShSBT1*, *-4*, *-8*) was carried out by comparing with homologs from *Arabidopsis*, *Lotus*, *Glycine Max,* tomato and *Nicotiana tabacum*. The parasite-induced PjSBTs were clustered together with the *Arabidopsis* group 5 SBTs. The group 5 was splitted into other two subgroups defined in this thesis as subgroup 5a and 5b. Thus PjSBT8 was in the subgroup 5b and *PjSBT2,-4,-7,* and *-11* placed in subgroup 5a. PjSBT3 and PjSBT6, which were not induced in parasitism, were placed in the group 3 and 6, respectively. Thus, it is possible that the group 5 SBT from *P. japonicum* may have acquired a function associated with successful connection between a parasite and a host. The alignment of the group 5 SBTs of parasites and enzymes from other plants is shown in Fig. 3.6. Among SBTs belonged to group 5 are soybean *Gm-1*, which the processed-peptide activates the plant immune response; *SbtS* from Lotus associated with symbiont interactions (Kistner et al., 2005; Takeda et al., 2009) and the *SBT* gene *AIR3* upregulated during lateral root formation (Neuteboom, 1999). The *AtSBT5.2* which is closest to *PjSBT2, -4, -7* and *-11*, has its expressed limited to xylem vessels of *Arabidopsis* (Zhao et al., 2005), consistently with the observation that parasitism-induced *PjSBTs* are conditioned with the formation of parasite xylem cells to connect it to its host.

The expression of apoplastic *SbtS* has a specific expression pattern associated with development of arbuscular mycorrhiza (Kistner et al., 2005) and rhizobia-mediated nodulation (Takeda et al., 2009). SbtS is phylogenetically distant with other described AM-induced SBTs (SbtM1, SbtM3 and SbtM4) from Lotus, which are closer to AtSBTs of Group 1 (Fig. 3.5). *SbtM1* and *-3* are exclusively induced by arbuscular mycorrhiza (AM) (Takeda et al., 2009), while the AM-induced *SbtS* and *SbtM4* are also modulated by root nodule symbiosis (RNS) (Takeda et al., 2009). *SbtM1*, *-3*, *-4* are activated by the common *SYM* pathway shared by by AM and RNS, whereas *SbtS* is regulated by a *SYM*-independent path. The homologies between haustoria-induced *PjSBTs*, *SbtS* and *AIR3* may suggest the existence of shared components between lateral root emergence, nodulation and haustorium formation. In this respect, it is interesting to know that AM- and RNS-released molecules (Nod and Myc factors) are able to promote lateral root formation (Oláh et al., 2005).

To investigate if the recruitment of SBTs for establishment a compatible host-parasite interaction is conserved among the parasitic plants, three subtilase genes from *S. hermonthica* were analyzed. Partial sequences were initially isolated from Parasitic Plant Genome Project (http://ppgp.huck.psu.edu/) according to their similarities with *PjSBT* and *AtSBT* genes. As showed in the phylogenetic tree (Fig. 3.5), *ShSBT1* is grouped into the group 3, while *ShSBT4* and *ShSBT8* are in the group 5 (Rautengarten et al., 2005). Their transcript levels were quantified in *S. hermonthica* tissues parasitizing *L. japonicum* or rice. *S. hermonthica* is able to establish xylem connection with rice, but it fails to form it in *Lotus* roots (Yoshida and Shirasu, 2009) (Fig. 3.7 B-C). As showed in Figure 3.7-A, the expression profiles of *S. hermonthica SBT4* (*ShSBT4*) *and SBT5* (*ShSBT5*), which show a high similarity to haustoria-induced *PjSBTs* from the group 5, were different from that of *ShSBT1*. These preliminary data may indicate that SBTs may have different roles in the parasitism in the obligate and in facultative parasites. The participation of SBTs with potentially different functions may be reflected by the fact that distinct haustoria structures are formed in obligates and facultative parasite roots. In the facultative parasites, the haustorium develops at the lateral side of roots similar to the lateral root or rhizobia-mediated nodule. In contrast, in obligate parasites, the first haustorium develops at the tip of a radicle (terminal haustorium) with reorganization of the meristematic cells at the root tip that affect the root growth; while the lateral haustorium develops at upper part of the meristematic zone which allows the haustorium formation without affecting the root growth.

3.4 CONCLUSION

In this chapter I described five *P. japonicum SBTs* which were expressed during establishment of parasitism. These *SBTs* belong to the *SBT* group 5 which contains *SbtS* and *AIR3* which are induced symbiosis and lateral root formation, respectively, possibly indicating a common component necessary for haustorium induction and the formation of nodule or lateral root. Furthermore, the localized expression of one of these *SBTs* was observed at the interface between *P. japonicum* and its compatible host at the penetrating stage of haustoria development. In addition, the parasitism-induced *S. hermonthica ShSBT1*, which is phylogenetically distant from the group 5 *PjSBTs,* potentially suggesting that SBTs may acquire distinct functions during the development of terminal or lateral haustoria.

3.5 TABLES AND FIGURES

Table 3.1 Classification of SBT and their distribution in human and in *Arabidopsis*

MEROPS classification		Homology groups	Distribution	
			Human	*Arabidopsis*
S8	S8A	Subtilisin	0	0
		Thermitase	0	0
		Lantibiotic leader peptidase	0	0
		Proteinase K	1	0
		Pyrolysin	2	56
	S8B	Kexin	7	0

* modified from Schaller et al., 2012

Table 3.2 Function of some members of subtilase family in plants

Name	GenBank accession	Organism	Function	SBT group	Reference
SbtS	Chr3.CM1144.130.r2.m **	*Lotus japonicus*	Induced by arbuscular mycorrhizal and root nodule symbiosis	5	(Kistner et al., 2005) (Takeda et al., 2009)
SbtM1, SbtM3 and SbtM4*	Chr2.CM0021.2780.r2.m** Chr2.CM0021.2790.r2.m** Chr4.CM0126.510.r2.m**	*Lotus japonicus*	Induced by arbuscular mycorrhizal (*SbtM4 is also induced by root nodule symbiosis)	1	(Takeda et al., 2009)
Phytaspase	GQ249168	*Nicotiana tabacum*	Involved in programmed cell death	1	(Chichkova et al., 2010)
Gm-1	Glyma18g48580.1	*Glycine max*	Contains the signal peptide GmSubPep that induces defense-related genes	5	(Pearce et al., 2010)
P69B and P69C	CAA76725, CAA06412	*Solanum lycopersicon*	Pathogen-related proteins	1	(Jordá and Vera, 2000)
SDD1	At1g04110	*Arabidopsis thaliana*	Controlling of stomata density	1	(Rautengarten et al., 2005)
ALE1	At1g62340	*Arabidopsis thaliana*	Required for cuticle formation and epidermal differentiation during embryo development	2	(Tanaka et al., 2001)
AIR3	At2g04160	*Arabidopsis thaliana*	Lateral root emergence	5	(Neuteboom, 1999)
AtSBT3.3	At1g32960	*Arabidopsis thaliana*	Switch for immune priming	3	(Ramírez et al., 2013)
Cucumisin	D32206.1	*Cucumis melo*	Involved in melon fruit maturation	4	(Nakagawa et al., 2010)

** miyakogusa project (http://www.kazusa.or.jp/lotus/)

Table 3.3 – Top genes exclusively expressed in the parasitic stage.

Feature ID	log_2 (1+RPKM)	Orthologs in other Plant Genomes						Description
		Chlre	Phypa	Selmo	Orysa	Vitvi	Arath	
Pj3contig_3988	7.96	0	16	91	64	113	56	subtilase family protein
Pj3contig_27632	6.91	-	-	-	-	-	-	** No description available **
Pj3contig_52149	6.74	0	16	91	64	113	56	subtilase family protein
Pj3contig_41594	6.59	0	16	91	64	113	56	subtilase family protein
Pj3contig_3531	6.54	0	16	91	64	113	56	subtilase family protein
Pj3contig_6540	6.02	0	0	2	8	4	5	disease resistance-responsive protein
Pj3contig_49579	5.86	0	16	91	64	113	56	subtilase family protein
Pj3contig_49870	5.74	-	-	-	-	-	-	** No description available **
Pj3contig_1736	5.60	0	16	91	64	113	56	subtilase family protein
Pj3contig_6068	5.57	0	16	91	64	113	56	subtilase family protein

Chre = *Chalamydomonas reinhardtii*; Phypa = *Physcomitrella patens*; Selmo = *Selaginella moellendorffii*; Orysa = *Oryza Sativa*; Vitvi = *Vitis vinifera*; Arath = *Arabidopsis thaliana*. RPKM = reads per kilobase per million.

Table 3.4 – *P. japonicum* Subtilase genes analyzed in this thesis

Feature ID (contig)	Contig size (bp)	Gene name	Full length size of coding region (bp)	log_2 (1+RPKM) parasitic stage	Fold change**
Pj3contig_3988*	376	PjSBT2	2313	7.96	-
Pj3contig_41594*	721			6.59	-
Pj3contig_51440	325			7.75	36.6
Pj3contig_2620*	342	PjSBT4	2277	2.26	-
Pj3contig_35289*	442			1.19	-
Pj3contig_49579*	565			5.86	-
Pj3contig_57758*	336			3.27	-
Pj3contig_1736*	1409	PjSBT7	2307	5.60	-
Pj3contig_6068*	729			5.57	-
Pj3contig_13637	360			0	?
Pj3contig_46128	515			1.32	2.7
Pj3contig_35234*	360	PjSBT8	2325	2.98	-
Pj3contig_19312*	534	PjSBT11	2334	4.32	-
Pj3contig_27842	689			4.84	24.2
Pj3contig_48672	617			5.68	17.6
Pj3contig_57900*	306			2.79	-
Pj3contig_22232	898	PjSBT3	2280	2.12	1
Pj3contig_17077	915	PjSBT6	3096	3.29	1.1
Pj3contig_39971	2581			2.84	0.95

* Exclusively expressed in parasitic stage

** fold change = log_2(1+RPKM-parasitic)/log_2(1+RPKM-control)

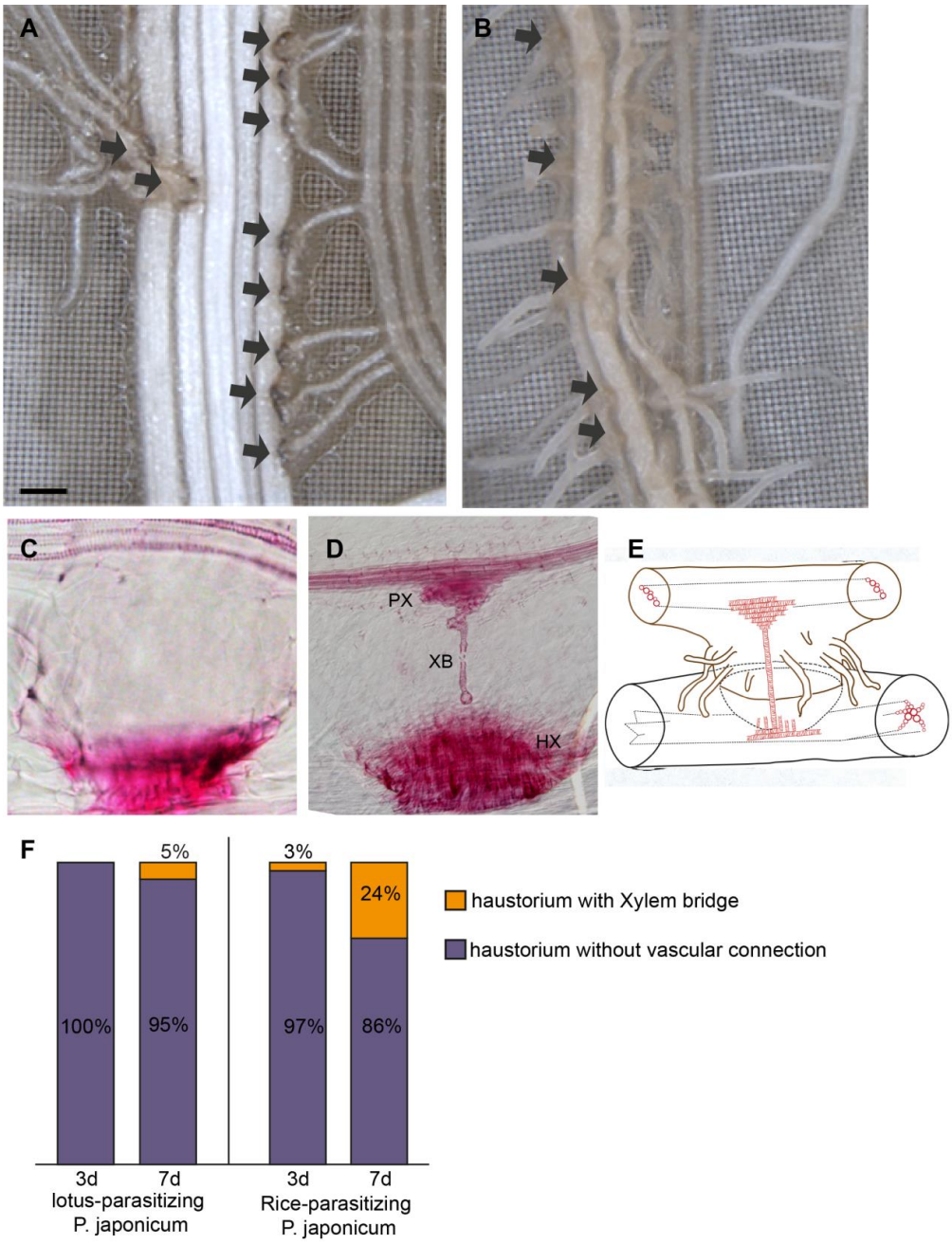


Figure 3.1 – Interaction of *P. japonicum* with nonhost *Lotus japonicus* and compatible host rice. (A-B) *P. japonicum* haustoria penetrating to *L. japonicus* (nonhost) (A) and rice root (host) (B). Arrows point to haustoria. Bar 0.5 mm. **(C-D)** Safrarin-stained of *P. japonicum* haustorium penetrating *L. japonicus* (C) and rice (D) after 7 days of interaction with parasite. **(E)** Schematic view of haustorium (modified from Heide-Jørgensen, H.S and KUIJT, 1995) **(F)** Frequency of xylem ridge formation in *P. japonicum* parasitizing host or nonhost at 3 days or 7 days post inoculation. XB = Xylem bridge; PX = Plate xylem; T = Trichome; Hx= Host xylem; G = graniferous tracheary element.

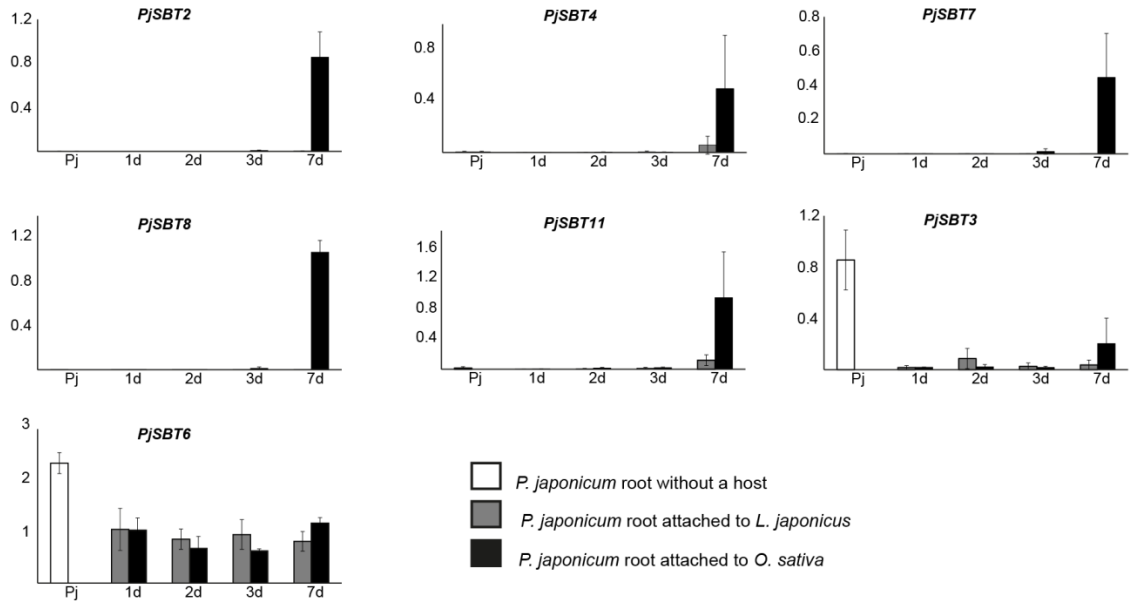


Figure 3.2 – Transcript levels of *PjSBT* in host and non-host plants. Expression of gene encoding Subtilases in the interaction of parasite with the non-host *L. japonicus* (grey) and host rice root (black) analyzed by qRT-PCR. Pj = non-parasitizing *P. japonicum* roots (white); 1d, 2d, 3d, and 7d means the number of days which *P. japonicum* was in contact with rice or *L. japonicus.*

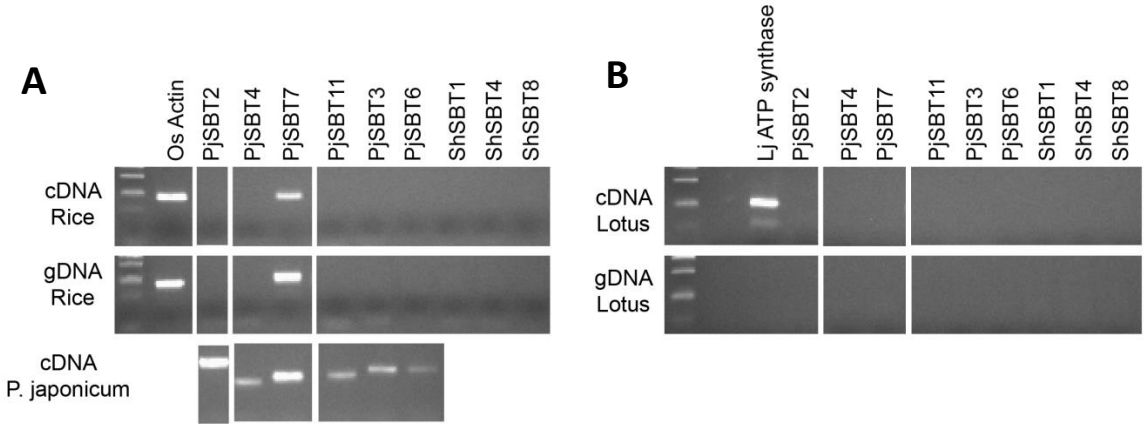


Figure 3.3 – Primers specificity for quantification RT-PCR. Genomic and cDNA of rice(A) and Genomic DNA and cDNA of *L. japonicus* (B) were used as template in a PCR reaction with the PjSBT and ShSBT primer pairs. cDNA samples were extracted from non-infecting root tissue. At the bottom, amplifications from *P. japonicum* roots cDNA after 7 days of infection are shown as positive controls. No DNA amplification was observed in all rice and *L. japonicus* cDNAs nor genomic DNAs, except for *PjSBT7* primer pair which amplified in rice tissues. Primer specificity for *PjSBT8* was not tested in this experiment

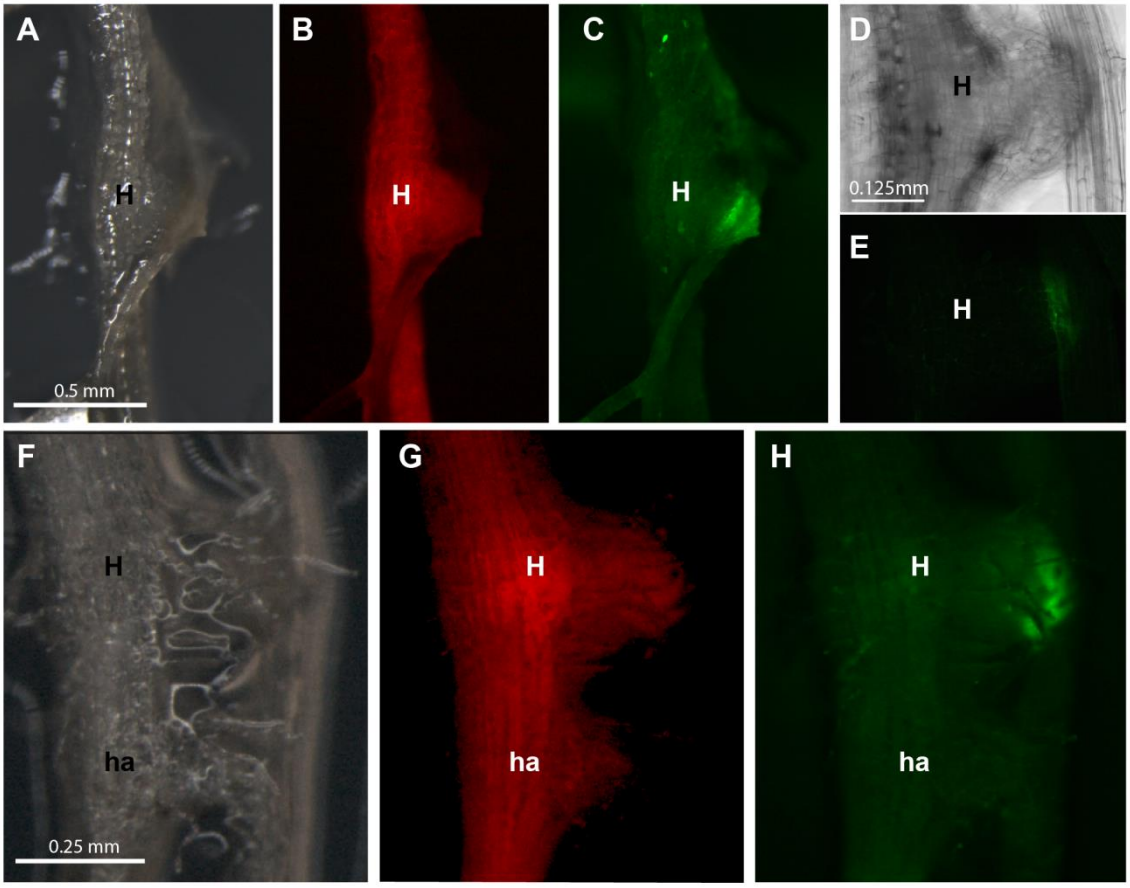


Figure 3.4 - Promoter analysis of *PjSBT8* gene. Transgenic *P. japonicum* roots transformed with plasmid R4L1-35SRFP containing the promoter region of *PjSBT8* fused to *GFP* penetrating Arabidopsis root (pro SBT8::GFP) are observed under bright field (A and F), or fluorescence microscopy with RFP (B and G) or GFP (C and H) filters. A pricture taken under confocal laser scanning microscope (E) and corresponding bright filed image (D) are shown.. H = mature haustoria; ha = immature haustoria. Photographs were taken 10 days after incubation with host roots.

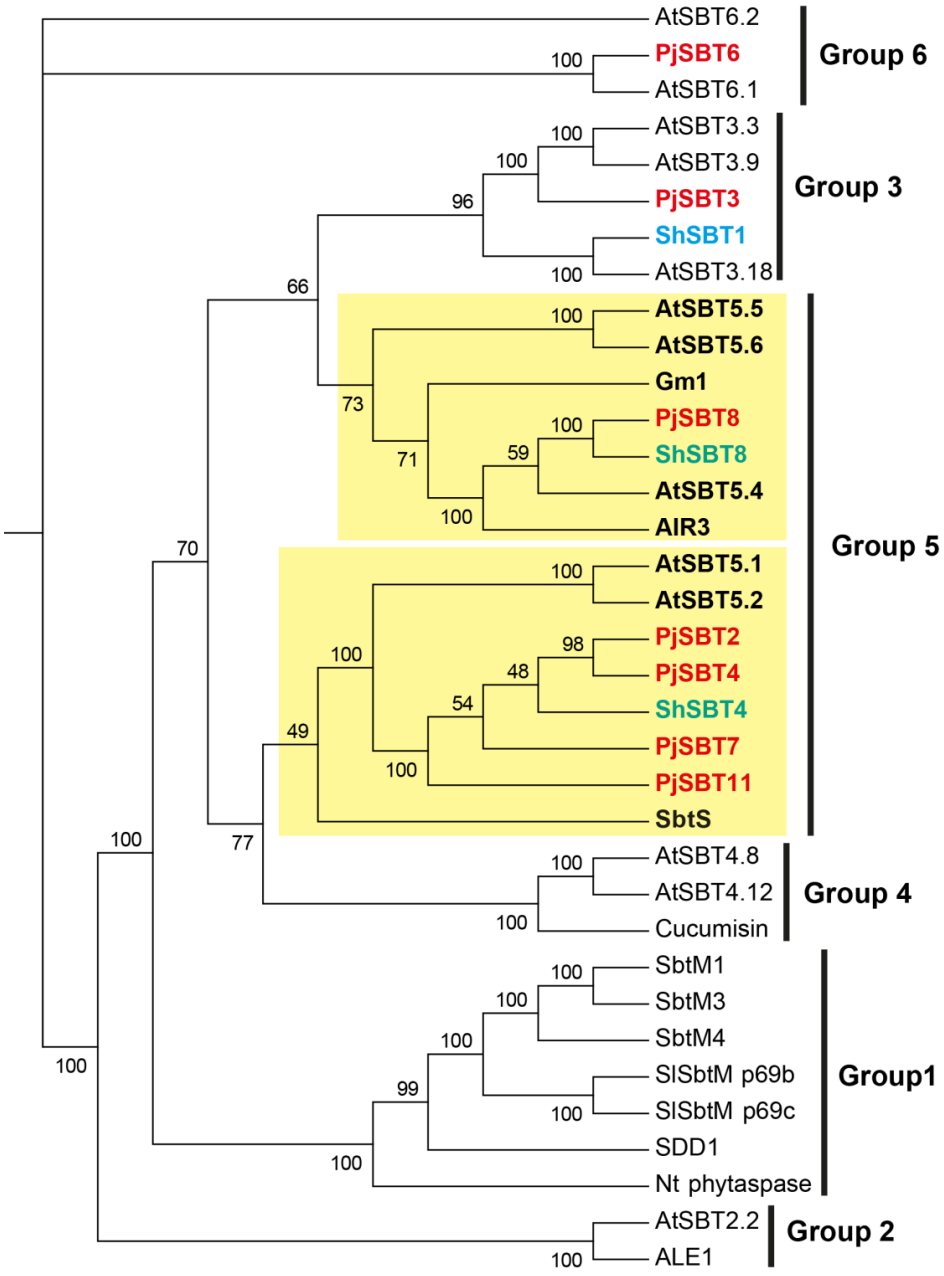
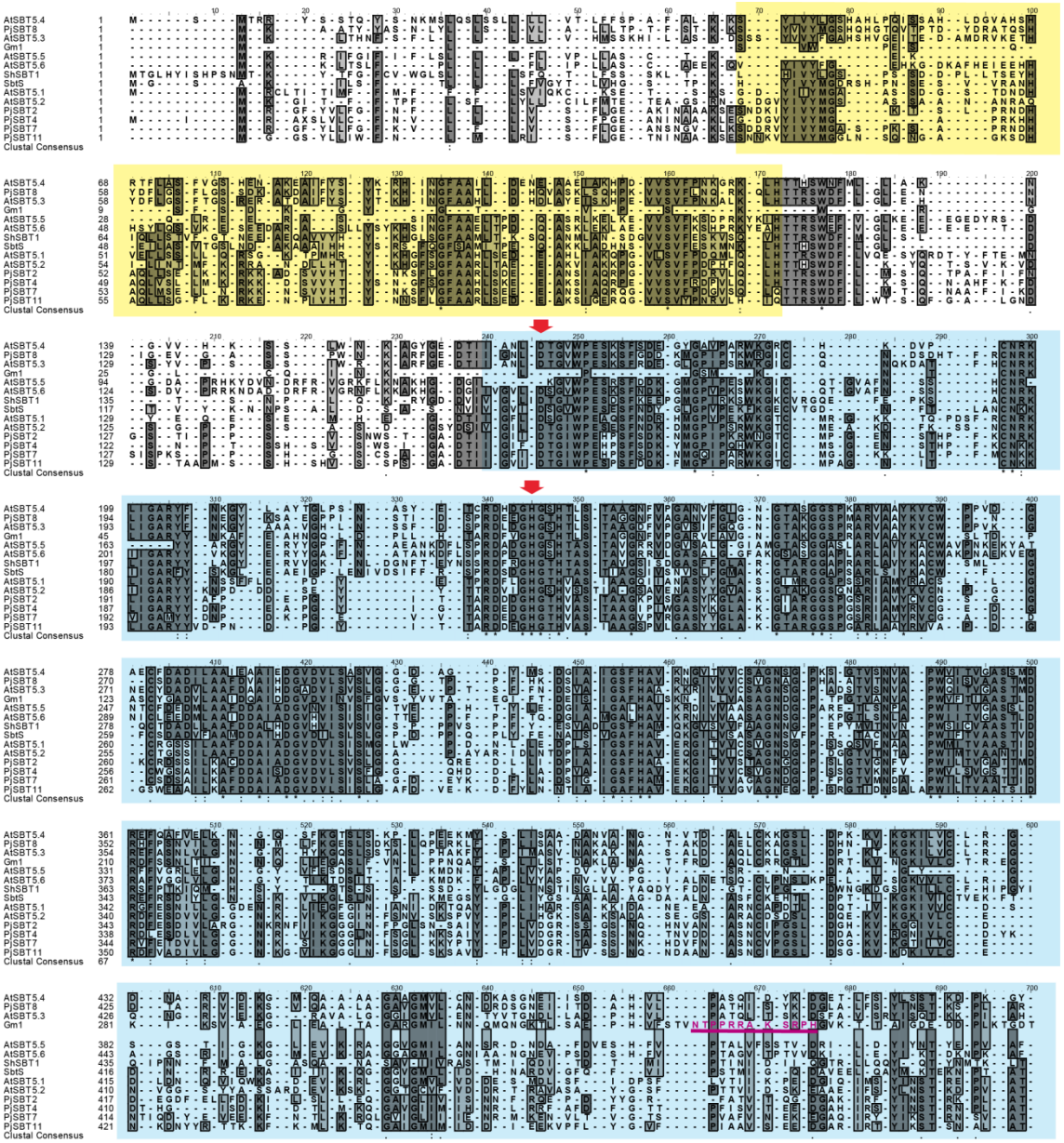


Figure 3.5 – Phylogenetic tree of plant Subtilases. UPGMA tree based on full length protein alignment of subtilases from *P. japonicum* (PjSBT, shown in red), *Striga hermonthica* (ShSBT, shown in blue), some SBT members from *A. thaliana* (AtSBT) and SBT genes whose the functions are known. The yellow boxes highlight the PjSBT in group 5. Arabidopsis gene accession numbers are followings; AtSBT2.2 (At4g20430); AtSBT3.9 (At4g10520); AtSBT3.18 (At4g26330); AtSBT4.8 (At5g58830); AtSBT4.12 (At5g59090); AtSBT5.1 (At1g20150); AtSBT5.2 (At1g20160); AIR3 (PjSBT5.3, At2g04160); AtSBT5.4 (At5g59810); AtSBT5.5 (At5g45640); AtSBT5.6 (At5g45650); AtSBT6.1 (At5g19660) and AtSBT6.2 (At4g20850). Further details in Table 3.2


AtSBT5.4
PjSBT8
AtSBT5.3
Gm1
AtSBT5.5
AtSBT5.6
ShSBT1
SbtS
AtSBT5.1
AtSBT5.2
PjSBT2
PjSBT4
PjSBT7
PjSBT11
Clustal Consensus

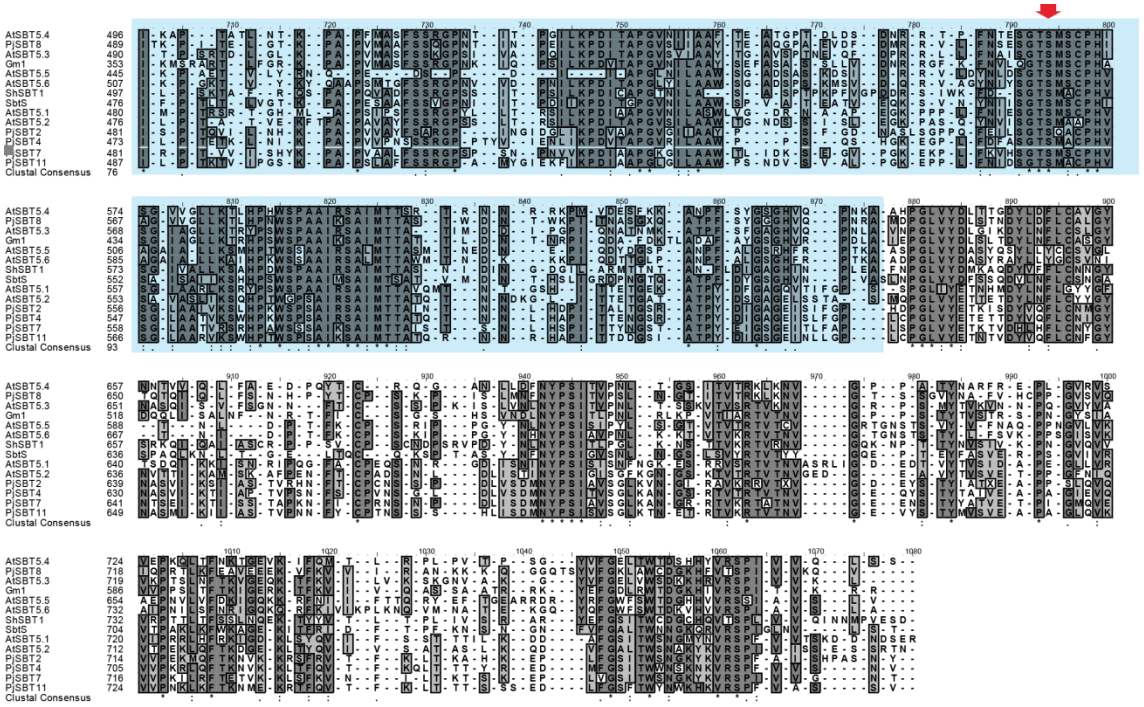


Figure 3.6 - Amino-acid alignment of SBTs clustered to AtSBT group 5. The sequences were aligned using CLC Genomics Workbench (v 4.8) and the figure generated by BioEdit. The inhibitor 9 domains are highlighted in yellow and the peptidase S8 domain in blue. The triad aminoacid D, H and S are pointed by red arrows. The 12 aa bioactive peptide located within Gm1 (Glyma18g48490.1) is shown in pink.

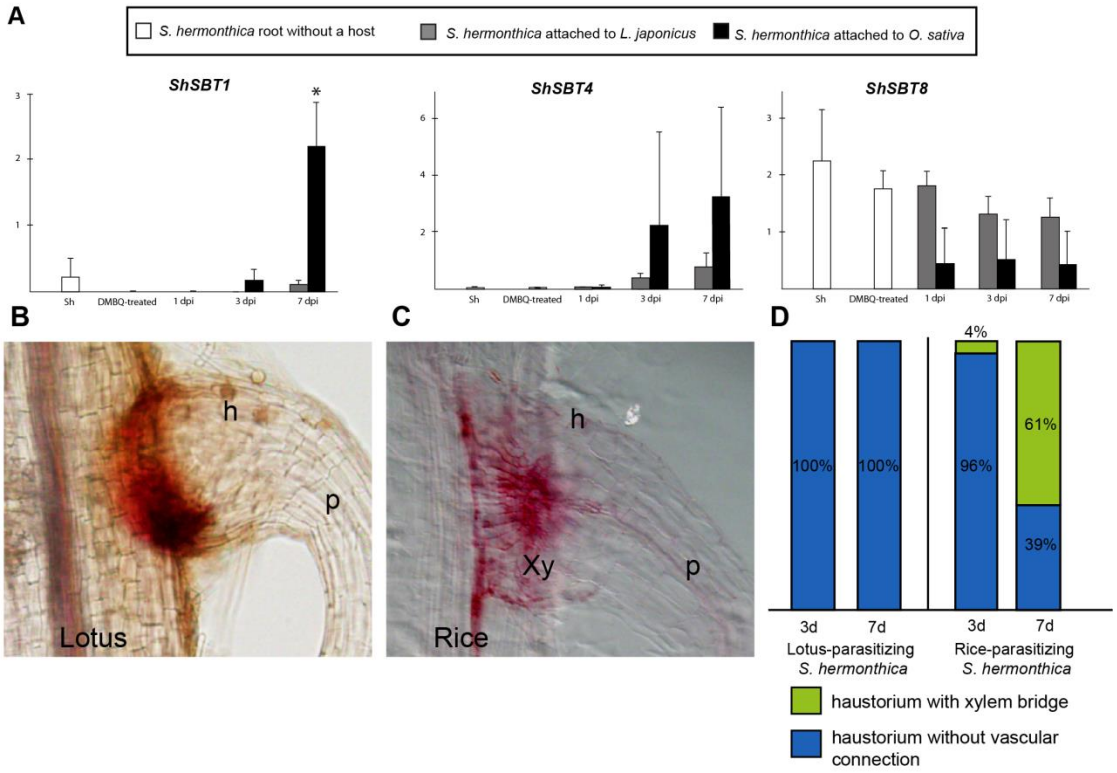


Figure 3.7 – Transcript levels of *ShSBT* in host and non-host plants. (**A**) Expression levels of gene encoding Subtilases quantified by qRT-PCR in the interaction of parasite with the non-host *L. japonicus* (grey) and host rice root (black). Sh = non-parasitizing *S. hermonthica* radicles (white); 1 d, 2 d, 3 d, and 7 d indicate number of days after the infection of *S. hermonthica* in rice or *L. japonicus* roots. Data correspond to two biological replicates. **(B - C)** Safrarin-stained of *S. hermonthica* haustorium penetrating *L. japonicus* (B) or rice (C) roots. h = haustorium; p = parasite root; xy = xylem of parasite. (**D)** frequencies of xylem bridge establishment in *S. hemonthica* parasitizing host or nonhost at 3 days or 7 days post inoculation.

4

CHAPTER
AUXIN BIOSYNTHESIS MEDIATED BY THE YUCCA FLAVIN MONOOXYGENASE HOMOLOG IS ESSENTIAL FOR THE HAUSTORIUM DEVELOPMENT IN ROOT PARASITIC PLANT *P. JAPONICUM*

4.1 SUMMARY

To understand the molecular events associated with host perception and haustorium development, I identified differentially regulated genes expressed during early haustorium development in a model Orobanchaceae plant, *P. japonicum*. I generated a *de novo* assembled transcriptome enriched with transcripts of *P. japonicum* in the parasitic stage and designed a customized microarray. The microarray-based expression studies identified 1,577 genes differentially regulated by the haustorium-inducing compound 2,6-dimethoxy-*p*-benzoquinone (DMBQ). Among them, *PjYUC3* encoding a flavin monooxygenase similar to Arabidopsis YUCCA proteins that are involved in auxin biosynthesis pathway. *PjYUC3* was upregulated around 24-h after haustorium-inducing treatments. Four YUCCA homologs were identified from the *P. japonicum* transcriptome but only *PjYUC3* was specifically up-regulated after the haustorium-inducing treatments. *P. japonicum* hairy roots transformed with the *PjYUC3* overexpression construct resulted in a typical auxin-overproducing phenotype, indicating that *PjYUC3* encodes a functional YUCCA enzyme. Promoter analysis showed that *PjYUC3* expression occurs in two different locations, at root apical meristem and at the haustorium initiation site. Silencing *PjYUC3* provoked reduction of haustorium numbers, emphasizing its relevance for the haustorium development. Our results suggest that the *de novo* auxin biosynthesis by *PjYUC3* at the root aplical meristem and/or haustorium initiation region is essential for the haustorium development.

4.2 INTRODUCTION

The maintenance of auxin gradients generated either by auxin flow or its synthesis is essential for root development. Auxin accumulation levels regulate length and number of root hairs as well as rates of lateral roots (reviewed in Jones et al., 2009). A large portion of auxin molecules are synthesized in the young leaves and cotyledons, and transported to underground parts through phloem channels, or via cell-to-cell basis, mediated by PIN transporters. PIN-mediated auxin transportation together with *de novo* synthesis generates local auxin gradients, which regulate local cell differentiation, root hair elongation, and maintenance of cell niche in root meristem. The main pathway of auxin biosynthesis involves two short steps, the TRYPTOPHAN AMINOTRANSFERASE OF ARABIDOPSIS (TAA) family of amino transferases converts tryptophan to idole-3-pyruvic acid (IPA), and subsequently *YUCCA* family of flavin monooxygenases catalyzes indole-3-acetic acid (IAA) production from IPA (Mashiguchi et al., 2011). The rate-limiting step for the IAA production is controlled by the *YUCCA* genes (Zhao et al., 2001). There are nine *YUCCA* gene family members in the *Arabidopsis* genome, and seven in rice (Yamamoto et al., 2007). Consistently, the phenotypes in the *YUCCA* multiple mutants are similar to known auxin transport and signaling mutants (Cheng et al., 2006)

The haustorium initiation in the facultative parasite *T. versicolor* is associated with local auxin accumulation (Tomilov et al., 2005). Disturbing auxin flow either by application of auxin efflux or activity inhibitors causes a reduction of haustorium number in *T. versicolor* (Tomilov et al., 2005). Similarly, application transport inhibitors or excess amounts of exogenous auxin disturb infection by the holoparasitic *Phelipanche aegyptiaca* (Bar-Nun et al., 2008). Dissection series of terminal portion of root, which included the region from apical meristem to root hair formation zones, indicated that the proximal area to the haustorium initiation site is a source of auxin, and auxin is transported toward root tip within a few hours after HIF treatment. Dissection of root from proximal region of haustorium formation before HIF treatment significantly reduces numbers of haustoria and this phenomenon is rescued by application of IAA, suggesting such auxin transportation from the proximal region is necessary for haustorium initiation. The local application of TIBA, an auxin efflux inhibitor, to a root tip suggests that the auxin transport determines the place of haustorium development, since this experiment resulted in misplaced development of a haustorium (Tomilov et al., 2005).

In this article we studied a model parasitic plant *P. japonicum* to unravel differentially regulated genes during initial stages of haustorium development. First, we established the *P. japonicum* transcriptome enriched with transcripts from haustoria at different developmental stages as a basis to design a custom microarray. The transcriptome/microarray analysis identified 1,577 genes differentially regulated by DMBQ. Next, we focused on a highly upregulated gene, *PjYUC3*, a member of the *YUCCA* family. We analyzed the *in vivo* expression patterns and functions of *PjYUC3* by using promoter-reporter analysis, overexpression and silencing experiments. Our functional studies suggest that *PjYUC3* encodes a key functional enzyme with expression localized at root apical meristem and at haustorium initiation region. Analysis with a prometer responsive to auxin indicated accumulation of auxin in early haustorium at similar site of *PjYUC3* expression, suggesting that accumulation of auxin may be driven by *PjYUC3*.

4.3 RESULTS

4.3.1 IDENTIFICATION AND CHARACTERIZATION OF *P. JAPONICUM* GENES INDUCED BY DMBQ

To identify the genes essential for haustorium formation I designed a customized microarray based on the *de novo* assembled *P. japonicum* transcriptome sequences. To avoid misassembled sequences that may interfere in the microarray analysis, unigenes with a putative coding region with at least 50 amino acids were selected and gene specific 60-mer probes were designed. *P. japonicum* roots were treated with DMBQ, and the gene expression was analyzed with microarray hybridization. To select the time points we carefully observed the morphological changes during the haustorium development. *P. japonicum* root was treated with DMBQ and monitored with time-lapse photographic series. Frames of the root tip which included root cap, cell division and elongation zones, were taken at 30-min intervals within a period of 48 h. The earliest morphological change occurred immediately behind the root tip and was observed at 9-12 h of exposure when the rate of the root elongation was reduced and a concomitant differentiation of epidermal cells into root-hair like haustorial hairs initiated immediately behind the root tip. This region preceded the establishment of the haustorium, which I termed prehaustoria site. At the point of 18 h, the haustorial hairs continued proliferating and a swelling started to develop. At 24 h, the haustorium was formed and its development was completed by 48 h after exposure to DMBQ (Fig. 4.1). The haustoria ontogeny events in *P. japonicum* showed similar to that described in *T. vesicolor* (Matvienko, 2001).

Based on the above observation, total RNAs were extracted from the *P. japonicum* roots at 0 h, 0.5 h, 1 h, 3 h, 6 h, 12 h, 24 h and 48 h after the DMBQ treatment and hybridized with the custom microarray slides. To remove the genes induced by treatment stresses, microarray slides were hybridized with labeled cRNAs from parasite roots under mock treatment and the genes which differentially regulated under mock treatments were subtracted from our analyses. The reliable differentially expressed genes were selected through statistical analysis across three biological replicates. Using as parameter a fold change larger than 2, I found that 481, 391, 461, 284, 326, 217 and 133 genes modulated exclusively by DMBQ for the 0.5, 1, 3, 6, 12, 24 and 48 h samples, respectively (Fig. 4.2-A). Large numbers of genes were up- or down-regulated during early stages (30 min to 3 h) of the haustorium development (Fig. 4.2-A). As some DMBQ-regulated genes were detected at more than one time point, a set of 1577 non-redundant genes were identified as DMBQ-up or down-regulated

genes. The 1577 non-redundant DMBQ-regulated genes were divided into three clusters according to their expression profiles. Cluster 1 contains 706 genes negatively modulated along the time course (Fig. 4.2-B). Cluster 2 has 396 positively regulated genes with a peak of expression before 3 h of treatment, designated as early responsive genes (Fig. 4.2-C). Cluster 3 contains 475 positively regulated genes with a peak of expression after 3 h of treatment, designated as late responsive genes (Fig. 4.2-D). The list of top differential expressed genes in each cluster is shown in table 4.1-4.3.

4.3.2 FUNCTIONAL CLASSIFICATION OF *P. JAPONICUM* GENES WITH ALTERED EXPRESSION PATTERNS IN RESPONSE TO HAUSTORIA-INDUCING FACTOR DMBQ.

Based on their best BLAST-hit annotations, the gene ontology (GO) terms were assigned for each gene, aiming to investigate functional gene populations in each cluster. Comparing with other clusters and all the entities, the cluster 1 showed higher percentage of sequences assigned as “other molecular functions” (Fig. 2-E), 139 sequences out of 701. In this category four genes identified as Glyoxal oxidase (GLOX), a enzyme is responsible for hydrogen peroxide production in the extracellular region (Kersten and Kirk, 1987; Leuthner et al., 2005) (Guan et al., 2011). The transcripts were downregulated from 3 h to 48 h after adding DMBQ with fold change varying from -2.4 to -13 times less expressed in DMBQ-induced haustoria than non-treated roots. Other four genes assigned as “other molecular function” category encoded to dirigent-like protein (Table 4.1), involved in the lignin biosynthesis pathway and related to plant defense since lignin deposition consists in a physical barrier against intruders (Burlat et al., 2001; Vance et al., 1980). The suppression of this gene was maintained from 1 h to 48 h during the haustoria development with a fold change ranging from -2.3 to -4.2. Sequences assigned as unkown protein appeared 46 times in this GO category. To gain insights about their potential biological function I searched among them for known domain in InterPro database (Apweiler, 2001), 17 showed a predicting domain. Out of them, 10 were associated with different domains of unknown function (DUF) and three sequences had Villin headpiece domain (IPR003128), typically found in Villin cytoskeletal protein. The gene encoding to Villin protein (CUST_12394_PI426107926) was found to be 2.2 times less expressed in DMBQ-treated roots compared with the control (Table 4.1). This protein is essential for bundling the actin filaments in plants (Tominaga et al., 2000).Thus, the analysis of cluster 1 suggested that within 48 h of DMBQ treatment GLOX was suppressed which potentially altered the redox status of extracellular region. The down regulation of genes encoding to proteins responsible

for lignification and strengthening of cell wall and maintenance of microvilli of cytoskeleton, suggested intense restructure of parasite cells in the DMBQ-induced haustoria.

Up-regulated transcripts in cluster 2 and 3 showed higher percentage of genes assigned in the "transcription factor activity" compared with the cluster 1 and all the entities in the microarray. In the cluster 2 was composed by the early-responsive genes (Fig. 2-E). In this GO category 7 sequences belonged to WRKY transcription factor, members of this family are associated with responses of abiotic and biotic stress. The transcript of *WRKY23* activated during the nematode plant parasitism, showed positively modulated until 6 h after adding DMBQ, and the stress-related *WRKY33* was activated from 1h to 3h of the chemical treatment. And the pathogen-induced transcription factor *WRKY40* had with a peak of 7.2 fold change at 1 h and *WRKY29* showed three times more expressed at 3 h time-point. Cluster 2 also showed higher percentage of sequences with unavailable function compared with other clusters and all the entities in the microarray. To increase the information about those genes I searched for predictable domains, 92 of 115 sequences were not assigned with any domain in InterPro database. It is noteworthy that in this cluster the homolog of Tv*Pirin* was one of the highest expressed genes in the microarray. These genes were previously described in *T. vesicolor* as essential for development of DMBQ-induced haustoria (Bandaranayake et al., 2012). In summary, the cluster 2 composed by up-regulated transcripts at the early moments of haustorium initiation concentrated sequences without biological information available, indicating that large amount of relevant genes for parasitism remain unexplored, among known sequences the transcription factors were more represented, with members of WRKY family modulated by plant interactions showed the highest expression levels.

In the cluster 3 sequences associated with "transcription factor activity" category, the gene enconding to beta HLH protein 93 (bHLH093) was the transcription factor gene with the highest modulation in the microarray, keeping its positive expression from 3 h to 48h of DMBQ treatment. The fold-difference of expression values varied from 3 to 19 times in treated root compared with non-treated. The homolog of bHLH093 in *Arabidopsis* (At5g65640) is bHLH from subclass IIIb in involved with abiotic response and stomata development (Heim et al., 2003). Its involvement with plant parasitism was first detected in this experiment and further experiments are necessary to clarify its function during the haustorium development. In addition, other class of GO terms with higher percentage in Cluster 3 was "hydrolases activity" (Fig. 2-E), which includes enzymes involved in cell wall modifications, such as pectinesterase, glycoside hydrolase, polygalacturonase, cellulases and peroxidases. In addition, this cluster

includes hormone-related enzymes such as YUCCA (see below), CYTOKININ OXIDASE 3 (CKX3) and GIBBERELLIN 2-OXIDASE 8 (GA2ox8), which function in auxin, cytokinin and gibbberellin metabolisms, respectively. Thus, the late-responsive genes grouped in cluster 3 showed activation of cell-wall-modifying e and hormone-metabolic enzymes, indicating that from 3 h to 48 h of DMBQ-induced haustoria development occurred intense degradation of cell wall, followed by activation of genes involved with auxin biosynthesis and degradation of cytokinin and gibberellin.

To validate the microarray results, we performed quantitative reverse-transcriptase PCR (qRT-PCR) analysis with representative genes in each cluster. For an internal control to normalize the expression value across the samples, the parasite-specific reference gene encoding to RNA-binding *Polypyrimidine tract- binding protein* (*PjPTB*) was selected because it was the closest homolog of *Arabidopsis* housekeeping gene At3g01150 (Czechowski et al., 2005) and showed a stable expression levels at eight time-points of *P. japonicum* microarray. Expression profiles generated with qRT-PCR showed similar patterns to those obtained from microarray analysis (Fig 4.3 A and C), verifying validity of the microarray data set. Although DMBQ is a natural compound identified from host root extracts (Chang and Lynn, 1986), DMBQ-inducible genes are not necessarily induced in other haustorium-inducing conditions such as host root exudate treatments (Bandaranayake et al., 2010). To confirm if the DMBQ-inducible genes identified in this study are also induced by host root exudates, this treatment severely induce the haustorium formation in *P. japonicum* roots as shown in figure 4.4 B to D. The gene expression profiles were monitored by qRT-PCR. The tested gene set showed similar expression profiles in host-root-exudate- and DMBQ-treated roots (Fig 4.3 B and D), confirming that the gene expression profile in DMBQ-treated samples is useful to understand the haustorium formation.

4.3.3 A *YUCCA* GENE HOMOLOG IS UP-REGULATED DURING HAUSTORIUM FORMATION.

Transcription of hormone-related enzymes such as YUCCA, CKX3 and GA2ox8 were turned on during the haustorium development. To further understand the involvement of hormone in plant parasitism I focused on homologous to *Arabidopsis YUCCA* genes that encode an essential enzyme to catalyze a rate-limiting step in the auxin biosynthesis. I reasoned that YUCCA might be important for haustorium development once it was already shown that auxin accumulation is observed in DMBQ-induced haustoria in *T. vesicolor* (Tomilov et al., 2005). At first, I searched in *P. japonicum* transcriptome for homologs of *Arabidopsis*

YUCCA (*AtYUC*) genes. Five *YUCCA*-homolog genes were identified in the *P. japonicum* transcriptome assembly and their full-length sequences were determined by RACE analysis. One of the five genes was excluded from our analysis since its full-length sequence could not be amplified by RACE. The amino acid level alignment of *YUCCA* homolos in *P. japonicum* is shown in figure 5. We named the *P. japonicum YUCCA* homologs as *PjYUC*1, *PjYUC*2, *PjYUC*3 and *PjYUC*4. *PjYUC*3 was identified as DMBQ-inducible in the microarray analysis. Phylogenetic tree revealed that *PjYUC*2, *PjYUC*3 and *PjYUC*4 are clustered together with *AtYUC* genes expressed exclusively in roots (*AtYUC5,-3,-7,-9, -8*) (Fig. 4.6 A). This is consistent with the fact that the *P. japonicum* transcriptome assembly was originated from root RNA sequences. Consistent to the microarray analysis, only *PjYUC3* was highly up-regulated after the host root exudate treatment, but other *PjYUC* genes showed constant low-level expressions (Fig. 4.6 B).

4.3.4 *PJYUC3* ENCODES A FUNCTIONAL AUXIN BIOSYNTHESIS ENZYME.

To determine whether the *PjYUC3* gene encodes a functional enzyme in auxin biosynthetic pathway, transgenic roots in which *PjYUC3* was overexpressed were analyzed. The *PjYUC3* cDNA was amplified and cloned behind the constitutively-expressed ubiquitin promoter in pUB-GW-GFP plasmid, which harbors green fluorescence protein (GFP) as a visible marker. The resultant construct was transformed to generate *P. japonicum* hairy roots using the transformation system via *A. rhizogenes* (Ishida et al PLOS ONE). The higher expression of *PjYUC3* did not affect the transformation rates, which ranged alike the transformation with control vector, from 20% to 40% of success rates. The qRT-PCR analysis verified that *PjYUC3* transformed hairy roots exhibited over 70 times higher expression than non-transformed roots (Fig. 4.7D). *PjYUC3* overexpressed roots showed clear morphological phenotypes, including high proliferation of root hairs (Fig. 4.7 E-F), short root length (Fig. 4.7 E-F) and promotion of lateral root formation (Fig 4.7 G-H), which phenocopy the auxin-overproducing plants (Boerjan, 1995). These observations highlight that the PjYUC3 protein most likely kept its catalytic ability for auxin biosynthesis.

4.3.5 TISSUE-LOCALIZED EXPRESSION OF *PJYUC3* IS SPECIFIC FOR ROOT TIP AND IN THE HAUSTORIUM REGION.

To further characterize the roles of *PjYUC3* in the haustorium development, localization of *PjYUC3* expression was investigated. The promoter region was cloned upstream to *GFP* and *beta-glucuronidase* (GUS) reporter genes and introduced to *P. japonicum* hairy roots using *A. rhizogenes*. To facilitate the screening of transgenic roots we modified

R4L1pGWB532 vector (Nakamura et al., 2009), inserting the *red fluorescent protein* gene (*mRFP*) driven by the *35S* promoter as a visible marker. Transgenic RFP-fluorescent hairy roots were screened under a fluorescent microscopy and the RFP-positive roots were stained for the GUS activity. Without DMBQ nor host root exudate treatment, no GUS staining was detected in the transgenic roots (Fig. 4.8 A). In contrast, when they were submitted to the host exudate treatment the GUS activity was detected in two different regions (Fig 4.8 B-D). The most pronounced localization was at root apical meristem expressed between 24-30 h after the treatment (Fig. 4.8 D). Interestingly, we also detected localized *PjYUC3* expression at the lateral side of the haustorium initiation region (Fig. 4.8 B). Two distinct local expression patterns of *PjYUC3* at meristematic zone and emerging haustorium indicate the *de novo* auxin biosynthesis in specific root cells during haustorium development.

To confirm the involvement of *PjYUC3* in auxin production during the haustorium development *P. japonicum* was transformed with the construct containing DR5 promoter, an artificial auxin-responsive promoter, fused upstream to VENUS-N7, a modified version of yellow fluorescent protein (YFP) with nuclear signal tag (Brunoud et al., 2012). As showed in Figure 4.8 E-H, DMBQ-induced haustorium in transgenic root showed auxin accumulation at specific set of epidermal and cortical cells localized at the region of emerging haustorium, which corresponds to the site where the promoter of *PjYUC3* gene was activated. Thus, these data strongly suggest that PjYUC3 enzyme is involved in accumulation and production of *de novo* auxin at the haustorium initiation site.

4.3.6 *PJYUC3*-RNAI REDUCES THE HAUSTORIUM DEVELOPMENT IN *P. JAPONICUM*.

The hairpin RNAi vector carrying the specific sequence fragment targeting *PjYUC3* transcripts (pHG8-YUC3) were generated and transformed to *P. japonicum* seedlings to investigate function of *PjYUC3* in the haustorium formation. Transcript levels of *PjYUC3* were measured by qRT-PCR in the transgenic roots under contact to the host root exudates. In the pHG8-YUC3 transformed roots the *PjYUC3* expression was significantly reduced comparing to the root transformed with the control pHG8-YFP vector (Fig. 4.9 A). In addition, the *PjYUC3* silenced roots and control transformed roots showed no detectable morphological differences, suggesting that *PjYUC3* is not the major factor for normal root development (Fig. 4.9 C-E). However, the frequency of haustorium formation was significantly lower in pHG8-YUC3 transformed roots compared with roots carrying an empty vector (Fig. 4.9 B). The host root exudate treatment provoked haustorium appearance in 100% of the normal non-transformed

roots as showed in Fig 4.4 A-D. In the non-fluorescent or empty-vector transformed hairy roots, which emergeed from cotyledons, formed haustoria in 59 to 67% of roots in the same inducing condition (Fig. 4.9 A). However, the *PjYUC3*-silenced roots dramatically reduced the haustorium development frequency to 31%. The figure 4.9 D-E shows an example of silencing roots without haustorium, while a non-fluorescent hairy root at same condition was able to develop a haustorium.

4.4 DISCUSSION

Parasitic plants are harmful threats in agriculture in worldwide. Despite its economic importance, molecular mechanisms controlling the establishment of plant parasitism are poorly known. For understanding the key factors underlying the transition from an autotrophic to parasitic lifestyle as well as increasing the knowledge to combat these parasitic pests, it is essential to access to essential genes regulated by the infection. The advent of next-generation sequencing has accelerated the amount of genomic information available for several non-model plants. In this study, a combination of *de novo* transcriptome and customized microarray of the facultative hemiparasite *P. japonicum* provided a comprehensive access to relevant genes on first 48 h of DMBQ-induced haustorium development. Here, using expression and functional analysis I showed that the parasite gene *PjYUC3* encodes to a functional auxin biosynthesis enzyme which is responsible for local auxin accumulation essential for haustorium development in *P. japonicum* roots.

4.4.1 COMPREHENSIVE SURVEY OF THE GENES REGULATED IN THE HAUSTORIUM DEVELOPMENT OF *P. JAPONICUM*

Microarray hybridization analysis revealed that 1577 genes were differentially regulated after the treatment of haustoria-inducer chemical DMBQ on first 48 h of haustorium development. Expression profile of those genes divided them into 3 clusters. In general there is a large amount parasite transcripts with no significant alignment with sequences in available databases, the percentage of those were higher among genes up-regulated on first 3 h of haustorium development, which indicated that genes recruited at early stages of parasitism is largely unknown. Similarly, micro dissection at interface of *T. vesicolor* (Honaas et al., 2013) with its hosts also pointed an enrichment of unknown parasite sequences on underground phases of parasite growth. Thus, those sequences represent a rich material for discovery and further characterization of novel gene functions.

Gene ontology analysis revealed that an early responsive gene from cluster 2 has overrepresentation of parasite genes assigned as Transcription factor activity. Based on this analysis I found members of WRKY family which respond to pathogen attack. For example, WRKY33 is a positive modulator of resistant against the *Alternaria brassicicola* and *Botrytis cinere* (Zheng et al., 2006), while the auxin-inducible WRKY23 increase resistant agains the plant-parastic Heterodera schachtii (Grunewald et al., 2008). In contrast, homolog of WRKY40 in *Arabidopsis* negativelly regulate the resistance against powdery mildew infection (Pandey et

al., 2010) and *Pseudomonas syringae*, but promotes resistant against *B. cinerea* (Xu et al., 2006)., The up-regulation of WRKY factors that activate the plant defense on first 3 h of haustorium development, might indicate that initial perception of host-derived molecules, such as haustoria-inducer chemical DMBQ, passed by rapid activation of defense-related genes. Consistently, the overexpression of plant genes responsive to pathogen attack is also described in parasite *T. vesicolor* during infection process (Honaas et al., 2013). It still remains elusive the mechanism that parasitic plant adopt to modulate its own plant defense responses to allow the infection in host plants.

Lignin is one of organic polymers most commonly found on earth, and it part of secondary cell walls filling spaces between cellulose, hemicellulose and pectins (Scurfield, 1973). Its deposition increases the cell wall rigidity and confers impermeability and resistance for plant tissues against microbial and other mechanical attacks (Vance et al., 1980). In the custom-designed microarray analysis I identified among downregulated parasite transcripts, the genes encoding to dirigent-like proteins involved with lignin/lignan biosynthesis pathway (Burlat et al., 2001). Those were assigned as “Other molecular function” in GO classification, a category which was overrepresented in downregulated cluster (Cluster 1) compared with all the entities of the microarray and with up-regulated parasite genes. Dirigent-like proteins involved with which confers insect resistance by conifers (Ralph et al., 2006) and cyst nematode resistance in soybean (Afzal et al., 2009). It may suggest that synthesis of protective lignin derivatives is suppressed in parasitic roots during the haustorium development. It is also possible that haustorial cells keep a certain degree of flexibility of their cell wall by suppressing genes involved lignin polymer formation during the penetration into the host tissue.

The flexibility of cell wall during the first 48 h of haustorium development is reinforced by the reduction of transcript levels of genes encoding to Villin or with Villin headpiece domain in their predicted protein sequences. Cross-linking Villin is the molecule responsible for arranging the bundles of actin filaments and keep the microfibrils of cytoskeleton (Tominaga et al., 2000). Alteration at expression levels of genes related with actin reorganization was further described in substrative cDNA library enriched with DMBQ-induced transcripts of *T. vesicolor* (Matvienko, 2001). Thus, the underlying mechanisms of haustorium induction during parasite infection may promote restructuring the actin cytoskeleton which contributes for intense cell shape changes.

Modifications of cell structure cannot take place without the alteration of parasite cell wall compounds, the molecular events to promote such biological process was observed from 3 h after adding DMBQ with up-regulation of cluster-3 transcripts encoding to proteins involved in cell wall digestion or its loosening (Table 4.3). Similarly *T. vesicolor* transcripts encoding to enzymes responsible for hemicellose degradation were up-regulated in DMBQ-treated roots for 2 or 5 h. Other enzyme in cluster 3 was class III peroxidases, a plant-specific enzyme acquired by terrestrial plants throughout evolution (Duroux and Welinder, 2003). In general these apoplastic secreted proteins affect cell wall though redox reactions (Passardi et al., 2004). For example, the hydroxyl radicals cleave various cell wall polysaccharides, including pectin and xyloglucan (Schweikert et al., 2000) which may cause non-enzymatic cell wall loosing and stimulate cell enlargement. Recruitment of peroxidases in haustorium development was previously described in hemiparasites *S. asiatica* (Kim et al.), *Phelipanche ramosa* (González-Verdejo et al., 2006) and *Viscum angulatum* (Das et al., 2011). The overexpression of cell-wall-modifying enzymes coincide with the cell enlargement and division that occurred on first 24 h of haustorium development (Ishida et al., 2011) (Fig. 4.1). Thereby, the transcriptional activation or suppression of parasite genes found in microarray followed the morphological changes observed in kinetics of haustorium ontology.

4.4.2 ACTIVE ENZYME ENCODED BY *PJYUC3* IS IMPORTANT FOR HAUSTORIUM DEVELOPMENT.

One interesting finding of our microarray analysis is the discovery of *PjYUC3* as a transcriptionally up-regulated gene in haustorium development. In *Arabidopsis AtYUC* genes encode flavin monooxygenases that catalyze the synthesis of IAA from tryptophan (Trp). IAA is a key regulator for various developmental processes, and as expected, the plants lacking multiple *YUCCA* genes show severe deficiencies in plant growth and development (Tobeña-Santamaria et al., 2002). For example, the *Arabidopsis yuc1 yuc4 yuc10 yuc11* quadruple mutants do not form basal embryonic regions (Cheng et al., 2006), and the *yuc1 yuc4* double mutants exhibit deficiencies in floral organs (Cheng et al., 2006). *AtYUC5*, *AtYUC8*, *AtYUC9*, *AtYUC7* and *AtYUC3* are particularly responsible for root development since the quintuple mutants show seriously disturbed root growth (Cheng et al., 2006). Here the knocking down the expression of *PjYUC3* did not affect root growth or its external morphology; phenotype was limited to reduction of haustorium formation. Expression and promoter-reporter gene fusions indicated that PjYUC3 is the unique *P.japonicum YUC* gene transcriptionally up-

regulated in haustorium development, which led me to hypothesized that *PjYUC3* gene may have acquired a new role associated with the haustorium formation in parasitic plants

The formation of newly haustoria organ is not randomicaly in parasitic plant roots, the area just behind the elongation zone is more sensitive for haustoria induction. The underlying morphological events leading to haustorium development suggested that selection of specific cells predetermined spatial localization of haustorium in parasite roots. Here, I applied a combination of expression and functional analysis to show that PjYUC3 was responsible for synthesizing *de novo* auxin which accumulated at DMBQ-induced haustoria. The auxin accumulation was limited to epidermal and cortical cells in the region just behind root tip, indicating that those cells were more permissive to sense the host-derived molecules than other parts of the root. It is possible that the localized *de novo* synthesis of auxin defines its gradients and cell-specific accumulations to further promote cell division and differentiation along haustorium formation.

The involvement of auxin accumulation during the DMBQ-induced haustoria is not new, it was already described for *T. vesicolor* (Tomilov et al., 2005) and *Phelipanche aegyptiaca* (Bar-Nun et al., 2008). Both reports suggested that auxin transported from proximal area of haustorium initiation site is key factor for haustorium formation. In this study, I complemented these previous analyses showing that a newly synthesized auxin mediated by PjYUC3 accumulated at specific cells in prehaustoria site. Potentially a combination of transported and *de novo* synthesized auxin might be relevant for induction of haustorium formation. The localized auxin accumulation is also the case for determining the position of newly lateral roots along the primary root (Moreno-Risueno et al., 2010). Presence of auxin in primary root is not uniform and the periodic ramification of roots preceded by temporary auxin accumulation at prebranch sites oscillates.

4.5 CONCLUSION

Here, using promoter-reporter gene fusions and reverse transcription-polymerase chain reaction, I could show that PjYUC3 is specifically regulated in plant parasitism and its expression is restricted to meristematic region and prehaustoria site. Knocking down the expression of *PjYUC3* resulted in a reduction of haustoria formation. *PjYUC3* encodes a functional auxin biosynthesis enzyme and its expression co-localized with local auxin accumulation at prehaustoria site. Although auxin is known to be involved in haustorium development in *T. vesicolor*, my results suggested that newly synthesized auxin catalyzed by PjYUC3 was accumulated at early stages of haustorium development.

In summary my main conclusions are:

(1) 1577 were genes differentially expressed in DMBQ-treated roots
(2) Large amount of sequences activated by parasitic infection were unknown.
(3) On first 3 h of haustorium development occurred induction of defense-related WRKY factors.
(4) 48 h along haustoria development occurred intense modification of cell shape structure with suppression of genes involved in lignin biosynthesis pathway, actin-bundling protein, and overexpression of cell-wall-modifying enzymes.
(5) Local auxin biosynthesis at apical root meristem and at haustorium initiation site mediated by *PjYUC3* is essential for haustorium development in *P. japonicum*.

4.6 TABLE AND FIGURES

Table 4.1 – List of representative DMBQ-induced genes divide into Cluster 1

Probe_ID	Cluster	0.5 h	1 h	3 h	6 h	12 h	24 h	48 h	GO term	Description
CUST_44793_PI426107926	1	-2.036	-2.610	-3.138	-2.623	-2.697	-2.025	-	0	calmodulin-binding protein-related
CUST_6948_PI426107926	1	-3.623	-4.083	-4.871	-2.878	-3.663	-2.541	-	0	calmodulin-binding protein-related
CUST_6120_PI426107926	1	-	-	-	-2.431	-	-	-	DNA or RNA binding	AGD15 (ARF-GAP DOMAIN 15); DNA binding
CUST_33727_PI426107926	1	-	-	-	-9.984	-10.550	-2.472	-	hydrolase activity	glycosyl hydrolase family 1 protein
CUST_1693_PI426107926	1	-	-2.898	-3.200	-	-	-	-	hydrolase activity	ARABIDOPSIS DYNAMIN-LIKE 2
CUST_30158_PI426107926	1	-	-	-6.365	-4.514	-	-	-	kinase activity	protein kinase family protein
CUST_20694_PI426107926	1	-3.319	-	-	-4.820	-3.344	-	-	kinase activity	leucine-rich repeat transmembrane protein kinase
CUST_19168_PI426107926	1	-	-	-	-4.975	-4.242	-	-	kinase activity	leucine-rich repeat transmembrane protein kinase
CUST_35747_PI426107926	1	-2.912	-2.138	-3.698	-2.546	-3.179	-	-	kinase activity	protein kinase, putative
CUST_19231_PI426107926	1	-	-	-2.768	-2.322	-2.910	-	-	nucleotide binding	RESISTANCE TO P. SYRINGAE PV MACULICOLA 1
CUST_37694_PI426107926	1	-	-	-2.790	-2.142	-	-	-	other enzyme activity	FERRIC REDUCTION OXIDASE 2
CUST_876_PI426107926	1	-	-	-2.517	-	-	-	-	other enzyme activity	FERRIC REDUCTION OXIDASE 2
CUST_44421_PI426107926	1	-2.599	-	-3.247	-2.929	-3.121	-	-	other enzyme activity	FERRIC REDUCTION OXIDASE 2
CUST_39743_PI426107926	1	-	-	-	-	-2.264	-	-	other molecular functions	allergen V5/Tpx-1-related family protein
CUST_21018_PI426107926	1	-3.203	-3.519	-4.167	-3.007	-3.098	-	-	other molecular functions	allergen V5/Tpx-1-related family protein
CUST_18025_PI426107926	1	-	-	-	-2.327	-	-	-	other molecular functions	dirigent-like protein (TAIR:AT5G42500.1)
CUST_18054_PI426107926	1	-	-2.953	-3.734	-	-2.113	-	-	other molecular functions	dirigent-like protein (TAIR:AT5G42500.1)

CUST_23602_PI426107926	1	-	-	-	-	-3.300	-	-	other molecular functions	dirigent-like protein (TAIR:AT5G42500.1)
CUST_44905_PI426107926	1	-	-2.849	-4.242	-	-	-2.530	-	other molecular functions	dirigent-like protein (TAIR:AT5G42500.1)
CUST_31168_PI426107926	1	-	-	-3.616	-2.592	-	-	-	other molecular functions	similar to unknown protein with Villin headpiece domain
CUST_31680_PI426107926	1	-	-	-2.462	-	-	-	-	other molecular functions	similar to unknown protein with Villin headpiece domain
CUST_46814_PI426107926	1	-	-	-2.160	-	-	-	-	other molecular functions	similar to unknown protein with Villin headpiece domain
CUST_1405_PI426107926	1	-	-	-13.766	-4.237	-	-	-	other molecular functions	glyoxal oxidase-related
CUST_45872_PI426107926	1	-	-	-5.065	-5.265	-	-	-	other molecular functions	glyoxal oxidase-related
CUST_34518_PI426107926	1	-	-	-3.150	-3.721	-3.735	-2.389	-2.482	other molecular functions	glyoxal oxidase-related
CUST_31003_PI426107926	1	-	-	-	-	-	-2.455	-	other molecular functions	glyoxal oxidase-related
CUST_6720_PI426107926	1	-	-	-2.433	-2.023	-2.075	-	-	other molecular functions	MRH6 (morphogenesis of root hair 6)
CUST_33865_PI426107926	1	-2.956	-2.046	-3.337	-2.332	-2.551	-	-	other molecular functions	TAIR:AT1G79760.1
CUST_2763_PI426107926	1	-	-	-	-3.016	-4.155	-3.017	-2.366	other molecular functions	similar to Os03g0309000
CUST_12394_PI426107926	1	-2.230	-	-	-	-	-	-	protein binding	VLN1 (VILLIN 1); actin binding
CUST_19197_PI426107926	1	-	-	-	-4.711	-	-	-	transferase activity	glycogenin glucosyltransferase (glycogenin)
CUST_39874_PI426107926	1	-	-	-	-2.263	-2.496	-	-	transferase activity	GATL8/LGT9 (Galacturonosyltransferase-like 8)
CUST_9349_PI426107926	1	-	-	-12.937	-3.821	-	-2.126	-	transferase activity	agmatine coumaroyltransferase
CUST_16896_PI426107926	1	-	-	-	-6.110	-10.453	-	-	transporter activity	Nitrate transporter 2.1)

Table 4.2 – List of representatives DMBQ-induced genes divide into Cluster 2

Probe_ID	Cluster	0.5 h	1 h	3 h	6 h	12 h	24 h	48 h	GO term	Description
CUST_46224_PI426107926	2	-	-	30.252	60.750	129.498	92.663	20.499	-	unknown protein
CUST_40047_PI426107926	2	-	6.161	39.831	60.058	79.655	82.948	38.147	-	unknown protein
CUST_44178_PI426107926	2	-	-	16.869	20.412	-	14.384	-	-	unknown protein
CUST_1315_PI426107923	2	-	13.964	-	-	-	13.130	-	-	unknown protein
CUST_45134_PI426107926	2	-	-	4.912	6.881	8.481	12.194	7.239	-	unknown protein
CUST_1458_PI426107923	2	-	4.465	11.939	-	-	-	-	-	unknown protein
CUST_17044_PI426107926	2	-	11.368	-	-	-	-	-	-	unknown protein
CUST_1966_PI426107923	2	-	4.945	4.955	-	-	-	-	-	unknown protein
CUST_2051_PI426107923	2	-	-	3.712	-	-	-	-	-	unknown protein
CUST_3597_PI426107923	2	-	6.069	2.770	-	-	-	-	-	unknown protein
CUST_7120_PI426107923	2	-	5.485	3.641	-	-	-	-	-	unknown protein
CUST_38354_PI426107926	2	7.277	-	-	43.386	-	-	-	-	unknown protein
CUST_7048_PI426107923	2	7.114	4.659	3.193	-	-	-	-	-	unknown protein
CUST_3985_PI426107923	2	5.657	-	-	-	-	-	-	-	unknown protein
CUST_11438_PI426107926	2	-	8.419	38.152	24.769	19.780	-	8.321	hydrolase activity	RESPONSIVE TO DEHYDRATION 19
CUST_17498_PI426107926	2	-	-	12.906	81.183	220.486	-	184.659	hydrolase activity	ARAB-1 (Arabidopsis lipase)
CUST_13350_PI426107926	2	-	3.038	2.974	-	-	-	-	kinase activity	FLAGELLIN-SENSITIVE 2
CUST_36557_PI426107926	2	4.901	3.388	-	-	-	-	-	kinase activity	RECEPTOR-LIKE PROTEIN KINASE 1
CUST_46945_PI426107926	2	-	2.316	-	-	-	-	-	nucleotide binding	RESISTANCE TO P. SYRINGAE PV MACULICOLA 1
CUST_18852_PI426107926	2	16.018	41.061	95.908	67.751	43.912	29.931	17.290	other binding	QUINONE REDUCTASE 2
CUST_7233_PI426107926	2	-	-	3.761	4.046	3.484	-	3.682	other binding	pentatricopeptide repeat-containing protein
CUST_18780_PI426107926	2	-	2.523	-	-	-	-	-	other binding	CAM8 (CALMODULIN 8)
CUST_41112_PI426107926	2	3.069	5.811	4.160	2.109	-	-	-	other enzyme activity	pyridine nucleotide-disulphide oxidoreductase
CUST_40645_PI426107926	2	24.796	63.343	52.226	26.291	17.302	14.374	9.234	protein binding	pirin, putative
CUST_40538_PI426107926	2	18.876	45.629	62.841	29.322	18.282	13.421	9.817	protein binding	pirin, putative
CUST_37862_PI426107926	2	-	4.882	-	-	-	-	-	protein binding	MLO2 (MILDEW RESISTANCE LOCUS O 2)

CUST_5769_PI426107926	2	-	2.237	2.135	-	-	-	-	protein binding	RIN4 (RPM1 INTERACTING PROTEIN 4)
CUST_42275_PI426107926	2	4.207	-	-	2.329	-	-	-	protein binding	MLO4 (MILDEW RESISTANCE LOCUS O 4)
CUST_32052_PI426107926	2	-	-	2.652	-	-	-	-	transcription factor activity	WRKY33
CUST_14411_PI426107926	2	2.996	3.749	3.627	2.585	-	-	-	transcription factor activity	WRKY23
CUST_27495_PI426107926	2	-	-	5.813	-	-	-	-	transcription factor activity	basix helix-loop-helix (bHLH) family protein
CUST_18350_PI426107926	2	-	3.426	-	-	-	-	-	transcription factor activity	WRKY33
CUST_21846_PI426107926	2	-	7.288	-	-	-	-	-	transcription factor activity	WRKY40
CUST_23937_PI426107926	2	-	2.532	-	-	-	-	-	transcription factor activity	WRKY23
CUST_35027_PI426107926	2	-	2.560	-	-	-	-	-	transcription factor activity	WRKY23
CUST_20810_PI426107926	2	-	13.796	23.296	7.079	-	-	-	transferase activity	GLUTATHIONE S-TRANSFERASE 11
CUST_20941_PI426107926	2	-	-	4.133	-	-	-	-	transporter activity	PGP11 (P-GLYCOPROTEIN 11)

Table 4.3 – List of representative DMBQ-induced genes divide into Cluster 3

Probe_ID	Cluster	0.5 h	1 h	3 h	6 h	12 h	24 h	48 h	GO term	Description
CUST_25696_PI426107926	3	-	-	8.360	18.770	-	-	-	hydrolase activity	glycoside hydrolase family 28
CUST_21266_PI426107926	3	-	3.059	-	-	-	-	-	hydrolase activity	glycosyl hydrolase family 1
CUST_3993_PI426107926	3	-	-	2.067	-	-	-	-	hydrolase activity	glycosyl hydrolase family 3
CUST_11316_PI426107926	3	-	-	-	-	2.038	-	-	hydrolase activity	glycosyl hydrolase family 79
CUST_30581_PI426107926	3	-	-	-	2.060	2.691	-	-	hydrolase activity	pectinesterase family protein
CUST_16111_PI426107926	3	-	2.709	2.545	-	-	-	-	hydrolase activity	pepsin A
CUST_6491_PI426107926	3	-	-	-	2.717	-	-	-	hydrolase activity	pectinase
CUST_6911_PI426107926	3	-	-	-	2.542	-	-	-	hydrolase activity	pectinase
CUST_37346_PI426107926	3	-	-	3.071	2.304	-	5.944	-	hydrolase activity	glycoside hydrolase family 28
CUST_26963_PI426107926	3	-	-	-	-	3.644	4.382	3.275	hydrolase activity	glycosyl hydrolase family 3
CUST_15557_PI426107926	3	-	-	2.390	-	-	-	-	kinase activity	WAK4 (WALL ASSOCIATED KINASE 4)
CUST_33571_PI426107926	3	-	-	5.888	-	6.704	4.410	3.024	kinase activity	PGIP1 (POLYGALACTURONASE INHIBITING PROTEIN 1)
CUST_13097_PI426107926	3	5.937	8.198	7.669	5.082	3.599	2.762	2.278	kinase activity	YDA (YODA); kinase
CUST_38364_PI426107926	3	-	-	2.518	2.479	2.566	2.251	2.055	kinase activity	SUB (STRUBBELIG)
CUST_17683_PI426107926	3	-	-	7.523	4.266	3.463	-	-	other binding	zinc finger (C2H2 type) family protein
CUST_33398_PI426107926	3	4.105	8.388	8.276	5.556	2.821	-	-	other binding	cytochrome P450
CUST_25752_PI426107926	3	-	-	-	-	-	4.250	3.779	other binding	GLP4 (GERMIN-LIKE PROTEIN 4)
CUST_12924_PI426107926	3	4.267	-	4.261	5.186	2.711	2.960	-	other binding	CAM4 (CALMODULIN 4)
CUST_6580_PI426107926	3	-	-	-	15.803	-	8.507	-	other enzyme activity	ATGA2OX8 (GIBBERELLIN 2-OXIDASE 8)
CUST_21965_PI426107926	3	-	-	3.481	3.328	4.140	4.386	2.037	other enzyme activity	YUCCA
CUST_45132_PI426107926	3	5.311	8.387	9.726	5.357	4.459	3.655	2.462	other enzyme activity	ATCM2 (CHORISMATE MUTASE 2)
CUST_30762_PI426107926	3	-	-	-	2.516	3.268	2.874	2.461	other enzyme activity	CKX3 (CYTOKININ OXIDASE 3)
CUST_45384_PI426107926	3	-	-	2.310	-	2.715	2.659	2.038	other enzyme activity	ACC oxidase
CUST_17538_PI426107926	3	2.833	5.286	4.748	3.679	3.080	2.412	-	other enzyme activity	pyridine nucleotide-disulphide oxidoreductase
CUST_43527_PI426107926	3	- 2.476	-	-	-	2.564	2.332	-	other enzyme activity	YUCCA
CUST_17818_PI426107926	3	-	-	-	2.492	-	-	-	other molecular functions	peroxidase
CUST_40323_PI426107926	3	-	-	-	-	3.873	-	-	other molecular functions	peroxidase
CUST_6684_PI426107926	3	3.979	4.436	6.386	3.176	-	-	-	other molecular functions	unknown protein (TAIR:AT1G58420.1)
CUST_4476_PI426107926	3	-	3.171	2.882	3.256	-	-	-	other molecular functions	unknown protein (TAIR:AT1G69510.2)
CUST_12531_PI426107926	3	-	-	-	-	2.496	4.928	3.231	other molecular functions	PEROXIDASE 33
CUST_31659_PI426107926	3	-	-	-	3.044	3.648	3.791	2.294	other molecular functions	peroxidase, putative
CUST_18247_PI426107926	3	-	-	-	-	2.421	2.655	2.345	other molecular functions	unknown protein
CUST_34400_PI426107926	3	-	-	5.630	-	-	-	-	transcription factor activity	ATERF10/ERF10 (ERF domain protein 10)

CUST_36413_PI426107926	3	-	-	6.058	-	-	-	-	transcription factor activity	AP2 domain-containing transcription factor
CUST_24123_PI426107926	3	-	-	13.652	6.929	6.476	-	-	transcription factor activity	BHLH093 (BETA HLH PROTEIN 93)
CUST_31948_PI426107926	3	-	-	-	6.824	4.523	-	-	transcription factor activity	ABR1 (ABA REPRESSOR1)
CUST_25859_PI426107926	3	-	-	19.455	9.645	8.014	5.573	3.193	transcription factor activity	BHLH093 (BETA HLH PROTEIN 93)
CUST_37051_PI426107926	3	-	-	-	4.486	6.755	-	5.733	transferase activity	ATGSTU18 (GLUTATHIONE S-TRANSFERASE 29)
CUST_15569_PI426107926	3	-	7.355	3.530	2.289	-	-	-	transferase activity	AtSerat2;1 (SERINE ACETYLTRANSFERASE 1)
CUST_17858_PI426107926	3	-	-	7.864	-	-	-	-	transferase activity	Glutathione S-transferase
CUST_4330_PI426107926	3	4.043	5.283	3.592	3.305	3.930	2.937	2.592	transferase activity	UGT73B5 (UDP-glucosyl transferase 73B5)

Table 4.4 – List of unaltered expressed genes with homology of Arabidopsis reference genes

ProbeName	Best_Blast_Hit	Description
CUST_11046_PI426107926	gnl|Vitvi1|GSVIVP00020992001	PTB (POLYPYRIMIDINE TRACT-BINDING)
CUST_3782_PI426107926	gnl|Vitvi1|GSVIVP00031192001	pentatricopeptide (PPR) repeat-containing protein
CUST_17482_PI426107926	gnl|Poptr1|666735	UBC27 (ubiquitin-conjugating enzyme 26)
CUST_4987_PI426107926	gnl|Vitvi1|GSVIVP00021722001	UPL7 (ubiquitin-protein ligase 7)

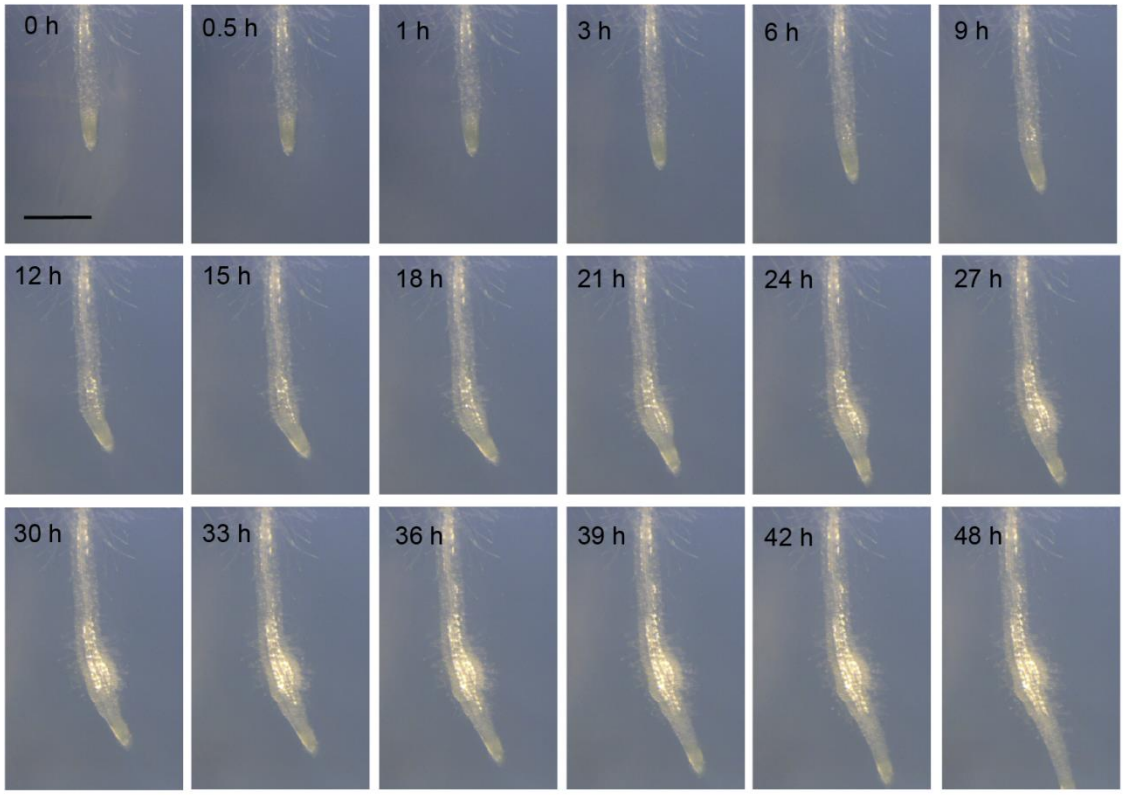


Figure 4.1 - Time-lapse images of haustorium development in response to rice root treatment. Agar 0.8% (w/v) containing 10µM DMBQ were placed on *P. japonicum* root tip, the photographs were taken in series every 30 minutes until complete 48h. Bars correspond to 500µm.

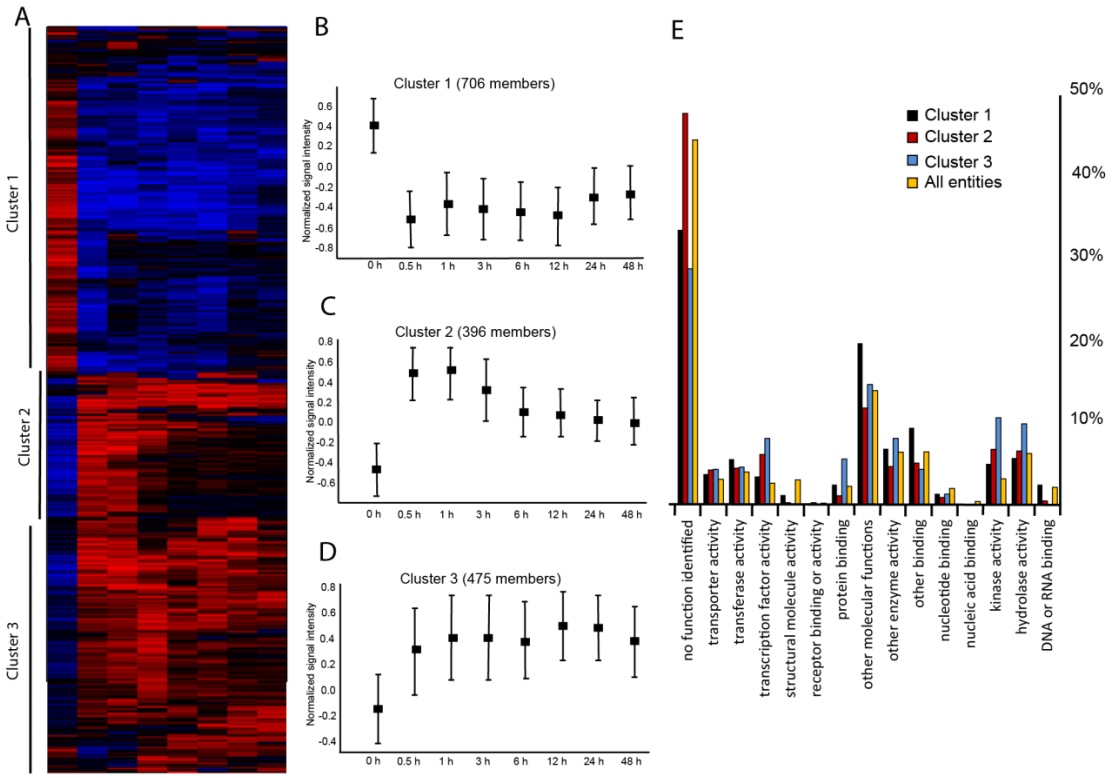


Figure 4.2 – Cluster analysis of 1577 genes regulated specifically by DMBQ. (**A)** *P. japonicum* genes were hierarchically clustered based on similarity of their expression profiles using the program CLC Genomics Workbench 4.8. Fold changes of expression level in DMBQ-treated versus non-treated roots (time 0 h) is represented by colors (red – upregulated; black – no regulation and blue – downregulated). (**B)** Expressed sequences organized by Self-Organization Map (SOM) using the program Gene Pattern 2.0. Cluster 1 contains 706 genes whose the expression is under regulated after the addition of DMBQ. Cluster 2 has 396 genes whose the peak of expression occurs before 3 h and cluster 3 grouped 475 genes whose the peak of expression occurs after 3 h of DMBQ addition. (**C)** Gene Ontology terms in molecular process category among the differentially expressed genes. The details about the clustering members are listed in Supplemental Table S4.

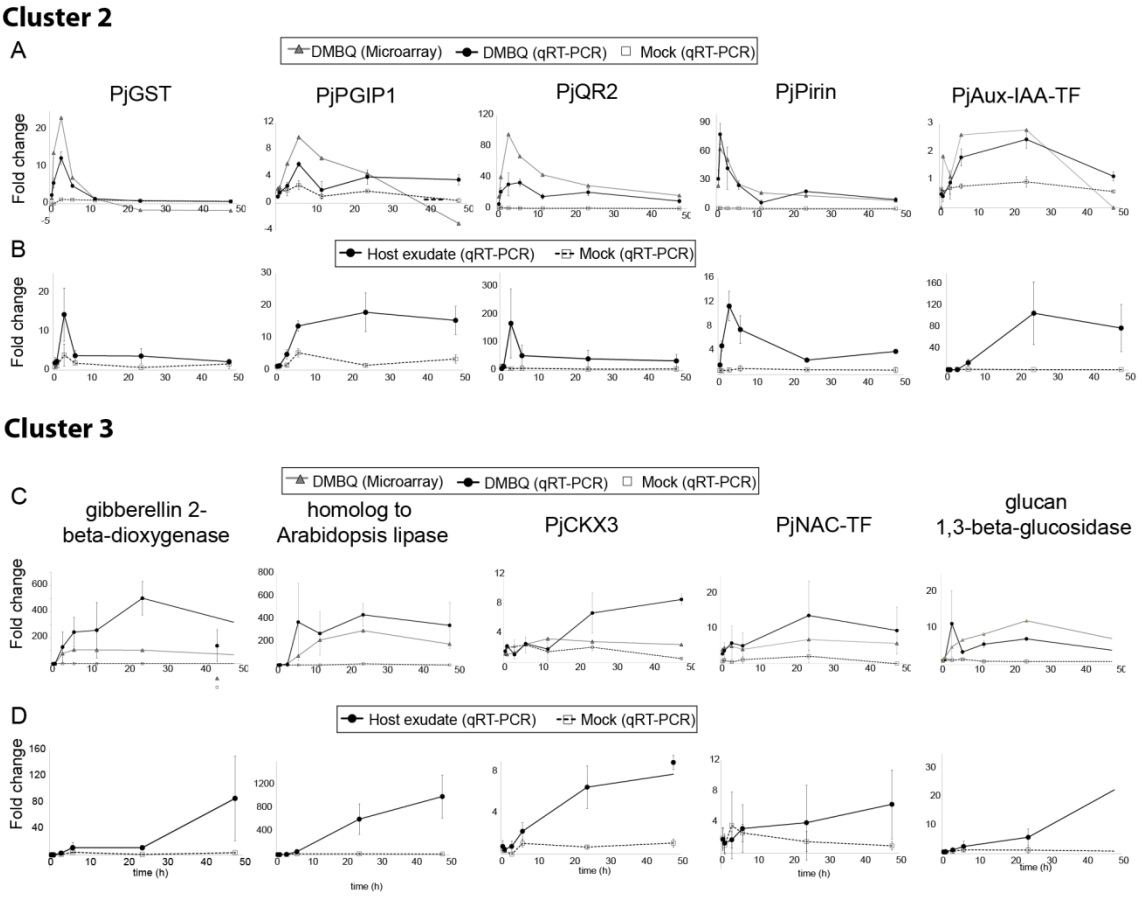


Figure 4.3. Quantitative reverse-transcriptase PCR validation of expression profiles of genes. Data are shown as fold of induction from treated roots versus non-treated roots at 30 min, 1 h, 3 h, 6 h, 24 h and 48 h of treatment period. 2 week-old *P. japonicum* plants were growing vertically to avoid the roots enter into the agar media, in this way the roots can be in contact with 10 µM DMBQ solution (**A** and **C**) or with rice root exudate (**B** and **D**). The genes from cluster 2 are represented in A-B, while the expression profile of transcript belonging to cluster 3 are in C-D. The values are the average of at least two biological replicates with three technical replicates each for DMBQ-treated samples and four biological replicates for treatment with rice exudate. Each experiment corresponds to RNA pool extracted from the roots of 20-30 plants divided in 2 square plate dishes. Primers specific for reference *PTB* gene was used for normalization

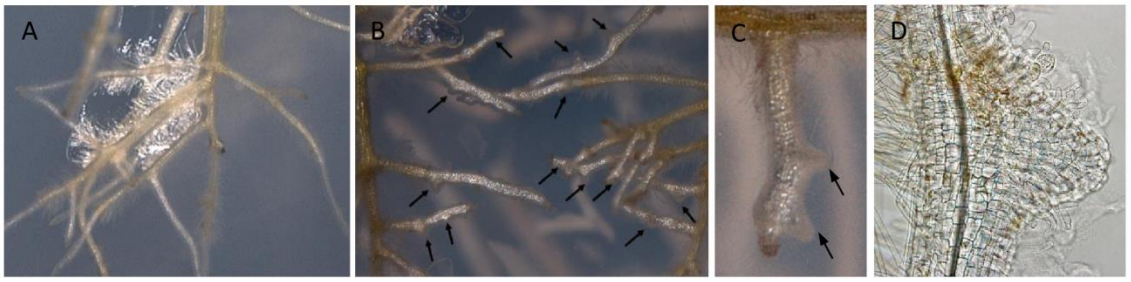


Figure 4.4. *P. japonicum* roots at 48h after rice root exudate treatment. (**A**) No-transformed root in normal growth condition and (**B**) 48 h-treated with rice root exudate. **(C-D)** Magnification of host-induced haustoria in non-transformed root.

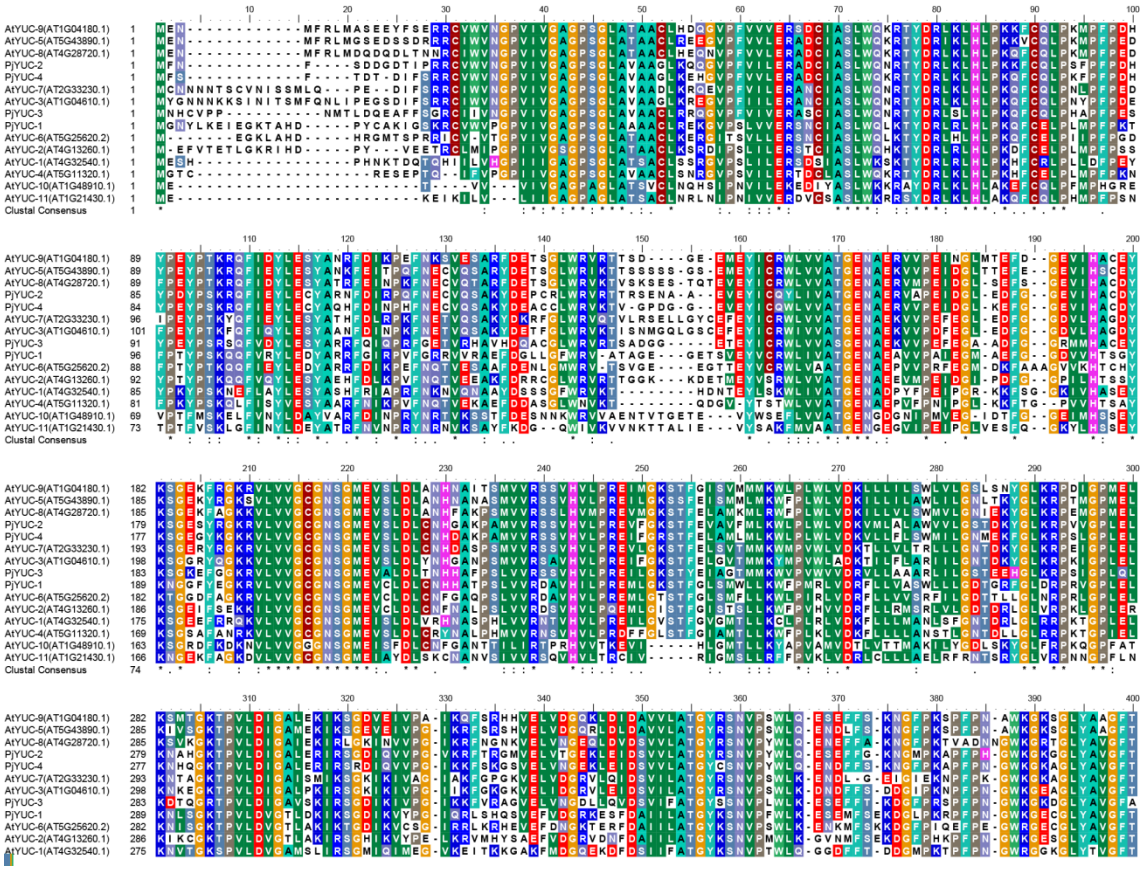


Figure 4.5. Alignment of *PjYUCs* and *AtYUC*s amino acid sequences. The sequences were aligned using CLC Genomics Workbench (v4.8) and the figure generated by BioEdit

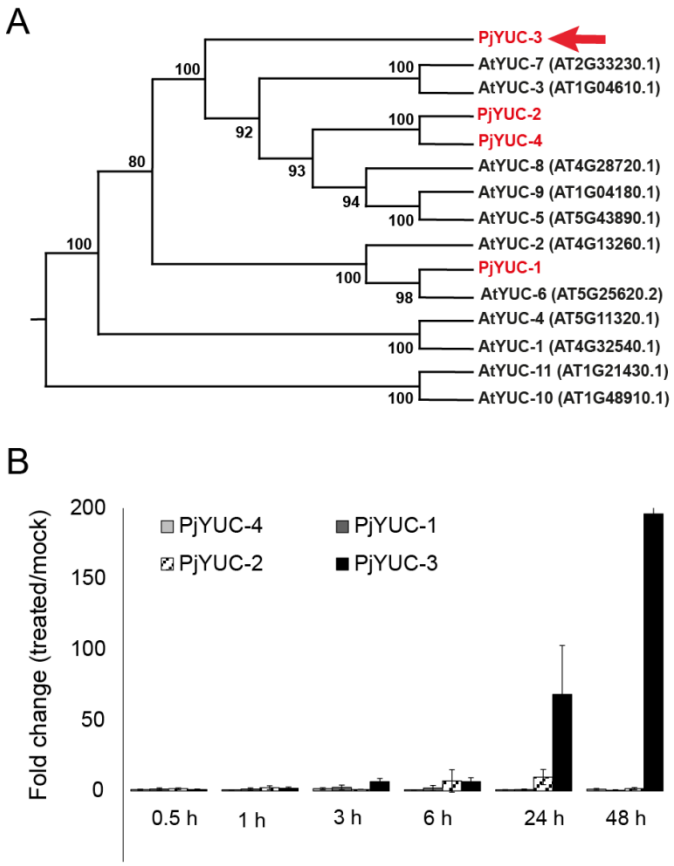


Figure 4.6 – Homologs of *YUCCA* genes in *P. japonicum* transcriptome. (**A**) Phylogenetic tree of all *AtYUC* genes and their homologs (*PjYUC*) found in root transcriptome. In red the *PjYUC* genes and pointed by the arrow *PjYUC3* that is differentially expressed in the microarray analysis. (**B**) Expression levels of *PjYUC3* homologs measured by qRT-PCR after host root exudate treatment (0 h, 0.5 h, 1 h, 3 h, 24 h and 48 h).

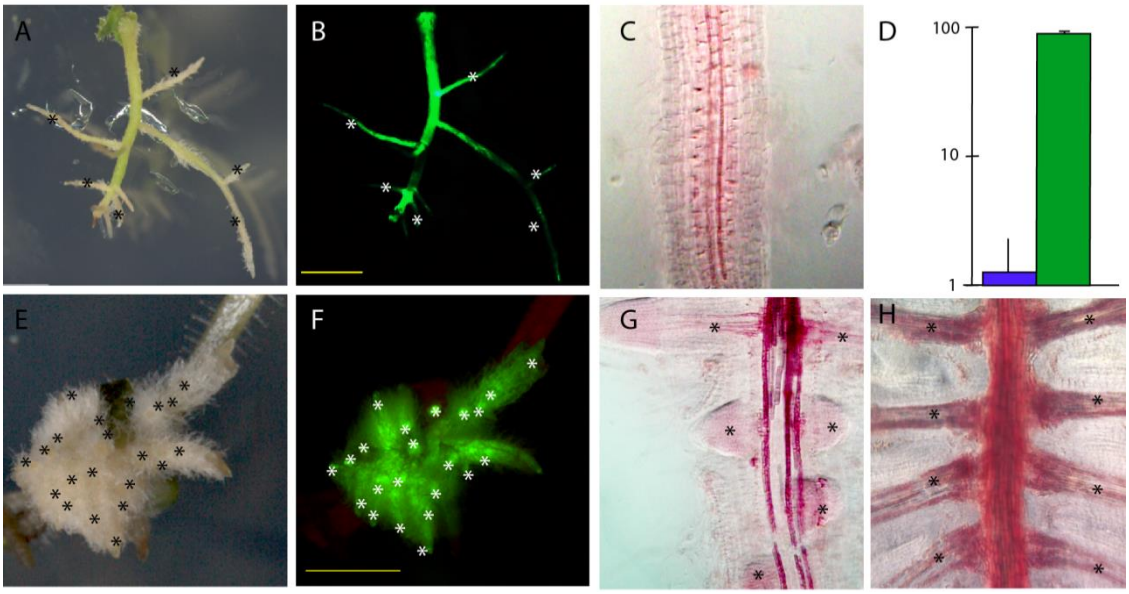


Figure 4.7 –Overexpression of *PjYUC3* in transgenic *P. japonicum* roots showed auxin-overexpressing phenotype. (**A**-**C**) Transgenic roots carrying 35S promoter upstream of *GFP* gene in bright field (A), under fluorescence microscopy (B) and stained with safrarin (C). (**D**). Expression levels of *PjYUC3* in non-transformed roots and in overexpressed root analyzed by qRT-PCR. Data are presented as means with Standard error of two technical replicates of two biological replicates containing 5 to 10 roots in each experiment. (**E**-**H**) Phenotype of transgenic roots overexpressing *PjYUC3* in bright field (E), under fluorescence microscopy (F) and stained with safrarin (G-H). Asterisks mark the lateral roots. Bars correspond to 2 mm

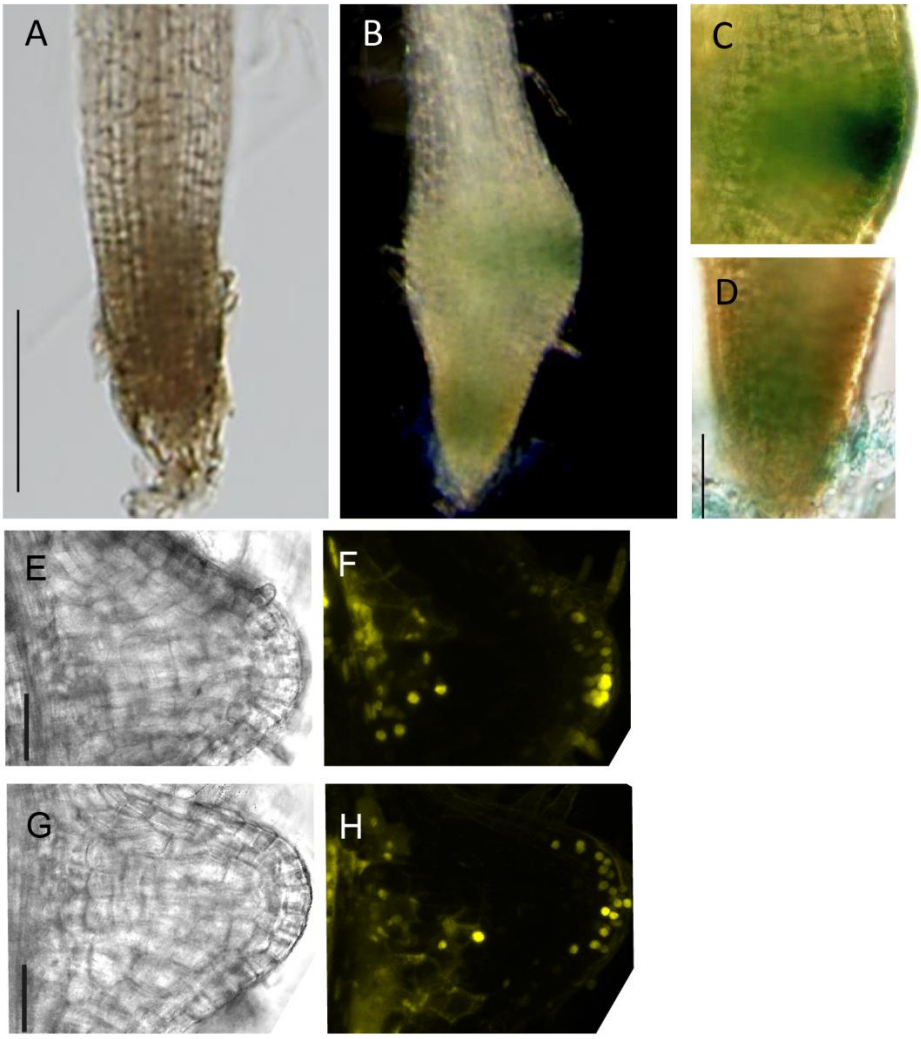


Figure 4.8 – Site of *PjYUC3* expression and Auxin accumulation coincide during haustorium development. Transgenic root carrying *PjYUC3* promoter region upstream to *GFP* and *GUS* reporter genes. (**A**) no-treated root. (**B-D**) Root after 24 h of rice root exudate treatment. Bar corresponds to 250 µm. **(E-H)** Auxin localization at DMBQ-induced haustorium. Confocal microscopic photo of transgenic roots carrying the construct of VENUS, a fast maturing form of yellow fluorescent protein (YFP), fused to a nucleus targeting signal. The reporter gene is driven by auxin-specific promoter DR5 (DR5rev::3xVenus-N7). (**E-F**) transformed root 40 h after the application of 10 µM DMBQ in bright field (E) and under fluorescence (F). (**G-H**) 48 h after the application of 10 µM DMBQ in bright field (G) and under fluorescence (H). The pictures corresponding to stacked of successive photographs taken with intervals of 5 µm of thickness. Arrows point to the auxin accumulation at the haustorium initiation site. Bar corresponds to 50 µm. Pictures of E to H were taken by Takanori Watakake.

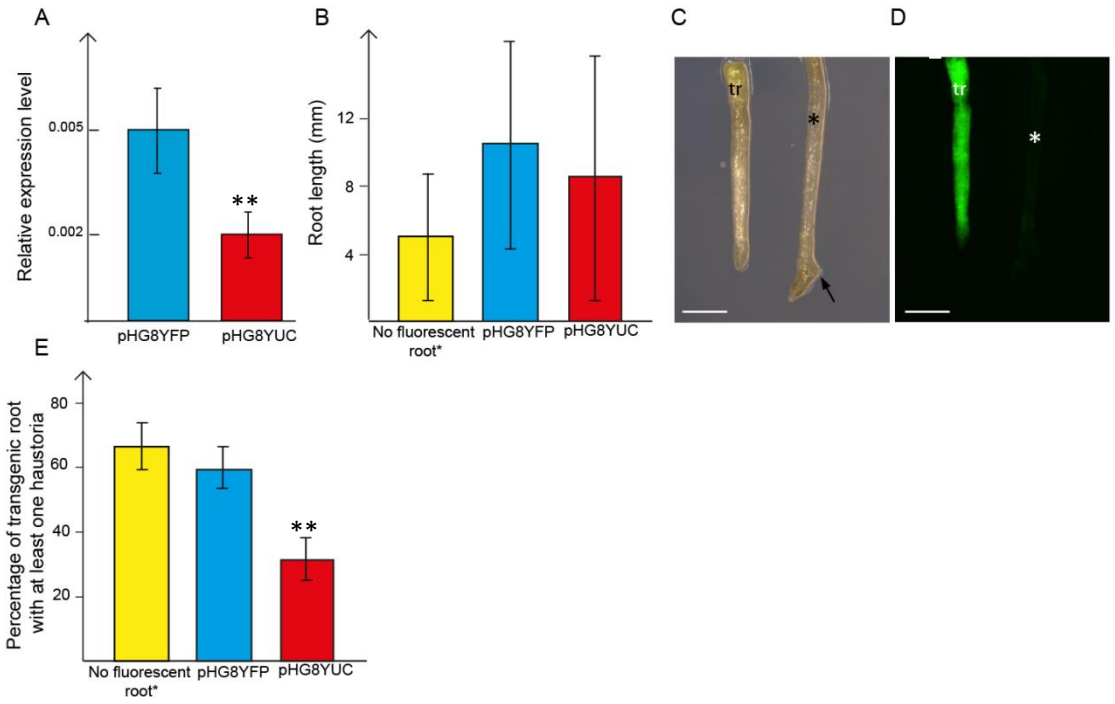


Figure 4.9. *PjYUC3* is essential for haustorium development. (**A**) *PjYUC3* transcript levels were measured by qRT-PCR in transgenic root carrying silencing construct targeting *PjYUC3* gene (red - pHG8-YUC) and in roots transformed with empty vector (blue - pHG8-YFP), after 24h of 10 µM DMBQ treatment. Data correspond to three biological replicates (±SD); each experiment contains a pool of at least 12-15 independently transgenic roots. (B) Phenotypic analysis of transgenic roots. (**B-C**) Phenotypic ana;ysis of transgenic roots. (B) Number of transgenetic root with at least one haustoria. (C) Root length measured in millimeters. No fluorescent root means those transgenetic hairy root without the T-DNA containing the *yellow fluorescent protein* gene. Double asterisk (**) means α = 0.01 by t-Test assuming unequal variances. Values are means ±SD of four to six independently experiment with 10 to 40 transformed roots each experiment. Total number of analyzed roots in all the biological replicates were as following: 99 no fluorescent roots, 90 pHG8-YFP and 89 pHG8YUC. (**D – E)** Morphologic aspect of a transformed hairy roots. Roots transformed with empty vector pHG8-YFPused as control (left) and no-fluorescent hairy root (right) in bright field (D) and under fluorescence microscopy (E). Fluorescent root silenced for *PjYUC3* gene (left) without haustorium and no-fluorescent hairy root (right) with a formed haustorium in bright field (G) and under fluorescence microscopy (D). Black arrows point to the haustoria and the asterisks mark no-fluorescent hairy root. Co: cotyledon; Tr: transformed root. Bars correspond to 1 mm.

5

CHAPTER
DEVELOPMENT AND CHARACTERIZATION OF SSR MARKERS FOR *P. JAPONICUM*

5.1 SUMMARY

De novo transcriptome of *P. japonicum* was used to identify 1,275 putative simple sequence repeats (SSRs). The repeated sequences were mainly found at putative untranslated region (UTR), with exception of tri- type, which were mainly found at putative coding sequences. The most frequent motif was the mono- type A/T (16.15%). To check if the SSRs were informative markers, 67 primer pairs were designed and used to amplify short sequences from genomic DNA extracted from *P. japonicum* collected at seven different locations in Japan. Among them 12 SSRs were informative markers and showed that Hiroshima-Okayama and Gunma-Nagano populations were genetically close and the group from Kyushu was genetically furthest, the data corroborating with their geographic positions.

5.2 INTRODUCTION

In addition to the high throughput functional genomic analysis, the identification of genetic patterns through molecular markers provides useful resources to combat the parasitic plant problem. Molecular markers scan can also inform the population history, their propensity for gene flow and genetic diversity among different population of parasitic plants. Simple Sequence Repeat (SSR) or microsatellites are tandem repetition sequence of 1 - 6 nucleotides and due to their high mutation ratio per locus per generation SSRs represent a useful tool for genetic analysis within individual genotypes (Enalik et al., 2012). The emergency of next-generation sequencing technology increased the availability of informative SSRs for innumerous plants (Enalik et al., 2012). SSR markers are useful for several applications, including genetic diversity and gene flow analyses. As *P. japonicum* is emerging as a model plant for plant parasitism studies, we used the assembled transcriptome to discovery simple sequence repeat (SSR) loci for enriching its genetic information.

5.3 RESULTS AND DISCUSSION

Using as basis *de novo* transcriptome 1,275 SSRs were identified, and among them 1,223 are located as single motifs per sequence and 25 are present in compound form (two or more motifs in interrupted or uninterrupted arrays) (Table 5.1). Most of the SSRs are in the putative untranslated region (UTR), with exception of tri- type, which are mainly found in coding sequences (Fig. 5.1). The most frequent individual motif was mono- type A/T (16.15%), followed by TAA/TTA (including ATT and AAT) (13.8%) and GA/TC (including CT and AG) (9%)(Table 5.2). To investigate if the identified SSRs are informative markers we designed 67 primer pairs in the flaking region of repeated motifs, taking into account that higher polymorphism rates are often associated with shorter motifs and higher number of the repeats (Enalik et al., 2012). As the PCR template we used genomic DNA extracted from *P. japonicum* populations collected at different locations in Japan (Fig. 5.2-A). Among 67 primer pairs tested, 12 amplified a band with size larger than expected, 18 were no-polymorphic and 6 amplified multiple bands. Of 31 working primer pairs, 12 showed reproducibility using two different techniques for band- size analysis, the microchip electrophoresis system MultiNa which allows the separation of bands smaller than 5 bp and the traditional gel-basedelectrophoresis. The 12 selected markers detected 23 alleles with average of 2.875 alleles per locus. The identification of these markers revealed the genetic relationships of different *P. japonicum* populations. *P. japonicum* from Hiroshima and Okayama were genetically closer compared with plants collected in other prefectures, which corroborates with their close geographic locations. The farthest group is from Kyushu which is localized in an island in the Japanese archipelago. The other genetically close group is composed by plants collected in Gunma and Nagano prefectures, reflecting their close geographic positions (Fig. 5.2-B). Thus, the assembled transcriptome allowed us to identify informative SSR markers, providing the first study of genetic relationships among different population of the parasitic plant *P. japonicum* in Japan

5.4 TABLES AND FIGURES

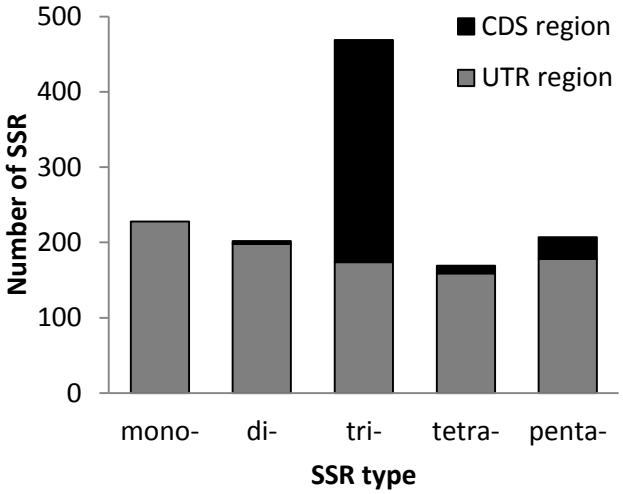


Figure 5.1 – Abundance of SSRs in non-coding and coding region. The distribution of SSRs in untranslated region (UTR) and in the coding region (CDS) according with their motifs sizes

Table 5.1 – Distribution of SSRs detected in *P. japonicum* transcriptome

Parameter	Number
Total SSR identified	1275
sequences containing one SSR	1223
sequences containing two SSRs	23
sequences containing three SSRs	2
Total contigs containing at least one SSR	1248

Table 5.2 – Frequency of different motifs among the SSR in *P. japonicum* transcriptome

SSR Motif	Number
A/T	206
TAA/TTA	176
GA/TC	115
ATC/GAT	87
AAAT/ATTT	58
AC/GT	56
CAA/TTG	49
GAA/TTC	45
ACC/GGT	37
AT/TA	31
CAG/CTG	28
AAAG/CTTT	24
C/G	21
AACA/TGTT	18
CATA/TATG	16
ATTA/TAAT	15
ACT/AGT	14
CCG/CGG	14
CTC/GAG	12
AAAC/GTTT	8
GAC/GTC	7
AGAT/ATCT	7
GTGA/TCAC	5
ATTC/GAAT	4
GTAG/CTAC	2
GTCA/TGAC	2
GTTG/CAAC	2
TTGA/TCAA	2
TGCG/CGCA	2
CTCC/GGAG	2
TGCA/TGCA	1
GTCG/CGAC	1
GCTT/AAGC	1
penta-	207
Total	1275

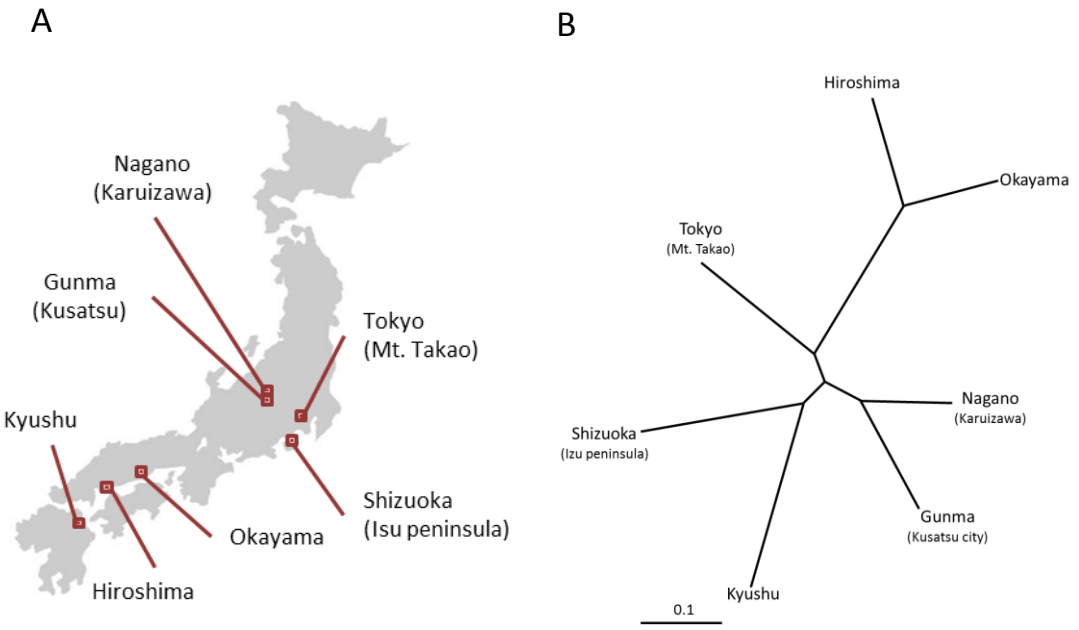


Figure 5.2 – SSR-based genetic analysis of different *P. japonicum* population. (A) *P. japonicum* populations were collected in different locations across Japan (Kyushu, Hiroshima, Okayama, Gunma, Nagano, Shizuoka and Tokyo) indicated by red crosses in the map. In parenthesis is specified the name of the area where the plants were collected. **(B)** Neighbor joining-based phylogeny tree derived from genotype data from 12 SSRs loci.

6

CHAPTER
TRANSCRIPTOMIC ANALYSIS IN HOST-PARASITE INTERACTION

6.1 SUMMARY

Deep-sequencing technologies have revolutionized transcriptome analyses to obtain high-coverage sequence reads of non-sequenced organims in economically feasible prices. In this chapter I describe a transcriptome analysis during the interaction of the parasitic plant *P. japonicum* with its host rice. I identified regulated genes in the parasitic as well as in the host tissues after 1 or 7 days of host-parasite interaction. At the first step I generated Illumina Hi-Seq sequences and *de novo* assembled the *P. japonicum* transcriptome. The transcript levels were determined by mapping those reads against the assembled transcriptome. The rice transcripts were mapped against available rice cDNAs and expression levels were accessed. Finally, statistical analysis identified 2917 parasite and 1155 rice genes regulated during parasitism. These genes were divided into nine clusters according to their expression profiles. Co-expression analysis identified the hub genes which are highly co-expressed with other genes. Among them, the hub transcription factor from parasitic tissues CRF was highly up-regulated. CRF is a member of AP2/ERF transcriptional factor family. In both parasitic and host tissues, up-regulation of cell wall-modifying enzymes were observed, inferring intense modification of cell shape during the establishment of plant parasitism. To identify the rice genes similarly regulated during its interaction with *P. japonicum* and *S. hermonthica*, I compared the RNA-Seq data described in this chapter with published microarray data from *S. hermonthica*-interacting rice roots. This analysis revealed that the rice homeobox-leucine zipper transcription factor, Oshox16, is up-regulated after interaction with both parasitic plants. Taken together, the genes relevant to plant parasitism, which include two transcription factors, were identified in parasitic and host tissues.

6.2 INTRODUCTION

The first barrier that a parasitic plant has to overcome to invade host tissues is rigid cell wall of host cells. During the haustorium development, host and the parasite cells needs to be modified. One of non-enzymatic proteins that participate in this process is Expansin; the recruitment of different types of Expansin during host infection was shown in the facultative parasitic plant *T. vesicolor* that can infect wide range of host speceis. When this parasite is infecting a Poaceae the β-type Expansin is up-regulated. In contrast, when the host is a dicot the α-type Expansin is highly expressed (Honaas et al., 2013). Such difference reflects the distinct activities of α- and β-Expansins in higher plants. β-Expansin specifically acts in cell wall of grasses, whereas α-Expansin expands cells in dicot plants (Cosgrove, 2005). Thus understanding of heterogeneity of cell wall structures may bring insights for how a parasitic plant recognizes its host. In general, cell wall is formed by a network composed by cellulose, hemicellulose, pectin polysaccharides and structural proteins such as glycoproteins. Cellulose is a simple structure composed of β-1,4-linked glucan chains. Hemicelluloses are polysaccarides that hydrogen-bond to cellulose. They are formed by xylans, xyloglucans and mannans with distinct distributions of microfibril types depending on plant species; arabinoxylan is mainly found in monocots while xyloglucan is typical in eudicot. Pectins are formed by homogalactyuronanand typeI and II rhamnogalacturonans (Albersheim et al., 2011).

Once a parasitic plant overcomes the barrier of host cell wall, they are likely face the host defense responses. Bacterial, fungal and oomycetes pathogens inject effector peptides into plant cells to perturb the plant immune system, allowing the establishment of the disease. Resistant host co-evolved Resistant (R) proteins that recognize directly or indirectly the pathogenic peptides of pathogens to block thier growth. A cowpea cultivar resistant against *Striga gesnerioides* expresses a CC-NBB-LRR R protein and knockdown of the R protein turns the cultivar susceptible (Li and Timko, 2009b),raising a hypothesis that the parasitic plants might be able to produce effectors similar to other pathogens.

To understand the interaction of parasitic plant with its host, global gene expression analysis was carried out using interacting and non-interacting parasite and host roots as inputs. To dig the genes having biological relevance out of thousands of sequences, I first selected genes that have statistical support across three biological replicates, and next clustered them according to their expression profiles to rebuilt the putative co-expression networks. In this chapter, I discuss the potential roles of the highly co-expressed genes in the parasitism in both parasites and hosts.

6.3 METHODS

Plant Material

One week-old *P. japonicum* and *O. sativa* (var Nipponbare) seedlings were placed on "rhizotron" chamber and grew separately for further 2 weeks under the conditions described in "Plant material and Growth conditions" section of Chapter 8. Rhizotron chamber was formed by two plastic plates filled with rockwool, nylon mesh and plant root. In one experiment set, the rice roots were removed from a rhizotron with forceps and placed next to *P. japonicum*, and parasite and host roots were collected on 1 or 7 days after host-parasite contact. Additionally, *P. japonicum* and rice roots were separately collected and served as controls. This set named as RIKEN library provided unnecessary stress for rice roots. Thus, on the subsequent experimental sets to avoid disturbance of the root system as much as possible, the plastic plates of rice and *P. japonicum* rhizotronswere removed and both roots were faced each other without removing from nylon meshes. Thus, this infecting rhizotron was formed by two plates filled with rockwool, parasite's nylon mesh, parasite root, rice root and rice's nylon mesh. Similarly, plants were kept in contact for further 1 or 7 days and used as samples. As control RNA from roots of *P. japonicum* or rice growing separately were also harvested. Plant materials with three biological replicates were collected and named in this thesis as NIBB libraries. The RNA extraction method utilized for library preparation is described in chapter 8.

Library preparation

RNA-Seq libraries were prepared as described in "Illumina TruSeq™ RNA Sample Preparation v2 Guide" using half-scale of the original protocol. Briefly, mRNA was isolated from ~10 µg of total RNA using two cycles of Dynabead-based purification as described in Dynabeads® mRNA Purification kit (Life technologies, Cat # 61006). The mRNA was quantified using Agilent RNA 6000 Nano kit with 2100 Bioanalyzer from Agilent Technologies Genomics (Cat # 5067-1511). The mRNA concentration was ranged from 2.6 to 5 ng/µl and ribosomal contamination stayed lower than 9.8% in all samples. Next, the samples were fragmented followed by the synthesis of first and second cDNA strands. The newly synthesized cDNA was passed through a further purification step using magnetic beads (Life Technologies), end- repaired and adaptor-ligated into adenylated 3'end. The table 6.1 shows the adaptor used for each NIBB library. After two cycles of magnetic beads purification the adaptor-ligated cDNA was submitted to PCR amplification step to enrich the cDNA fragments with initial step of 98^{o}C for 30 sec, 15 cycles of 98^{o}C (10 sec), 60^{o}C (30 sec), 72^{o}C (30 sec), followed by extension step of 72^{o}C for 5 minutes.

Quarter scale was used for enrichment step compared with recommended protocol described in "Illumina TruSeq™ RNA Sample Prepration v2 Guide". ABI GeneAmp PCR system 9700 was used for all reactions. After the enrichment step the amplified fragments were twice purified using magnetic beads. Final samples are at a concentration of about 40 ng/µl and having a peak size of 300 bp (measured by Bioanalyzer DNA 1000 chip, Cat # 5067-1504). The precise quantification of library concentration was carried out with cDNA KAPA Library Quantification kit for Illumina Genome Analyzer platform (Cat # KK4824). Equal molar amounts of each sample were mixed in one tube and submitted for multiplex-based Illumina sequencing. Overview of RNA-Seq libraries construction is shown in figure 6.1.

Assembly and Annotation

The sequence reads were assembled with CLC Genomics Workbench (ver 4.8). Different assembly strategies were performed. In Strategy "A" three preliminary assemblies corresponding to reads from "Pj roots", "1d" and "7d" tissues were performed separately using the reads from four experimental sets (RIKEN, 1st NIBB, 2nd NIBB and 3rd NIBB) using 20 bp as parameter for the overlap length between sequences (word). To increase size of the transcripts the previous assembled contigs described in chapter 2 was also included in these assemblies. Rice reads were removed and final assembly was generated using an overlap size of 64 bp and minimum contig size of 300 bp. In the strategy "B" was similar to "A" with the difference that the option automatic word size adjustment was activated. And in the strategy "C", the previous *de novo* assembly from Chapter 2 was included only at the final assembly. Arabidopsis-trained Augustus program (Stanke et al., 2004) was used to determine the quality of contigs. To increase the program performance optimization using previous *de novo* assembly transcriptome of *P. japonicum* described in chapter 2 were performed as following the instruction provided in the website (http://bioinf.uni-greifswald.de/bioinf/wiki/pmwiki.php?n=Augustus.Augustus). Best blast hits for each assembled contig was assigned using blast2go (Conesa et al., 2005). Clustering and differential expression analysis was performed using CLC Genomics Workbench (Grabherr et al., 2011).

Secretome and mining for parasite effectors

Protein sequences of the assembled *P. japonicum* contigs were predicted using the ESTScan2 software. The predicted protein sequences were used as input for SignalP (ver 3.0) and TargetP (v 1.1) softwares for identification of signal peptides targeting to apoplastic region under default parameters. To remove those potential proteins with transmembrane domains the software TMHMM (v.2.0) was applied. Next a homemade perl-based scripts were written to select protein sequences that start methionine and have a sequence size smaller than 300 amino acids.

Comparison of Rice genes induced by *P. japonicum* and *S. hermonthica*

The rice genes induced after the interaction with *P. japonicum* were compared with published rice microarray data which was hybridized with sequences from root tissues infected by the parasitic plant *S. hermonthica* for 2, 4 and 11 days (Swarbrick et al., 2008). Swarbrick et al. (2008) compared two rice cultivars IAC165 and Nipponbare, which are susceptible and resistant to *S. hermonthica*, respectively. I used the data from the susceptible cultivar IAC165 for the comparison analysis to identify the shared parasite-induced genes. Microarray data was downloaded from Gene Expression Omnibus (GEO) under the code of GSE10373 and imported into CLC Genomics Workbench software. The raw values assigned as absent flag and with p-value lower than 0.05 were removed from the analysis. Next the scaling normalization method was applied and expression values of infected root compared with mock-treated samples. The sequences showing a ratio higher than 2 times fold were selected and shown in figure 6.8.

6.4 RESULTS AND DISCUSSION

To identify genes involved in host-parasite interaction, a large-scale transcriptome analysis of *P. japonicum* roots parasitizing rice for 1 or 7 days was carried out. In this experiment, the whole *P. japonicum* and rice roots were collected together, and for control RNAs from parasite (Pj roots) and rice roots (Rice roots) growing independently were included. Thus, four Illumina libraries were sequenced in three biological replicates. The overview of RNA-Seq experiments is shown in Figure 6.1 and the details is described in "Methods" section. In total 60.89 Gb were sequenced (Table 6.1) in collaboration with National Institute for Basic Biology (NIBB); 34.98 Gb, 15.45 Gb and 10.46 Gb for the first, second and third replicates, respectively. To increase the assembly quality another library set was prepared independently and sequenced at RIKEN Omics Center. The strategies for *de novo* assembly of *P. japonicum* transcriptome are summarized in figure 6.2 under different parameters. To assess the quality of assembly, I verified the values for N50 which is the length of 50% of the contigs, average length of the transcripts, number of putative coding sequences and number of complete transcripts (Table 6.2). Strategy "A" showed longer transcripts and higher percentage of sequences with putative coding region compared with those found in "B" and "C" . Therefore, it was chosen for the annotation and expression analyses.

Once *P. japonicum* contigs were *de novo* assembled, they were used as a reference to map the reads of "Pj roots", "1 d" and "7 d" conditions from three biological replicates. In addition, the rice host reads were mapped against the public available Rice transcriptome (www.phytozome.net). Therefore, this experiment allowed the identification of regulated genes during host-parasite interaction from parasite and from host tissues. The projections of expression values were plotted in two-dimensional space spanned by the first and second principal component analysis to evaluate variability among the biological replicates (Fig. 6.3). In the graph groups corresponding to parasite genes expressed in non-interacting roots, 1 day or 7 days of host interaction are indicated by different colors (Fig 6.3-A), and the proximity of dots represents homogeneity of samples. The dots of *P. japonicum* non-interacting root are grouped near to each other, showing good reproducibility among replicates. However, the dots of third replicate set of 1 day and 7 days of interaction are distant from their respective group, indicating that these libraries have higher variability compared to first and second replicates. Expression values of rice genes showed higher similarities among all biological replicates, the exception occurred with second biological replicate of 1 day and 7 days of Rice-

parasite interaction. This data pointed that there is a certain level of variability between the biological replicates although they were sampled under the same growth conditions.

6.4.1 *P. JAPONICUM* RESPONSE TO RICE-INTERACTION

To identify genes in parasitic and host tissues differentially regulated during the parasitism, the ANOVA statistical analysis was performed on the RNAseq data. In total, 2,917 parasite sequences showed to have statistical support evaluated through p-value lower than 0.05. They were divided into nine clusters as showed in figure 6.4. Cluster 1 has 38 transcripts with high expression at 1 dpi (day post infection) which expression continuous high at 7 dpi. Cluster 2, grouped 58 genes with low expression in non-infected roots and at 1 dpi with expression activation at 7 dpi. In the cluster 3 I detected 61 sequences with expression similar in non-infected roots and at 1 dpi, but at 7 dpi their expression was down regulated. The cluster 4 comprised 208 genes with expression profile similar to cluster 2, which means low expression in non-infected roots and at 1 dpi, but with transcript activation at 7 dpi, with the difference that in cluster 4 the expression values ranged between -0.6 to 0.6 while cluster 2 ranged from -1.2 to 1.8. Cluster 5 combined the 549 genes downregulated during infection compare with non-infecting roots. Cluster 6 has 599 genes with peak of expression at 1 dpi. Cluster 7 grouped 225 genes downregulated at 1 dpi with recovered of expression at 7 dpi to a level similar of non-infected roots. Cluster 8 compromises 44 genes up-regulated at 1 dpi. And cluster 9 with 105 genes, whose expression profiles were similar to those found in cluster 7 with downregulation at 1 dpi and expression levels at 7 dpi similar to non-infecting roots. The difference between cluster 7 and 9 relied on the values in y-axis, the gene expression in cluster 2 ranged from -0.6 to proximately 1, and the cluster 9 from -1.9 to 1.8. All the expression values were shown in base-2 logarithm of normalized RPKM (Read Per Kilo base per Million). Out of 2,917, 1,030 parasite genes had low expression values, with a normalized RPKM lower than 1 and were disconsidered in the clustering analysis.

In order to select genes elevant to plant parasitism, the gene regulatory network was reconstructed based on the expression levels. This approach allows identification of highly connected genes (hub genes) which are expected to play a regulatory role in network. Although it remains controversial whether these hub genes have biological importance, this analysis sprovide a clue for further empirical studies (Langfelder et al., 2013). Among the top 10 hub genes from cluster 8, Pj2contig_16920 (*PjCRF-like*) assigned as AP2/ERF transcriptional factor, showed co-expression with other 40 parasite genes. This contig contains single AP2

domain and has as its closest homolog (e-value 9.8e-56) in *Arabidopsis*, the *Cytokinin Response Factor 4* (*AtCRF4*; At4g27950), a gene responsive to cytokinin (Rashotte et al., 2006). Other similar *Arabidopsis* sequence is *Cytokinin Response Factor 2* (*CRF2*; At4g23750) with e-value of 4.6e-44. The CRF transcription factors are characterized by the presence CRF domain which is divided into TEH and CRF-core regions upstream to AP2 domain (Rashotte and Goertzen, 2010). The CRF domain is found in several genes of land plants, but lacks in algae genome (Rashotte and Goertzen, 2010), which indicates its relevance to conquest the land environments. It has been shown that the expression of CRF-domain-containing transcription factors are linked to positive regulator of cytokinin type-B ARR (Arabidopsis Response Regulator), consistent with the hypothesis that *CRF* genes might be direct target of type-B ARR. The amino acid alignment of PjCRF-like with AtCRFs is shown in the figure 6.5. PjCRF-like (378 aa) shares the conserved core (36 aa) of CFR domain and TEH region (19 aa) with the previously described sequences. Alignment includes the CRF-like Pti6 (also known as SlCRF1) and Tsi1 from *Solanum lycopersicum* and *Nicotiana tabacum*, respectively. Interestingly, these genes regulate expression of various genes encoding to Pathogenesis-Related (PR) protein though direct biding to PR promoters (Zhou et al., 1997; Gu, 2002; Park, 2001), indicating that *PjCRF-like* might have function for activation of *PR* expression, a defense-related gene which expression is mediated by Salicylic Acid (SA) (Yalpani, 1991). These data may suggest that parasite defense responses are activated during the host interaction.

Another examples is Pj2_contig_20810 in cluster 2, which was co-expressed with other 53 genes and annotated as Subtilases (SBT) (Table 6.3). The expression profile of Pj2_contig_20810 (Fig. 6.6) showed that this gene is recruited only at 7 days of host-interaction, consistent with my previous finding (this thesis, Chapter 3). At the chapter 3, I described that at least 5 *PjSBTs* (PjSBT2, -4, -7, -8 and -11) showed expression at7 days of compatible host infection, but no expression was detected when the parasite is infection a non-host *L. japonicus*. The sequence of Pj2_contig_20810 still needs to be confirmed by PCR amplification and Sanger sequencing. For instance, it seems that this contig might belong to a *PjSBT* gene which is not described in chapter 3 (Table 3.4). This data pointed that might exist at least six *PjSBT* genes whose transcriptional regulation is linked with parasite invasion into host root tissue.

Another interesting hub is a gene encoding xyloglucan endotransglucosylase (XET) occurs. This gene belongs cluster 4 and connected with 207 edges which is up-regulated at the 7th day in the host-interaction (Table 6.3). XET functions on the rearragment of hemicellulose

through cutting and rejoing the xyloglucan microfibrils (Nishitani and Tominaga, 1992), resulting in loosening cell walls (Fry et al., 1992). Thereby, the up-regulation of XET at invasion stage emphasizes the intense enzymatic restructuring of hemicellulose chains. This molecular evidence is consistent with histological studies which showed that host as well parasite cells whose shapes are continuously modified during the parasite invasion processes (Riopel and Musselman, 1979; Wm. Vance Baird and James L. Riopel, 1984; MUSSELMAN and DICKISON, 1975).

6.4.2 P. JAPONICUM SECRETOME ANALYSIS

To further identify enzymes involved in such cell structure modification I first identified the putative secreted proteins of *P. japonicum*. The putative proteins identified from *P. japonicum* transcriptome were selected based on their signal peptide and absence of transmembrane domains. To exclude partial protein sequences from analysis, a homemade Perl-based script was used to select only those that start with methionine. Applying this bioinformatics approach, 630 putative secreting protein coding sequences were identified. Among them, total 170 proteins encode enzymes that degrade or modify carbohydrate catalogued by CAZy database (http://www.cazy.org/). The glycoside hydrolase family, which is responsible for the hydrolysis of glycosidic bonds of cellulose/hemicelluloses, is the largest CAZy-family found in *P. japonicum* secretome and contains 75 enzymes. Members of this group were highly (Pj2_contig_43910, Pj2_contig_24090, Pj2_contig_22549, Pj2_contig_37633, Pj2_contig_31980, Pj2_contig_1688) or exclusively (Pj2_contig_29220) expressed in the parasitic stage of *P. japonicum*. Other enzymes that degradate pectin (Pj2_contig_33022) and cellulose (Pj2_contig_22408) were also highly recruited by the parasite (Table 6.5). To establish vascular connection with the host, parasitic plants in Orobanchaceae have to conquer space inside the host tissue between its cells without destroying them (Alejandro, 2013), this strategy requires drastic rearragment and loosing of the components of cell wall in parasite and host tissues. Thereby, it was expected the recrutament of large amount of cell wall-modifying enzymes, and my analysis succeeded to identify the potential molecules responsible for such modifications. However, it remains elusive whether these apoplastic enzymes can distinguish the substrate from host or parasite.

Pathogens have evolved an extensive arsenal against plant defense. Among them are the effectors, which are small secreted peptides (SSPs) that s manipulate the plant systems in order to facilitate the pathogen infection. To investigate whether parasitic plants are able to produce effectors I selected from secretome peptides with size smaller than 300 amino acids, as many of previously identified pathogen effectors are within this size (Table 6.4). Comparative analysis with in *S. hermonthica* transcriptome (Yoshida et al, in preparation) I identified 130 common SSPs in *P. japonicum* and *S. hermonthica*. Out of them, 24 showed ANOVA statistical analysis across three biological replicates and 15 returned a hit in the blast database. The expression profiles of potential effector candidates in *S. hermonthica* and *P. japonicum* transcriptome are in figure 6.8 A-B and those exclusively found in *P. japonicum* in figure 6.8 C-D. Although this analysis was limited to bioinformatics approach and further wet

experiments are required to prove the existence of those effectors, I could succefully identify sequences that may encode potential effectors in plants.

6.4.3 RICE RESPONSE TO PARASITE-INTERACTION

RNA-Seq analysis identified 1155 rice genes with p-value less then 0.05 according to ANOVA statistic test calculated across the biological replicates. These genes grouped into 9 clusters according to their expression profiles (Fig. 6.9). Cluster 1 grouped 61 genes with expression levels similar between non-parasitizing rice roots and at 1 days of parasitism, with downregulation at 7 dpi. Cluster 2 has 96 genes downregulated at 1 dpi and showing expression at 7 dpi similar to non-parasitizing rice roots. Cluster 3 counted with 212 down-regulated genes during parasitism. Cluster 4 compromised 237 up-regulated genes at 1 dpi. Genes from cluster 5 showed an expression profile similar to those genes of cluster 4, with the difference that in cluster 5 the 51 genes showed an amplitude of expression higher, ranging from -1.2 to 2.0, while cluster 4 varied between -0.4 to 0.5 in logarithmic scale in base 2 of normalized RPKM values. The 73 rice genes positivelly modulated at 1 and 7 days of *P. japonicum* infection were grouped in cluster 6. The cluster 7 counted with 49 genes with peak of expression at 7 days and transcript levels similar between non-parasitizing root and at 1 dpi. Cluster 8 has 39 genes with high expression kept during the parasite infection. The cluster 9 has 52 down-regulated genes. In both clusters 9 and 3 grouped genes with transcript levels negatively modulated iby parasite infection, with the difference that cluster 3 the down-regulation occurred gradually and in cluster 9 it occurred abruptally at 1 dpi and kept low at 7 dpi. Out of 1155, 285 rice genes showed low variance across biological samples and the normalized RPKM stayed lower than 1, consequently the log-transformation calculated in this analysis resulted in a not sensitive value (McIntyre et al., 2011). So, I decided to disconsider those genes in the clustering analysis.

In order to identify the key genes from each cluster the co-expression network analysis was carried out. Table 6.6 shows highly connected (more than 50) genes from each cluster designated as hub genes. Similar to the expression analysis of parasite genes, the rice sequences encoding cell-wall modifying enzymes are up-regulated during the parasitism. At 1 day post infection, the rice wall-loosening agents, such as expansin and XET from the cluster 4 with 235 edges each are positively expressed. As I mentioned ealier in this thesis, the rearragment of cell wall is required to accommodate the parasite cells inside the host tissue and homologs of those proteins were also found to be up-reguated in parasitic cells. However,

it is not clear the reason that a host may activate the expression of genes encoding proteins that will loosen itself cell wall and facilitate a parasite infection.

Among up-regulated hub genes, two appear to encode members of the Leucine-rich repeat protein kinase family protein (LRR-RLK) (Table 6.6), the LOC_Os02g05910 from the cluster 4 (235 edges) and LOC_Os05g40270 from the cluster 6 (50 edges). The LRR-RLK family consisits of 165 members in rice with distinct key functions in plant developmental process (Hwang et al., 2011). The expression of these two LRR-RLK genes are widely expressed in various tissues, ranging from seedlings to reproductive structures (Rice Genome Anntotaion Project http://rice.plantbiology.msu.edu). To have a closer insight about the potential role of LOC_Os02g05910 and LOC_Os05g40270 in the plant parasitism, I searched for their orthologs with known functions in other plants. The closest orthologs of LOC_Os05g40270 in *Arabidopsis* is At1g06840, a transmembrane leucine-rich repeat receptor-like kinases (LRR-RLKs) involved in plant immune response though the recognition of microbe-associated molecular patterns (Jehle et al., 2013). For LOC_Os05g40270, the closest ortholog with a described biological role is Glyma14g38670 in soybean (www.phytozome.net) which is associated with plant defense response in soybean (Suh et al., 2011). Interestingly, genes encoding NBS-LRR type R proteins were also up-regulated in the post parasite infections. For instance, LOC_Os04g21890 (Cluster 2 and 91 edges) with a peak at 7 dpi (days post infection) encodes an NB-ARC protein, LOC_Os02g47500 (Cluster 4 with 235 edges) at 1dpi Nodulin MtN21 (Kottapalli et al., 2007) and LOC_Os11g37740 (Cluster 6 and 52 edges) at 7 dpi and 1 dpi, Hopz-activated resistance 1 (Lewis et al., 2010). Taking together these data suggest that rice is able to recognize *P. japonicum* as an itruder and activates its defense arsenal, although not effective enough to confer resistance.

To further understand the role of rice host genes during infection of parasitic plants, I compared the 1155 rice genes regulated during the contact with *P. japonicum* with the previously published rice array data set on rice cultivar IAC165 infected with *S. hermonthica* (Swarbrick et al., 2008) (GSE10373). After 2 days of interaction with *S. hermonthica*, there are 3 induced genes which were also regulated in our RNA-Seq data (Fig. 6.10A). Similarly, at 4 dpi I identified 8 genes (Fig. 6.10B) and at 11 dpi 6 rice genes which were also upregulated in *P. japonicum*-rice interaction (Fig. 6.10C). The list of total 17 genes is shown in Table 6.7. For example, the homeobox gene encoding the leucine zipper-type transcription factor Oshox16 (LOC_Os02g49700), was up-regulated in rice tissues upon contact with *S. hermonthica* or *P. japonicum* . The biological role of Oshox16 is unkown, but is appeared to be expressed in

various tissues at different developmental stages (leaves 20 days, inflorescence, anther/pistil, endosperm and embryo) based on the rice expression database available on Rice Genome Anntotaion Project (http://rice.plantbiology.msu.edu). Other members of the Oshox group are involved in controlling leaf shape (Oshox1) (Meijer et al., 1997) or in suppression of gibberillin responses (Oshox4) (Dai et al., 2008). The activation of a homeobox transcription factor in host tissues upon the infection of *S. hermonthica* and *P. japonicum* may indicate intensive modulation of gene expression during parasitism. This statement was also reflected in my RNA-Seq analysis which I identified 1155 rice genes regulated in interaction with *P. japonicum*.

There are 7 out of 17 genes commonly downregulated after the contact of both parasites. Among them, genes encoding serine carboxypeptidase (LOC_Os02g46260) and multicopper oxidases (LOC_Os1g03530) are worth noted. Serine carboxypeptidases are often associated with plant responses against oxidative stress and pathogen infection (Liu et al., 2008), while the multicopper oxidases are responsible for lignin degradation (Hoegger et al., 2006). It is possible that both parasitic plants target these genes to suppress to avoid host counter attack. Thus these genes can be used as marker for functional analyses of potential plant effectors in the future.

6.5 CONCLUSION

This study identified 2917 parasite and 1155 host genes regulated in plant parasitism. The expression analysis concluded that:

(1) Upregulation of genes encoding expansin and cell wall-modifying enzymes, including xyloglucan endotransglucosylases and glycoside hydrolases are consistent with the fact that both parasite and host cells are under intense modification of their shape.
(2) The defense-related genes encoding PR1 and R proteins are up-regulated in the parasite and in host tissues, indicating that both plants are in a defence mode. It is yet unclear how parasites evade the host defense to invade its tissue to establish the parasitism.
(3) The transcription factor encoding genes *PjCRF-like* (Ap2/ERF family) and *Oshox16* (HD-ZIP family) are up-regulated in parasitic and host tissues, respectively, during the parasitism, potentially indicating their involvement in this biological process.

6.6 TABLES AND FIGURES

Table 6.1 – Number of reads sequenced by Illumina Hi-Seq

Experimental set		RIKEN library*	1st NIBB	2nd NIBB	3rd NIBB
Pj roots	Gb	~5,3	~8,4	~4,55	~2,45
	# total reads	55 098 160	91 247 333	47 569 983	25 734 731
	Average (bp)	97,13	92,9	95,63	95,52
Rice roots	Gb	5,8	9,8	4,47	2,49
	# total reads	60 207 027	107 505 964	47 166 098	26 501 595
	Average	96,23	91,41	94,79	94
Pj parasitizing Rice for 1 day	Gb	~5,5	~7,4	~2	~2,8
	# total reads	57 241 804	81 279 943	21 794 860	30 379 279
	Average	96,73	92	94,98	94,69
	rice reads (Gb)	2,4	4	0,9	~1,9
	Pj reads (Gb)	3,1	3,4	1,16	0,9
Pj parasitizing Rice for 7 days	Gb	5,88	9,38	4,426	2,72
	# total reads	61 266 058	102 039 609	46 451 361	28 849 694
	Average	96,07	91,92	95,3	94,36
	rice reads (Gb)	4,5	4,77	1,24	1,85
	Pj reads (Gb)	1,36	4,6	3,181	0,87

* RIKEN library was used to improve the assembly, but it not included in statistical significant analysis since the plant material was collected in different way and do not represent identical biological replicate.

Table 6.2 - Summary of *de novo* assembly of *P. japonicum* transcriptome

	Software	Strategy A	Strategy B	Strategy C
N75	CLC	1 008	902	932
N50	CLC	1 775	1 561	1 597
N25	CLC	3 010	2 526	2 552
Minimum length	CLC	300	300	300
Maximum length	CLC	19 599	16 753	16 753
Average length	CLC	1 323	1 202	1 231
Count	CLC	54 710	71 615	70 822
Putative coding sequence	ESTScan	49 821 (91,06%)	63 417 (88,55%)	62 726 (88,57%)
Number of complete transcripts	Augustus	16 194 (29, 59%)	27 947 (39, 02%)	22 375 (31, 6%)
Number of transcripts with:	Augustus			
Start codon		22 713 (41, 51%)	33 013 (46, 09%)	28 777 (40, 6%)
Stop codon		26 340 (48, 14%)	20 854 (29, 12%)	33 013 (46, 61%)

Table 6.3 – Hub genes in *P. japonicum* tissues divided into clusters.

	ID	hub	ctrl - Pj	1d - Pj	7d - Pj	p-value	Description
CLUSTER1	**Pj2_contig_28888**	**37**	**0.20**	**3.43**	**3.07**	**0.0019**	**pathogenesis-related protein 1a**
	Pj2_contig_4319	37	0.23	1.83	1.60	0.0039	elongation factor 1-alpha
	Pj2_contig_17540	36	0.13	1.17	1.03	0.0001	u1 small nuclear ribonucleoprotein 70k splice variant 1
	Pj2_contig_36179	36	0.27	1.23	0.77	0.0046	adf-like domain-containing protein
	Pj2_contig_35956	36	0.17	1.10	0.93	0.0169	elicitor-responsive gene
CLUSTER2	Pj2_contig_14253	56	0.77	0.87	1.50	0.0132	zeatin o-glucosyltransferase-like
	Pj2_contig_4790	55	0.40	0.40	1.27	0.0003	peroxidase 4-like
	Pj2_contig_15025	55	0.43	0.53	1.13	0.0006	alpha-ketoglutarate-dependent dioxygenase alkb
	Pj2_contig_2342	55	1.53	1.73	3.10	0.0007	lipoic acid synthase
	Pj2_contig_14769	55	0.63	0.67	1.57	0.0015	#N/A
	Pj2_contig_20810	**54**	**0.23**	**0.20**	**1.83**	**0.0004**	**subtilisin-like protease**
	Pj2_contig_14608	54	0.40	0.50	1.47	0.0007	lignan glucosyltransferase
	Pj2_contig_10835	53	0.27	0.50	1.73	0.0003	peroxidase
	Pj2_contig_13448	53	0.60	0.53	1.27	0.0056	wall-associated receptor kinase-like protein
	Pj2_contig_9483	53	0.60	0.53	1.13	0.0121	protein
CLUSTER3	Pj2_contig_3519	49	0.93	0.83	0.57	0.0038	sigma factor sigb regulation protein rsbq
	Pj2_contig_910	48	0.87	0.77	0.53	0.0064	#N/A
	Pj2_contig_17504	47	0.67	0.63	0.30	0.0056	gras family transcription factor
	Pj2_contig_2876	46	0.57	0.60	0.37	0.0173	3-dehydroquinate synthase
	Pj2_contig_13720	46	0.93	1.03	0.47	0.0332	myb transcription factor
	Pj2_contig_38545	45	0.47	0.60	0.30	0.0001	pap-specific mitochondrial-like
	Pj2_contig_6891	43	1.23	1.43	0.87	0.0107	trichome birefringence-like 43 protein
	Pj2_contig_14854	43	0.57	0.47	0.30	0.0129	protein
	Pj2_contig_13173	42	0.57	0.57	0.37	0.0370	ubiquitin carboxyl-terminal hydrolase family protein
	Pj2_contig_5958	42	1.27	1.30	0.90	0.0469	dna-binding protein
CLUSTER4	Pj2_contig_17405	207	0.57	0.57	0.87	0.0010	conserved hypothetical protein [Ricinus communis]
	Pj2_contig_2186	207	1.30	1.30	1.47	0.0012	homeobox-leucine zipper protein athb-15
	Pj2_contig_2158	207	0.80	0.83	1.13	0.0028	serine threonine-protein kinase oxi1-like
	Pj2_contig_9654	207	0.80	0.80	1.07	0.0039	xyloglucan galactosyltransferase
	Pj2_contig_3665	207	1.77	1.67	2.30	0.0059	xyloglucan endotransglucosylase
	Pj2_contig_8279	207	0.90	0.87	1.27	0.0079	heat stress transcription factor a-4a-like
	Pj2_contig_11544	207	1.27	1.27	1.70	0.0079	transcription factor bhlh144
	Pj2_contig_17227	207	0.83	0.83	1.03	0.0080	rop guanine nucleotide exchange factor 1-like
	Pj2_contig_18883	207	1.17	1.13	1.80	0.0170	peroxidase 4
	Pj2_contig_9785	207	1.33	1.20	1.80	0.0171	cytochrome p450
	Pj2_contig_15282	207	0.97	0.93	1.63	0.0189	zinc finger
	Pj2_contig_15231	207	1.33	1.33	1.57	0.0194	#N/A
	Pj2_contig_578	207	1.07	1.07	1.33	0.0261	heat shock like protein
CLUSTER5	Pj2_contig_11416	546	1.30	1.17	1.20	0.0066	protein
	Pj2_contig_8486	546	1.00	0.87	0.90	0.0066	ring u-box protein
	Pj2_contig_9082	546	1.20	1.07	1.10	0.0066	PREDICTED: uncharacterized protein LOC100255640 [Vitis vinifera]
	Pj2_contig_9878	546	1.00	0.87	0.90	0.0066	pyruvate kinase
	Pj2_contig_7092	546	1.40	1.27	1.30	0.0066	k() efflux antiporter chloroplastic-like
	Pj2_contig_7911	545	1.03	0.87	0.90	0.0110	zinc finger protein
	Pj2_contig_2355	545	1.77	1.60	1.63	0.0110	proteinase inhibitor serpin emp24 gp25l p24
	Pj2_contig_3126	544	1.57	1.27	1.37	0.0020	auxin response factor 9
	Pj2_contig_43331	544	1.17	0.97	1.03	0.0144	acid phosphatase vanadium-dependent haloperoxidase-related protein
	Pj2_contig_1432	543	1.60	1.37	1.43	0.0024	pseudouridine-5 -monophosphatase-like
	Pj2_contig_15516	543	0.80	0.67	0.70	0.0066	s-acyltransferase tip1
	Pj2_contig_9277	543	0.80	0.67	0.70	0.0066	hypothetical protein MTR_1g023970 [Medicago truncatula]
	Pj2_contig_2454	543	1.10	0.97	1.00	0.0066	shaggy-related protein kinase theta-like

	Pj2_contig_755	543	1.10	0.97	1.00	0.0066	protein
CLUSTER7	Pj2_contig_382	224	1.07	0.67	0.87	0.0005	myb domain protein 88
	Pj2_contig_24966	224	0.93	0.40	0.60	0.0009	duf246 domain-containing protein at1g04910-like
	Pj2_contig_14476	224	0.97	0.67	0.77	0.0020	duf21 domain-containing protein at1g47330-like
	Pj2_contig_17006	224	0.93	0.63	0.77	0.0021	lipase class 3 family protein
	Pj2_contig_17168	224	0.97	0.67	0.83	0.0021	l-type lectin-domain containing receptor kinase -like
	Pj2_contig_14980	224	0.63	0.37	0.53	0.0037	heparan-alpha-glucosaminide n-acetyltransferase-like
	Pj2_contig_11001	224	1.13	0.70	0.87	0.0039	protein shoot gravitropism 2
	Pj2_contig_44352	224	1.07	0.60	0.80	0.0048	#N/A
	Pj2_contig_10053	224	0.67	0.47	0.60	0.0055	probable receptor-like protein kinase at5g39020-like
	Pj2_contig_38756	224	0.80	0.37	0.53	0.0085	#N/A
	Pj2_contig_5459	224	0.80	0.47	0.60	0.0101	btb poz domain-containing protein
	Pj2_contig_3291	224	0.60	0.37	0.50	0.0147	probable inactive receptor kinase at4g23740-like
	Pj2_contig_13330	224	1.07	0.63	0.87	0.0179	cc-nbs-lrr resistance protein
CLUSTER8	Pj2_contig_19313	42	0.43	1.20	0.90	0.0012	protein
	Pj2_contig_36432	40	0.43	1.60	1.03	0.0020	#N/A
	Pj2_contig_3311	40	0.50	1.73	1.37	0.0036	senescence-associated protein
	Pj2_contig_11922	40	0.47	1.50	0.93	0.0059	glycosyl hydrolase family 98 carbohydrate binding module
	Pj2_contig_16920	40	0.60	1.03	0.90	0.0062	ap2 erf domain-containing transcription factor
	Pj2_contig_16585	40	1.47	2.97	2.60	0.0249	elongation factor 1-alpha
	Pj2_contig_28236	40	1.10	2.40	2.13	0.0286	acyl- -binding protein
	Pj2_contig_8302	40	1.47	2.50	2.23	0.0487	translation initiation factor eif3 subunit
	Pj2_contig_8021	39	0.93	2.40	2.20	0.0013	conserved hypothetical protein [Ricinus communis]
	Pj2_contig_13886	39	0.40	0.80	0.77	0.0025	protein kinase g11a-like
CLUSTER9	Pj2_contig_18255	104	0.93	0.50	1.00	0.0002	predicted protein [Populus trichocarpa]
	Pj2_contig_6331	104	1.83	0.87	1.70	0.0003	o-acyltransferase wsd1
	Pj2_contig_13919	104	0.93	0.47	0.93	0.0006	sulfate bicarbonate oxalate exchanger and transporter sat-1
	Pj2_contig_3864	104	0.90	0.37	0.83	0.0009	pentatricopeptide repeat-containing protein
	Pj2_contig_20720	104	0.90	0.37	0.87	0.0017	carbonic anhydrase
	Pj2_contig_8425	104	1.07	0.47	0.93	0.0047	protein
	Pj2_contig_13554	104	1.40	0.43	1.47	0.0056	potassium transporter 5-like
	Pj2_contig_19683	104	0.40	0.20	0.40	0.0370	pyridine nucleotide-disulfide oxidoreductase domain-containing protein 2-like
	Pj2_contig_19307	103	0.40	0.20	0.47	0.0002	pentatricopeptide repeat-containing protein
	Pj2_contig_16297	103	1.23	0.47	1.17	0.0003	alpha beta fold family protein

* Only gene showing number of connections higher than 50 and the 10 top hub-genes from each cluster are shown

Table 6.4 – Summary of *P. japonicum* putative secretome

	P. japonicum
All contigs	54 710
Predicted protein	49 821
Protein with first Met	22 075
Secreted protein (By SignalP and TargetP)	1 262
No TM (by TMHMM)	630
Less than 300 aa	361
Homolog in Sh	130 (Sh)
Upregulated during penetration	24 (pvalue < 0.05)
Annotated	15 (pvalue < 0.05)

Table 6.5 – Top Summary of Carbohydrate-active enzymes (CAZymes) potentially secreted by *P. japonicum*.

Feature ID	Pfam domains	Log (fold change 1d_vs_Pjcrtl)	Log (fold change 7d_vs_Pjcrtl)
Pj2_contig_33022	Plant invertase/pectin methylesterase inhibitor	∞	∞
Pj2_contig_29811	Carbohydrate phosphorylase	∞	2.5625
Pj2_contig_43910	Glycosyl hydrolase family 28	∞	9.2
Pj2_contig_29220	Glycosyl hydrolase family 38	∞	∞
Pj2_contig_24090	Glycosyl hydrolase family 79	∞	1.304348
Pj2_contig_22549	Glycosyl hydrolase family 16	2.769230769	4.461538
Pj2_contig_37633	Glycosyl hydrolase family 20	2.263157895	∞
Pj2_contig_31980	Glycosyl hydrolase family 3	2.15	1.25
Pj2_contig_20113	alpha/beta hydrolase fold	1.586206897	1.310345
Pj2_contig_22408	Cellulase (glycosyl hydrolase family 5).	1.45	2.075
Pj2_contig_1688	Glycosyl hydrolase family 16	1.265625	∞

Table 6.6 – Hub genes in Rice tissues divided into clusters.

	ID	hub	ctrl - Rice	1d - Rice	7d - Rice	p-value	Description
CLUSTER 1	LOC_Os09g39660.2	54	5.13	5.57	2.73	0.016	C2H2-like zinc finger protein
	LOC_Os06g50910.2	51	1.30	1.23	0.50	0.006	RAD3-related
	LOC_Os09g29820.1	51	13.20	14.90	8.40	0.009	basic leucine-zipper 42
	LOC_Os05g34150.1	50	8.97	8.83	5.10	0.031	glutathione S-transferase
CLUSTER 2	LOC_Os04g43430.3	95	0.47	0.13	0.73	0.0155	Pentatricopeptide repeat (PPR) superfamily protein
	LOC_Os02g43314.1	94	1.47	1.00	1.50	0.0019	-
	LOC_Os01g57020.1	93	0.83	0.37	1.07	0.0408	Domain of unknown function (DUF966)
	LOC_Os04g54220.1	92	4.77	2.03	6.10	0.0303	-
	LOC_Os09g39740.1	92	3.13	0.93	3.00	0.0025	ACT-like superfamily protein
	LOC_Os04g21890.1	91	1.93	1.13	2.27	0.0358	NB-ARC domain-containing disease resistance protein
	LOC_Os07g10840.1	91	13.63	8.33	17.53	0.0064	seed imbibition 2
	LOC_Os11g14230.1	91	3.80	2.77	4.87	0.0067	bacterial hemolysin-related
	LOC_Os01g01720.1	90	11.10	8.10	11.43	0.0248	peroxin 14
	LOC_Os07g41370.1	90	15.10	10.37	13.90	0.0004	AGAMOUS-like 8
CLUSTER 4	LOC_Os02g13530.1	236	169.97	207.53	154.63	0.0399	Ribosomal protein
	LOC_Os02g50960.1	236	34.27	38.70	32.90	0.0239	Auxin efflux carrier family protein
	LOC_Os03g13760.1	236	46.27	54.37	40.17	0.0071	Small nuclear ribonucleoprotein
	LOC_Os03g16400.1	236	26.73	31.73	23.93	0.0145	-
	LOC_Os03g57950.1	236	1.37	2.10	1.27	0.0173	CVP2 like 1
	LOC_Os04g30780.1	236	46.93	57.03	42.87	0.0175	translation initiation factor 3 subunit H1
	LOC_Os07g42950.1	236	290.10	356.77	261.27	0.0233	Ribosomal protein S6e
	LOC_Os07g44740.1	236	117.30	140.83	105.47	0.0191	chaperonin 10
	LOC_Os02g05910.1	235	0.83	1.17	0.70	0.0053	Leucine-rich receptor-like protein kinase family protein
	LOC_Os02g47500.1	235	11.73	13.87	11.03	0.0274	nodulin MtN21
	LOC_Os05g30750.7	235	3.70	6.07	3.17	0.0207	Calcium-dependent lipid-binding
	LOC_Os05g32600.1	235	8.77	10.10	7.90	0.0331	cyclin-dependent kinase D1;3
	LOC_Os05g45480.1	235	6.03	9.37	5.13	0.0500	Extensin family protein
	LOC_Os06g48030.1	235	18.73	26.77	17.30	0.0061	Peroxidase superfamily
	LOC_Os08g06200.1	235	29.70	32.57	28.03	0.0349	SPFH/Band 7/PHB domain-containing membrane-associated protein family
	LOC_Os08g14200.1	235	47.60	68.40	40.17	0.0432	xyloglucan endotransglucosylase
CLUSTER 6	LOC_Os03g55890.1	59	2.93	3.53	3.90	0.0009	Protein of unknown function, DUF547
	LOC_Os09g36110.1	58	1.20	1.57	1.80	0.0005	Subtilase family protein
	LOC_Os03g58540.1	56	0.43	0.60	0.80	0.0016	GTP1/OBG family protein
	LOC_Os05g07900.2	56	2.87	3.60	4.30	0.0023	Mitochondrial substrate carrier family protein
	LOC_Os01g67520.1	55	1.20	2.03	2.80	0.0284	galactose-1-phosphate guanylyltransferase
	LOC_Os01g62070.3	52	0.53	0.93	1.33	0.0156	Cation efflux family protein
	LOC_Os11g37740.1	52	3.70	4.60	5.67	0.0074	HOPZ-ACTIVATED RESISTANCE 1
	LOC_Os05g28830.1	51	1.53	2.00	2.70	0.0084	TRICHOME BIREFRINGENCE-LIKE 33
	LOC_Os04g39900.1	50	10.10	14.70	17.13	0.0236	beta glucosidase 15
	LOC_Os05g40270.10	50	1.47	2.07	2.67	0.0220	Leucine-rich repeat protein kinase family protein

Table 6.7 – Rice genes induced upon interaction with *P. japonicum* and *S. hermonthica*.

Gene	Fold Change	Rice plants challenged with *P. japonicum*				Description
	Rice* challenged with *S. hermonthica* for 2 days	Rice-ctrl	Rice1d	Rice7d	P-value	
LOC_Os04g38950	-2.53	0.03	-	0.30	0.0422	class I glutamine amidotransferase
LOC_Os01g58550	2.21	0.63	0.20	0.13	0.0412	methyladenine glycosylase
LOC_Os03g09970	2.74	11.03	8.07	4.80	0.0214	sulfate transporter
	Rice* challenged with *S. hermonthica* for 4 days					
LOC_Os03g20380	-3.44	0.27	0.23	0.70	0.0485	calcium/calmodulin depedent protein kinasese
LOC_Os02g46260	-3.34	0.80	1.57	0.53	0.0379	Putative Serine Carboxypeptidase
LOC_Os12g31860	-3.00	3.20	7.90	3.90	0.0036	ureide permease
LOC_Os03g05060	-2.51	0.60	0.33	1.47	0.0142	exostosin family domain containing protein
LOC_Os01g58550	-2.21	0.63	0.20	0.13	0.0412	methyladenine glycosylase
LOC_Os01g03530	-2.03	1.47	2.93	0.93	0.0242	multicopper oxidase domain containing protein
LOC_Os02g49700	2.00	1.10	3.20	2.77	0.0428	homeobox associated leucine zipper
LOC_Os04g38950	3.60	0.03	-	0.30	0.0422	class I glutamine amidotransferase
	Rice* challenged with *S. hermonthica* for 11 days					
LOC_Os03g20380	-3.20	0.27	0.23	0.70	0.0485	calcium/calmodulin depedent protein kinasese
LOC_Os02g46260	-3.09	0.80	1.57	0.53	0.0379	Putative Serine Carboxypeptidase
LOC_Os01g03530	-2.10	1.47	2.93	0.93	0.0242	multicopper oxidase domain containing protein
LOC_Os06g51520	-2.05	0.50	1.83	1.30	0.0143	lysine ketoglutarate reductase trans-splicing related 1
LOC_Os01g65550	-2.03	11.60	12.70	16.07	0.0137	RNA recognition motif containing protein
LOC_Os04g38950	5.84	0.03	-	0.30	0.0422	class I glutamine amidotransferase

* Rice cultivar IAC165. Reference in Swarbrick et al., 2008

*

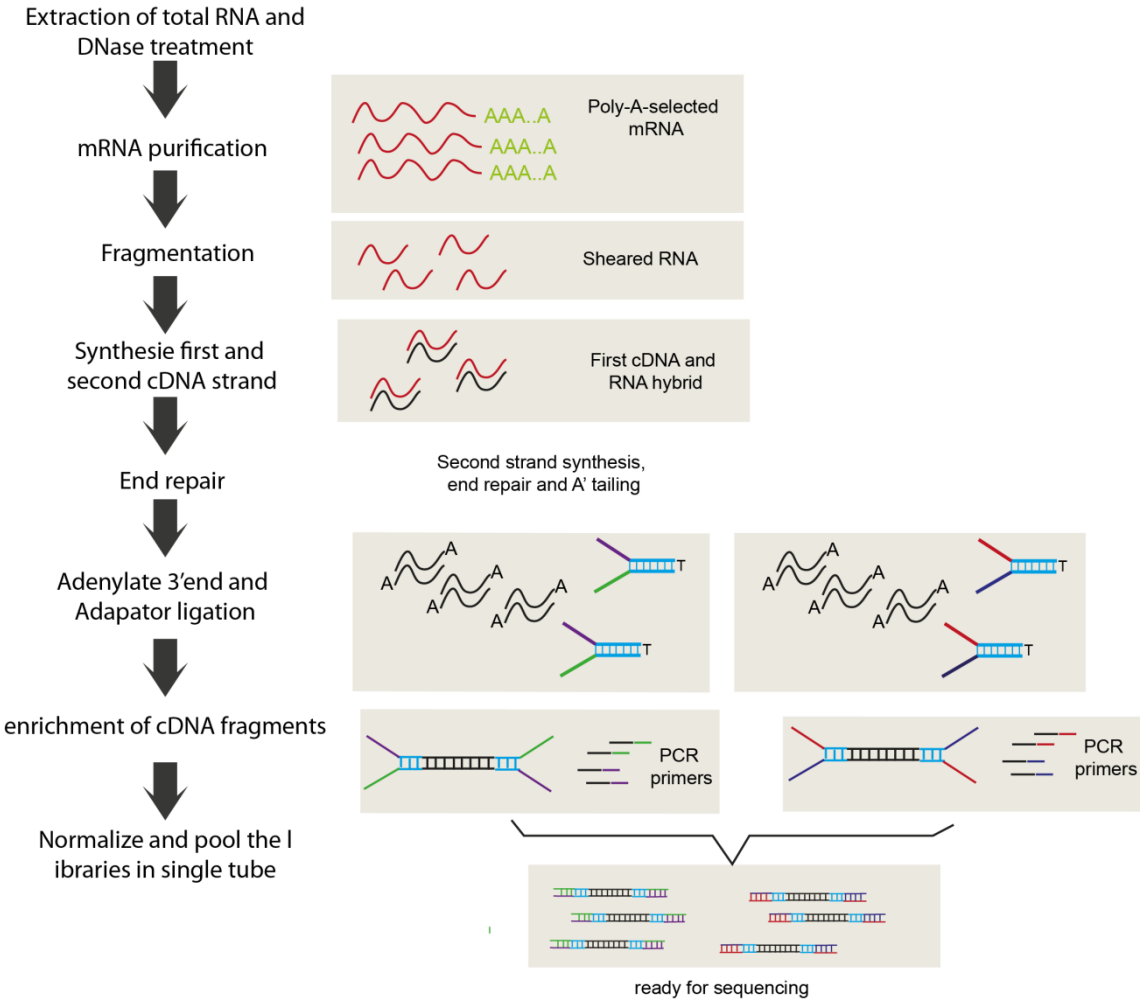


Figure 6.1 – Overview of RNA-Seq library construction. mRNA was purified from total RNA fragmented to 300 bp size range. Next the sheared RNA is reverse-transcribed to cDNA. The double-stranded cDNA was end-repaired and adenylated. Illumina adaptor were ligated, in the figure the construction of two libraries with two different adaptors are shown. Finally, the ligated DNA is amplified and normalized amount of each library was mixed in a single tube and the samples are ready for sequencing. Modified from Wang et al., 2011.

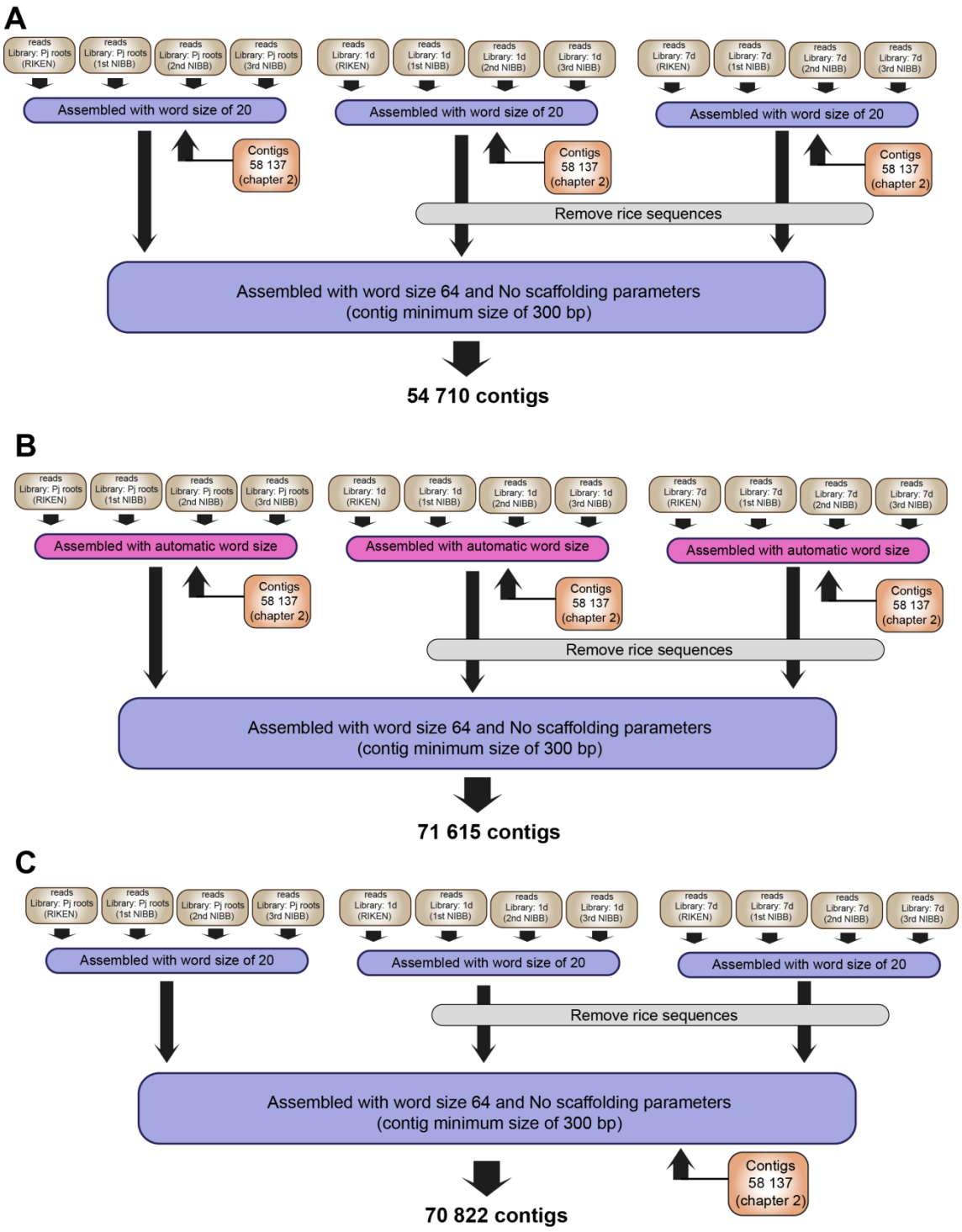


Figure 6.2 –Different strategies to de novo assemble P. japonicum transcriptome. In light brown the reads from different libraries sequenced by Illumina Hi-Seq methodology. In blue the parameters used in CLC genomics Workbench (ver. 4.8) to assemble the sequences. To improve this assemble the contigs described at chapter 2 was also included (showed in orange).

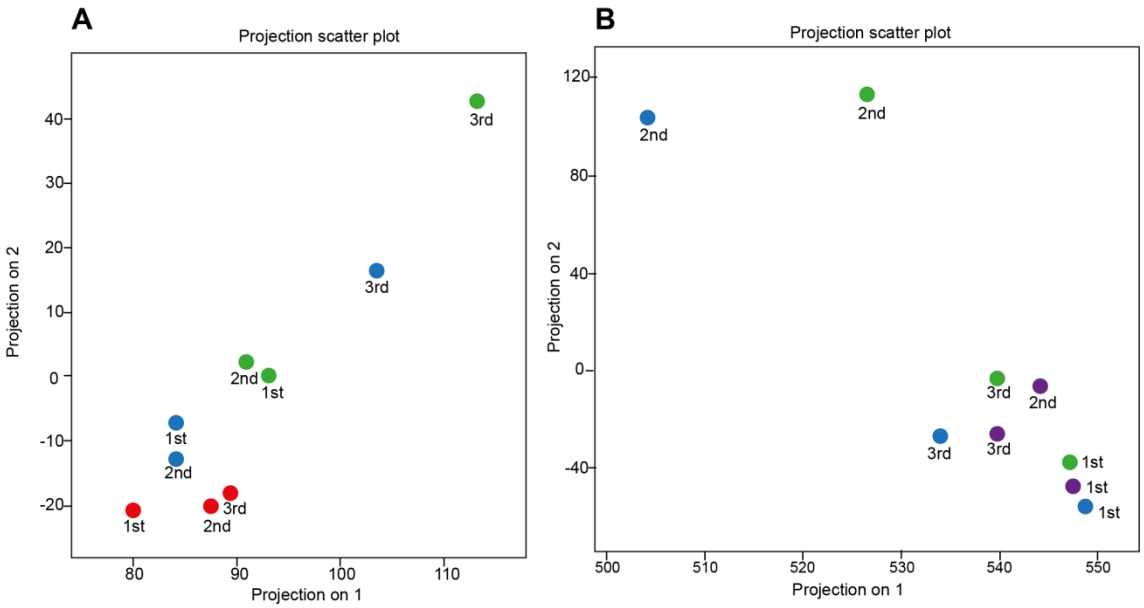


Figure 6.3 Multidimensional scaling plot and differential expression analysis. Library corresponding to *P. japonicum* root withou host (●), Rice roots (●), 1 day of rice-parasite interaction (●) and 7 days of rice-parasite interaction (●). **A** *P. japonicum* genes and **B** Rice sequences The first, second and third biological replicates are indicated in the figure.

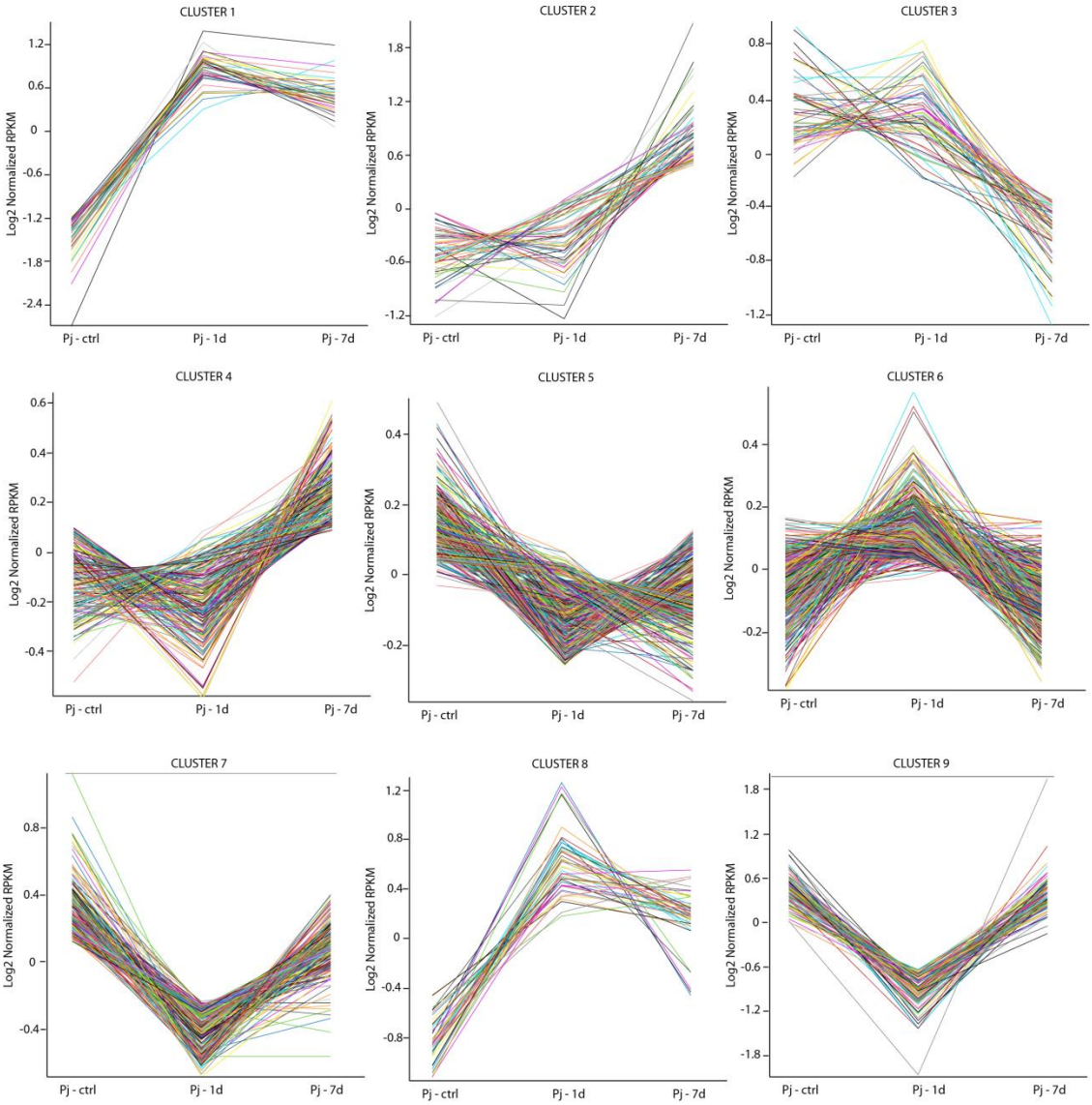


Figure 6.4 –Clusters of *P. japonicum* sequences regulated at rice-host-interaction. Differentially regulated rice sequences grouped according to their expression profile after 1 day (Pj-1d) and 7 days (Pj-7d) of interaction with the host Rice (*O. sativa* var. Nipponbare). Pj-ctrl means *P. japonicum* roots growing without host. The expression values are shown as $\log_2$ of normalized RPKM (Reads Per Kilo base per Million) value.

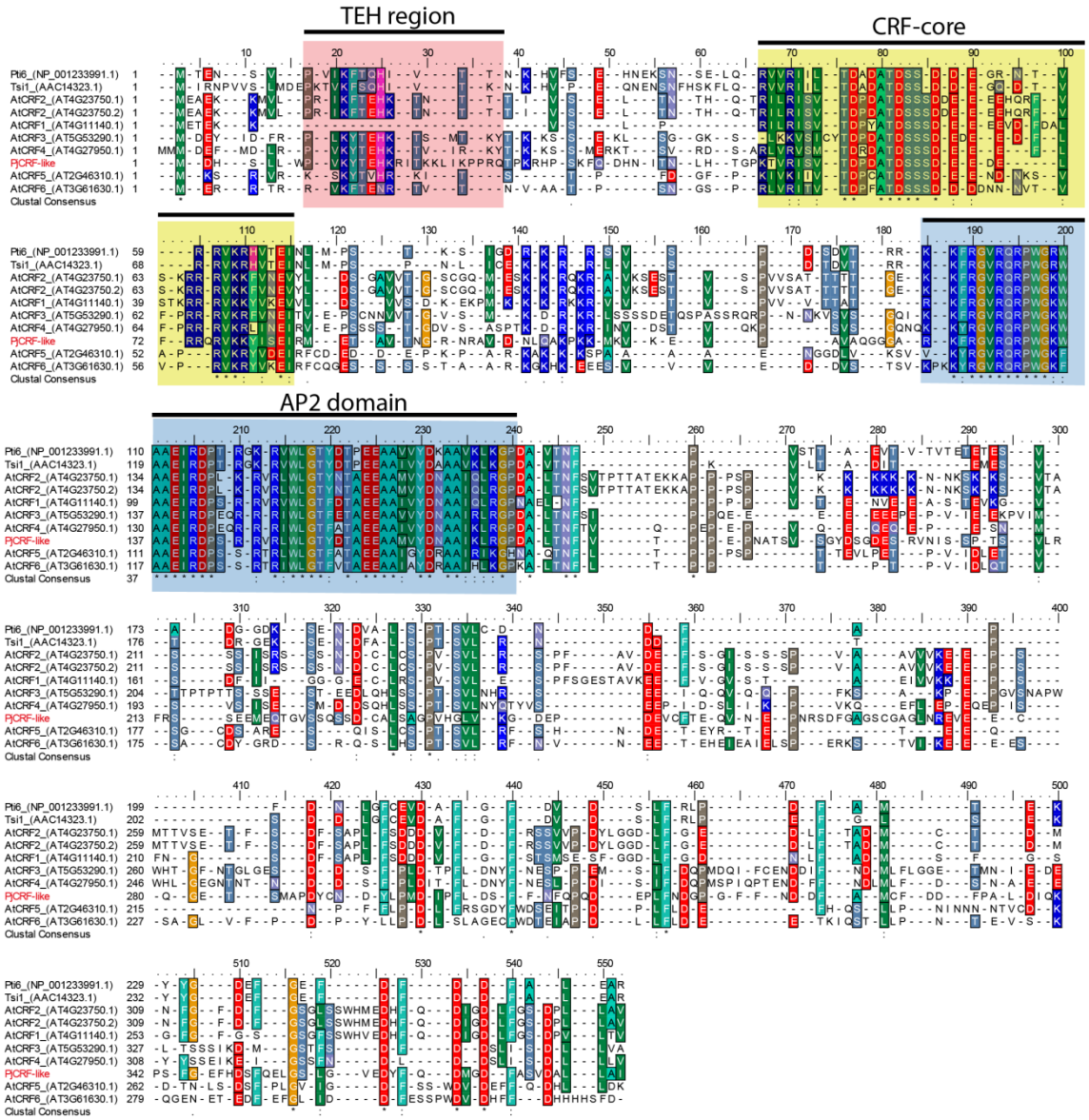


Figure 6.5 - Amino-acid alignment of Pjcontig16920 (PjCRF-like) and CRF-like AP2/ERF transcription factors. The sequences were aligned using CLC Genomics Workbench (v 4.8) and the figure generated by Bio Edit. The alignment includes PjCRF-like (in red), *Arabidopsis* CRF Transcription factors, Pti6 (*Solanum lycopersicum*) and Tsi1 (*Nicotiana tabacum*). In light red the TEH region, Yellow box corresponds to CRF-core domain, blue box the AP2 domain.

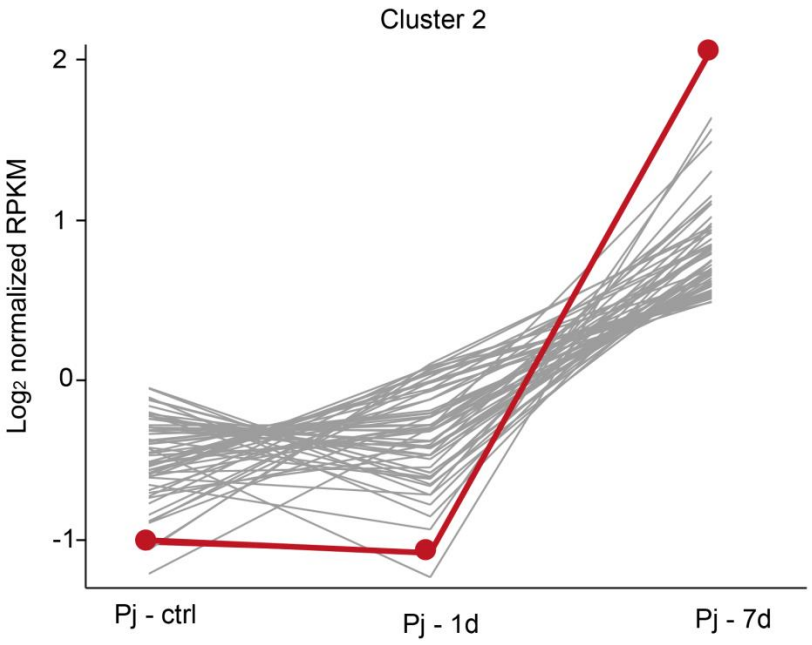


Figure 6.6 – Host-induced Subtilase gene. The contig Pj2_contig20810 a subtilase gene was grouped into cluster 2 (red line). The peak of expression occurs at 7 days after the contact with rice host. The y-axis shows the $\log_2$ of normalized expression values measured as RPKM (Reads per Kilo base per Million).

Figure 6.7 - Amino-acid alignment of Peptidase S8 domain of Pj2_contig20810 and other PjSBT sequences. The sequences were aligned using CLC Genomics Workbench (v 4.8) and the figure generated by BioEdit. The alignment is restricted to peptidase S8 domain. * Means that this sequence is from assembly and it still need to be confirmed by PCR amplification.

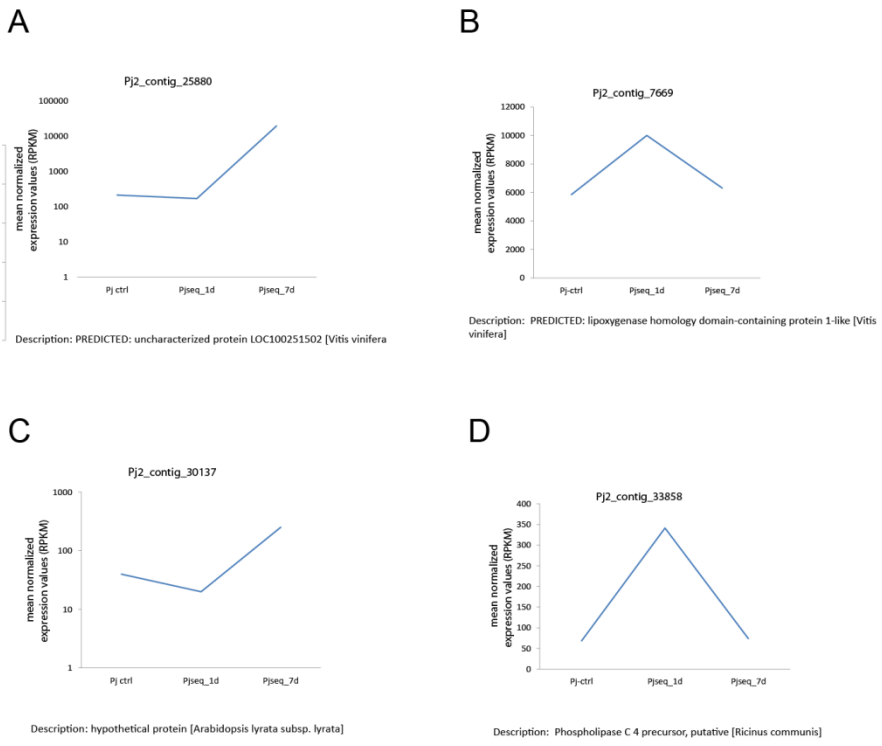


Figure 6.8 – Putative candidates effectors of the *P. japonicum*. The putative small secreted peptide of *P. japonicum* transcriptome predicted using bioinformatics approach. **A-B** expression profiles of predicted SSPs shared by *Striga hermonthica* and *P. japonicum*. **C-D** expression profiles of potential candidates to be effector exclusively found in *P. japonicum* sequences.

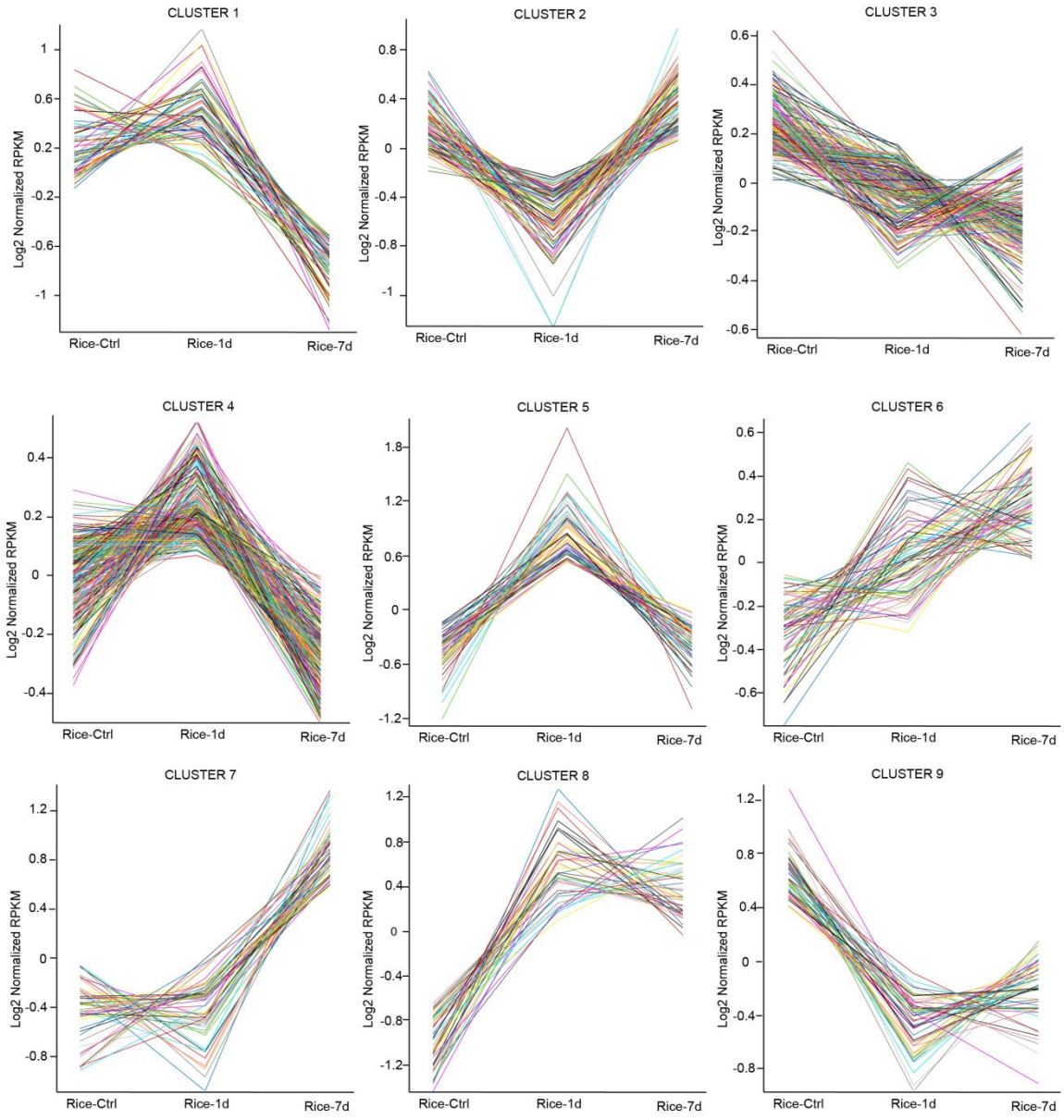


Figure 6.9 –Clusters of rice sequences regulated at *P. japonicum*-interaction. Differentially regulated rice sequences grouped according to their expression profile after 1 day (Rice-1d) and 7 days (Rice-7d) of interaction with the parasitic plant P. japonicum. Rice-ctrl means rice roots growing without contact with the parasite. The expression values are shown as $\log_2$ of normalized RPKM (Reads per Kilo base per Million) value.

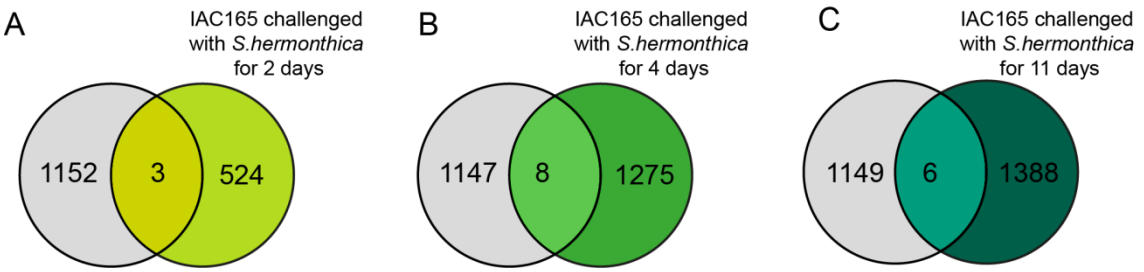


Figure 6.10 – Comparison between rice genes regulated upon the challenge with parasitic plant *P. japonicum* and/or *Striga hermonthica*. Venn diagram showed the number of rice genes induced genes after the interaction with *P. japonicum* (in grey and the reference is this study) and *S. hermonthica* (in different tons of green. Reference in Swarbrick et al., 2008). The *S.hermonthica*-induced rice genes are described in table 6.7.

7 MATERIAL AND METHODS

7.1 PLANT MATERIAL AND GROWTH CONDITIONS

P. japonicum (Thumb.) kanitz seeds were surface-sterilized and germinated as previously described in Ishida *et al.*, 2011. The outer coat of rice seeds were mechanically removed (*Oryza sativa* cultivar Nipponbare). The seed-coat-removed seeds were immersed in 70% ethanol for 3 minutes with agitation, washed with 50% (v/v) commercial bleach solution (including xx% of sodium hypochlorite, Kao, Japan,) for 30 minutes, and extensively rinsed with de-ionized water. Fifty rice seeds were sown in a petri dish with a moisten filter paper (Advantec No. 2) with 10-15 ml sterile water. Seeds of *Lotus japonicum* Miyakojima MG20 were scarified for 10 seconds and seed-surface sterilized through vigorous agitation in bleach solution 10% (v/v) for 10 minutes, followed by at least 5 times washing with de-ionized water. The seeds were kept overnight in darkness immersed in water at 4^{o}C. Next they were sown in petri dishes with moisten paper. Plants were kept growing at 25^{o}C under photoperiod 16 h light/ 8 h dark. For the expression analysis in adult tissues, 1-week-old seedlings of P. japonicum were transferred to soil and kept in long day condition (16h light and 8h dark) at 25^{o}C for 2-3 weeks. Then plants were transferred to short day (8h light and 16h dark) at same temperature for 4 months. A pool of green fruits, flowers at different developmental stages, roots, stem sand leaves were collected and RNA extracted as described below.

7.2 *DE NOVO* TRANSCRIPTOME SEQUENCING

About 10-12 seedlings of rice and 6-10 seedlings of *P. japonicum* with 1 week old were carefully placed in “rhizotron” system, which is described with details bellow and also in Yoshida and Shirasu, 2009. Parasite plants were growing together with rice for 4 to 6 weeks in rhizotron under growth conditions described above. Haustoria were identified in the rice-parasitizing *P. japonicum* root assisted by stereoscope. They were carefully excised and immediately frozen in liquid nitrogen. To enrich the sample with haustorial tissue (parasitic stage library) the root segment which did not develop into haustoria and rice roots were avoided as maximum as possible. When haustoria were not possible to be disconnected from rice root, the both tissues (from parasite and host) were collected. RNA from non-parasitizing *P. japonicum* roots were also collected as a control (autotrophic stage library). RNA were extracted using CTAB/LiCL based-protocol as previously reported (Yoshida, Ishida, et al., 2010).

Genomic DNA was removed from RNA samples by treatment with DNase (QIAgen. Cat No. 79254), followed by a cleaning step with RNeasy Mini Spin Column from RNeasy Plant Mini Kit (QIAGEN, Cat No. 74904). RNA was quantified with Nanodrop ND-1000 spectro-photometer (NanoDrop Technologies). RNA quality was assessed with Agilent 2100 bioanalyzer (Agilent Technologies, USA) using RNA 600 Nano kit (Cat.No.5067-1511). RNA quality for sequencing was set at minimum of RIN 8.0 (RNA Integrity Number)

For sequencing RNA-seq libraries 20 ng of total RNA (concentration above 400 ng/µl) from haustorial tissue (parasitic stage library) and the non-parasitizing root tissues (autotrophic stage library) were sequenced using Illumina-HiSeq2000 platform. The third library composed by mRNA normalized by mixing tissues was sequenced using Roche 454 platform. Around 4 µg of mRNA (concentration above 550 ng/µl) was isolated from parasitic and non-parasitizing root using the Illustra^{TM} mRNA purification kit (GE healthcare. Cat. No. 27-9258-01). Library construct and sequencing were performed by the Genome core facility at RIKEN. Sequence reads will be deposited in the NCBI sequence read archive after publishing the manuscript.

For assembly, the adapter oligonucleotide sequences were trimmed off from raw Roche 454 reads. Next the trimmed reads and the paired-end Illumina-HiSeq reads from parasitic stage library were mapped against the annotated coding DNA sequences of *O. sativa* using Mosaik program (http://bioinformatics.bc.edu/marthlab/Mosaik), the aligned reads were filtered out and the unaligned reads were assembled using combination of CAP3 (Huang and Madan, 1999) and CLC Genome Workbench Software 4.9. BLAST annotation and Gene Ontology analysis of translated sequences were done by blast2GO. The assembled contigs were then queried against 10 proteome in PlantTribes 2.0 (Wall et al., 2008) and assigned to PlantTribes 2.0 orthogroups, tribes and supertribes based on the blast results. The expression values were computed through mapping of reads against the assembled contigs using CLC Genome Workbench 4 software using the RPKM (reads per kilobase of transcript per million) estimation. The identification of differentially expressed sequences was carried out calculating for each transcript the ration between $\log_2^{(1+RPKM)}$ from parasitic stage and $\log_2^{(1+RPKM)}$ of autotrophic stage library.

The ultraconserved orthologs (UCOs) genes were downloaded from http://compgenomics.ucdavis.edu/compositae_reference.php website and their representation in assembled transcriptome was estimated by aligning the 357 UCOs protein

sequence against to *P. japonicum* transcriptome using blastx with threshold of e-value < $1e^{-10}$ using the local BLAST tool available in the CLC Genome Workbench 4. The calculation of full length was made based on the quantity of contigs that match to UCO showing the start and stop codon. We also considered those sequences that show almost perfect alignment, missing only one amino acid, the first Met or the stop codon.

The orthologs of single-copy nuclear genes shared in *Arabidopsis thaliana*, *Populus trichocarpa*, *Vitis vinifera* and *Oryza sativa* called APVO genes were identified using the PlantTribes 2.0 program which identified the Tribes (close to gene family concept) and Orthogroup (close to orthologs concept) based on cluster analyses of the inferred proteomes of ten sequenced angiosperms (*Arabidospsis thaliana*, *Carica papaya, Populus trichocarpa, Medicago truncatula, Oryza sativa, Sorghum bicolor,* the fern *Selaginella moellendorfii,* the moss *Physcomitrella patens* and the green alga *Chlamydomonas reinhardtii*) as described in Wickett et al (REF). The unigenes classified as belonging to the same tribe as the single-copy gene in these four plant genomes were assumed to be orthologs of APVO genes.

7.3 CUSTOM MICROARRAY DESIGN, LABELING AND HYBRIDIZATION.

Oligonucleotide probes (60 mers) fixed on Agilent 8x60K microarrays slide were generated using eArray tool (http://www.chem.agilent.com - design number 253261110001/253416810002) based on *de novo* assembled *P. japonicum* transcriptome. For custom microarray slide, putative coding regions of assembled contigs were detected by ESTScan2 (Iseli et al., 1999) and contigs without putative coding regions were filtered out . Homemade perl script was written to exclude the sequences encoding putative open reading frame with size smaller than 50 amino acids. Other processing tasks and sequences analyses were performed using the scripts available as part of the Biopieces (www.biopieces.org). For control of the background and uneven hybridization, ten probes for randomly selected fifty unrelated sequences were designed and fixed in different locations of the array. Next the generated *P. japonicum* probes were plotted in duplicate to complete 8x60K array. The raw microarray data, which are MIAME (minimum information about a microarray experiment) compliant, will bedeposited at the ArrayExpress microarray data archive (http://www.ebi.ac.uk/microarray-as/ae/) at the European Bioinformatics Institute. For microarray hybridization, total RNA was extracted from *P. japonicum* roots treated with DMBQ or DMSO (mock), and hybridization, washing and scanning were performed according to Agilent instructions. Shortly, 25 ng of total RNA from each sample was used to synthesize

fluorescent-labelled cRNA using Cyanine 3-CTP according to the manufacture's protocol (Agilent One Color RNA Spike-In Kit, Product Number 5188-5282). The labeled RNA solution and microarray slides were set in gasket slides loaded in Agilent microarray hybridization chambers and placed in hybridization incubator for 16 h at 65^{o}C. The hybridized slides were carefully transferred into the washing solution buffer 1 and all the subsequent washing steps were performed following the instructions in Agilent technologies protocol. Microarray slides were scanned on an Agilent DNA microarray scanner. Raw data was extracted using the Feature Extraction Software (v 10.7.3); spots without signal or under the background were removed from analysis. Feature extraction files were imported to GeneSpring GX software (v. 11.0) using baseline transformation to median of all samples and normalization option set at 75% of median shift. The differentially expressed genes were statistically identified though Student's T-Test on normalized ratio data with the significance threshold of $p < 0.05$. DMBQ-specific differently expressed genes were clustered using CLC genomics workbench 4 for Heat Map hierarchical algorithm and GenePattern version 2.0 (Reich et al., 2006) for Self-organizing method.

7.4 RNA EXTRACTION AND QUANTITATIVE RT-PCR ANALYSIS

RNA were extracted using CTAB/LiCL based-protocol as previously reported (Yoshida, Ishida, et al., 2010). Genomic contamination was removed using DNase I Amplification Grade (Cat.No.18068-015) and the cDNA first strand was synthesized with kit ReverTra Ace-α- (Toyobo, Code FSK-101), following the manufacture instructions. For qPCR analysis the kit Thunderbird™ SYBR® qPCR Mix (Toyobo, Code A4251K) was used. The primers used in this study are displayed in Appendix. All the experiments were average of at least two biological replicates with two-three technical replications each. For identification of constitutive genes in *P. japonicum* the whole transcriptome was aligned against the well described housekeeping gene of *Arabidopsis thaliana* (Czechowski et al., 2005) using blastx algorithm. The primer pair designed for Polypyrimidinetract- binding protein (PTB), a RNA-binding protein, showed high amplification efficiency and specificity for *P. japonicum*

7.5 FINDING *PJYUC* GENES IN *P. JAPONICUM* TRANSCRIPTOME

To identify homologs of *Arabidospis YUCCA* genes in *P. japonicum,* the amino acid sequence of Arabidopsis *YUCCA* genes (AT4G32540, AT1G04180, AT1G04610, AT1G48910, AT2G33230, AT4G13260, AT4G28720, AT5G25620, AT5G43890) were locally aligned against *P. japonicum* transcriptome database using tblastn algorithm. Sequences showing alignment score below an e-value of 1E-35 and amino acid identity over 35% were considered as homologs to *Arabidopsis YUCCA* genes. Rapid amplification of cDNA ends (RACE) method was used to clone cDNA containing the 5'-end and 3'-end of *PjYUC* homologs (GeneRacer™ Core kit, Invitrogen – Cat.No. 045-0168). The RACE library was constructed based on a pool of RNA extracted from *P. japonicum* roots in the parasitic and autotrophic stage. To clone specifically the 5'end and 3'end region of each gene two rounds of a touchdown PCR reaction () was performed using two gene specific primers with TM=72-74°C. The RACE product was subcloned and sequenced. To clone a full length cDNA primers corresponding to the sequences at both ends of the cDNA were designed. PCR amplifications were performed using as template first strand cDNA library from parasite root. All PCR amplification reaction in this manuscript were performed using Advantage® 2 PCR Enzyme (Clontech – Cat.No. 639137) and the PCR fragments cloned with TA Cloning® kit (Invitrogen – Cat.No.K2000-01).

7.6 TRANSFORMATION OF P. JAPONICUM

Transformation of *P. japonicum* was performed as previously described in Ishida *et al.*, (2011) with minor modifications. Three day-old *P. japonicum* seedlings were immersed in bacterial solution (OD_{600} 1.0) and submitted by ultra-sonication using a bath sonicator (Ultrasonic automatic washer, AS ONE, Japan) for 10-20 s. Next a continuous vacuum was applied in sonicated-treated seedlings during 5 minutes. Seedlings were transferred to co-cultivation media (Gamborg B5 agar media, 1% sucrose, 450 µM Acetosyringone) and kept in the dark at 22°C for 2 days. After the co-cultivation period the seedlings were transferred to B5 agar media containing cefotaxime (300µg/ml). After 3-4 weeks the transformed root were analyzed. To identify the transgenic root all plasmids used in this manuscript carry in the T-DNA region a constitutive promoter fused upstream of a gene encoding fluorescent protein. The fluorescence was detected through LEICA M165 FC stereoscope.

7.7 VECTOR CONSTRUCT FOR PROMOTER ANALYSIS

The 35S promoter and *red fluorescent protein* (*mRFP*) sequences were PCR amplified from pEZS-CL vector (Yoshimoto et al., 2004). Fragment was inserted at HindIII site into linearized R4L1pGWB532 vector (Nakamura et al., 2009) using In-fusion cloning kit (Clontech - Cat No. 638910). The modified vector, called RFP-R4L1, allows the visual screening of transgenic roots by detecting the fluorescence of mRFP protein.

The coding region of *PjYUC3* was compared in nucleotide level with roughly sequenced and assembled *P. japonicum* genome by BLAST algorithm. The putative genomic region showing nucleotide similarities over 95% with coding region of *PjYUC3*was identified. Oligonucleotides franking the putative promoter region and the second exon of PjYUC3 were designed and used for PCR amplification using P. japonicum genome as a template. The 4,924 bp genomic fragment was cloned and the sequence was confirmed by primer walking sequencing method with ABI3770 sequencer. Then, the promoter region (3,137 bp) was identified and amplified by primers containing attB4 and attB1R recombination arms (Supplemental Table S1). The fragment with recombination sites was cloned into pDONRG-P4P1R vector (Oshima et al., 2011). Finally the promoter of *PjYUC3* was transferred to RFP-R4L1 vector through reaction catalyzed by Gateway LR clonase® (Cat.No. 11791-019). The final construct was double-checked by restriction enzyme digestion and sequencing. The promoter activity was verified through Leica TCS-SP5 II confocal microscopy and beta-glucuronidase assay (Vitha et al., 1995).

7.8 OVEREXPRESSION AND SILENCING CONSTRUCTS

For overexpression analysis, the plasmid pUB-GW-GFP (Maekawa et al., 2008) was provided from National BioResource Project of *Lotus japonicus* (http://www.legumebase.brc.miyazaki-u.ac.jp). Fragment amplified by the primers designed at the beginning of the open read frame and at the end of transcript of *PjYUC3* (PjYUC-pUB-GW-F, PjYUC-full-R) was cloned into pUB-GW-GFP using Gateway technology (Invitrogen). Three day-old *P. japonicum* seedlings were genetically transformed and the transgenic roots selected by fluorescent emission of green fluorescent protein (GFP) marker, whose expression is controlled by 35S promoter.

Silencing vector pHG8YFP modified from pHellsgate 8 harboring the gene encoded to *yellow fluorescent protein* (*yfp*) (Bandaranayake et al., 2010) was used to silence the *PjYUC3* genes. The PjYUC3 gene was locally blasted against transcriptome. Those sequences in the database which showed nucleotide identity over 90% with targeted gene were aligned and the

conserved region was selected for silencing. The targeting region (339 bp) was amplified using the primes described in Supplemental Table S1 and inserted into pENTR^{TM}/D-TOPO vector using TOPO cloning kit (Invitrogen – Cat.No. K240020SP). The resultant clone was checked by sequencing and the targeted sequence was inserted into pHG8YFP plasmid through recombination reaction catalyzed by LR clonase II. The fragment orientation was double checked by digestion with XhoI and/or XbaI and by sequencing using primers-pairs specific for pHG8YFP plasmid (35S-pHGT-F and OscTer-pHGT-R). For control, *P. japonicum* plants were transformed with pHG8YFP empty vector.

The silencing and control constructs were transformed to P. japonicum hairy roots as described above and the hairy roots emerged after 3-4 weeks were used for haustorium-inducing assays. Haustorium development in transgenic roots was monitored after 48 h host root exudate treatment. The number of fluorescent roots that formed at least one haustorium was counted and compared with total number of fluorescent roots in contact with host root. For statistical analysis, at least 50 independent fluorescent roots were analyzed across 3 to 6 biological replicates

7.9 CROSS-SECTIONING, STAINING AND TIME-LAPSE PHOTOGRAPHY

3 day-old and 7 day-old haustorial tissues surrounding the rice and *L. japonicus* roots infected with *P. japonicum* were Safrarin-stained as described previously in (Yoshida and Shirasu, 2009). For cross-sectioning we followed the instructions in the manual of Technovit 7100 (Heraeus Kulzer, Hanau, Germany), more details about the procedures can be find in (Yoshida and Shirasu, 2009). *P. japonicum* was transformed with CYCB1;2pro::YFPnuc construct, the transformation protocol and the detail of the construct are previous described in Ishida et al 2011. Transgenic *P. japonicum* and *Arabidospsis* roots were placed touching each other in a petri dish with agar (1% (w/v)). The petri dish was sealed with vinyl tape and the photographs were taken with Leica MZ16FA stereoscope within an internal of 30 min for 72 h.

7.10 ALIGNMENT AND PHYLOGENETIC TREE ANALYSIS

Transcriptome is composed by mostly by incomplete sequences. To identify full length transcript four oligonucleotides were designed for each sequence. The primers were projected to be highly specific with melting temperature over 70^{o}C annealing and size bigger than 27 mers. Library was constructed using a normalized pool between RNA from autotrophic and parasitic stage using the GeneRacer kit (Invitrogen, Code L1500-01). To increase the specificity

two rounds of touch-down PCR amplification was carried out using decreasing TM values as described in the manual of GeneRacer® kit. All the amplification was performed using the high fidelity Advantage 2 Proofreading Polymerase (Clontech, code 639137). The single amplified bands were cloned into p-GEM®-T Easy vector (Promegra, code A1360) or using TA Cloning® kit (Life technologies, code K2000-01), the reagents were used proportionally as described by the manufacturers to have a minimum volume of 2 µl. The cloned fragments were sequenced by Sanger, and additional primer pair flanking the coding region was designed to confirm the full length sequence. The putative amino acid sequence were aligned using the MUSCLE algorithm available as option tool in the CLC Genomics Workbench and the UPGMA-tree designed with MEGA 5 software (Tamura et al., 2011).

7.11 TRANSCRIPT PROFILES OF S. HERMONTHICA GENES

The raw Illumina GAIIx sequence reads from different developmental stages of *S. hermonthica* were downloaded from Parasitic Plant Genome Project website (http://ppgp.huck.psu.edu/download.php). The reads from nine libraries (StHe0, StHe1G, StHe2G, StHe3G, StHe4G, StHe51G, StHe52G, StHe61G and StHe62G) were mapped against the full length transcript sequence of *ShQR1* and *ShQR2* using CLC Genomics Workbench 4. The parameters adopted were “minimum length fraction” = 1, “minimum similarity fraction” = 1, “maximum number of hits for a read” = 10, for paired settings we used “minimum distance” = 100 and “maximum distance” = 250.

7.12 GENETIC STUDIES OF *P. JAPONICUM* POPULATIONS ACROSS JAPAN

The assembled transcriptome was used as input file for SSRLocator software available in http://www.ufpel.tche.br/faem/fitotecnia/fitomelhoramento/faleconosco.html. The SSR description and primer pairs used in this study were provided by the program. As a template we used genomic DNA extracted from *P. japonicum* populations collected at different regions across Japan (Gunma prefecture in the Kusatsu city, Nagano prefecture in Karuizawa city, Tokyo region around Mount Takao, Shizuoka in the Isu peninsula, Hiroshima, Okayama and Kyushu prefectures). Total genomic DNA was extracted from 100 mg leaves using the Phytopure DNA extraction kit (GE healthcare, Little Chalfont, England). The PCR amplification was performed with hot start at 95^{o}C for 3min, then 40 cycles of denaturation (95^{o}C 15 s), annealing (55^{o}C 15 s), extension (72^{o}C 30 s), and a final extension of 5 min at 72^{o}C. The informative band size were screening using two methods the traditional gel-based

electrophoresis with low melting temperature agarose (Lonza Nusieve®GTG, code 50080) at 4% (w/v) and MCE-202 MultiNa Microchip Electrophoresis System (Shimadzu, Japan) with DNA-500 premix. The 23 alleles detected by 12 SSR markers were analyzed in the DARwin software using as input file a matrix composed by “1” or “0” for allele presence of absence, respectively . The tree was visualized by Mega 5 software (Tamura et al., 2011).

8

APPENDIX

8.1 PRIMER LIST

The list of primers used in this thesis is shown as table with their respective length, amplicon size and their melting temperature (TM), followed by the figure localizing the primer sequence on their respective target.

Table 9.1 Primer sequences target to *PjQR1* gene

Primer	Sequence	Direction	length	Amplicon (bp)	TM	Experiment
PjQR1-Race-F	CCAAGGCCCGGGAGTTATATCGATCAC	Forward	27	831	72	RACE
PjQR1-Race-R	TCGCAAAGCTAGGAAACGCACACGTAA	Reverse	27	615	71.65	RACE
PjQR1-Race-Nested-F	GAGCTTCATGCCAAAGGATTGGGTGAG	Forward	27	516	72	RACE
PjQR1-Race-Nested-R	TGGTGCCGCTACTTGTTGAAGTCAAGG	Reverse	27	324	71.65	RACE
PjQR1-Full-F	TGATCGAACAAGATATAAAATACGACATAA	Reverse	30	1204	61.3	Amplify full length
PjQR1-Full-New-F	CAAACCCTCTACATAACACACAAAGG	Forward	26	1204	62.6	Amplify full length
PjQR1-9469-qPCR-F	TATTTGCCACCCGTAACCAC	Forward	20	78	60.6	qRT-PCR
PjQR1-9469-qPCR-R	GGAGACATGGCTGATTTTGG	Reverse	20	78	60.5	qRT-PCR

Table 9.2 Primer sequences target to *PjQR2* gene

Primer	Sequence	Direction	length	Amplicon (bp)	TM	Experiment
PjQR2-Race-F	CCAGCTGGCATCTTTTGTAGCACTGGA	Forward	27	463	71.73	RACE
PjQR2-Race-R	AGCGCCATATGGACTGCCACCCTTAAT	Reverse	27	576	72.14	RACE
PjQR2-Race-Nested-F	GGACATCGAACTCGCACAAGCTTTTCA	Forward	27	251	71.77	RACE
PjQR2-Race-Nested-R	CCAAACGTGTAACCGATGGGGACAAAA	Reverse	27	516	72.32	RACE
PjQR2-Full-F	CCAACCAACTCATACTAAACCAAA	Forward	24	857	59.37	Amplify full length
PjQR2-Full-R	GGCAAGAAATCATTGGCATC	Reverse	20	857	60.42	Amplify full length
Pj_22303-F	AAGGTGGCGGCCAAGAAACCAC	Forward	22	151	60.1	qRT-PCR
Pj_22303-R	AACGTGCCAGCGCCATATGGAC	Reverse	22	151	60.1	qRT-PCR
Pj2-11944qPCR-QR2-F	ATGTACATCGCAGGCATCAC	Forward	20	71	59.6	qRT-PCR
Pj2-11944qPCR-QR2-R	GGATGCAATTAGCATGATCG	Reverse	20	71	59.1	qRT-PCR

Table 9.3 Primer sequences target to *ShQR1* gene

Primer	Sequence	Direction	length	Amplicon (bp)	TM	Experiment
ShQR1-qPCR-F	CGACAAAGTTGTAGCCATGC	Forward	20	82	59.3	qRT-PCR
ShQR1-qPCR-R	TACGGTCTGGTTTTCCTTGG	Reverse	20	82	60	qRT-PCR
ShQR1-full-F	CGCCGCTCATCAGTTCTACT	Forward	20	1375	60.56	Amplify full length
ShQR1-full-R	CCCAATTGCCAACTTTATTCA	Reverse	21	1375	59.82	Amplify full length

Table 9.4 Primer sequences target to *ShQR1* gene

Primer	Sequence	Direction	length	Amplicon (bp)	TM	Experiment
ShQR2-qPCR-F	ACGACGTGCCGATTATTAGC	Forward	20	84	60.1	qRT-PCR
ShQR2-qPCR-R	CCATCATCCCGAATCTTGTC	Reverse	20		60.3	qRT-PCR
ShQR2-full-F	CACACTTCACACACCAAATCAA	Forward	22	967	59.51	Amplify full length
ShQR2-full-R	GAATATGCAATTGCCCGAGT	Reverse	20		59.93	Amplify full length

Table 9.5 Primer sequences target to *PjYUC1* gene

Primer	Sequence	Direction	length	Amplicon (bp)	TM	Experiment
PjYUC-12867-Race-F	GCAACAGTTCGTGCGCTACCTGGAGGA	Forward	27	667	75.1	RACE
PjYUC-12867-Race-R	CCCGTATCACCCAGCAAGAGCCATGAA	Reverse	27	920	74.61	RACE
PjYUC-12867-Race-Nested-F	CAACTTTCGGGCTGTCCATGTGGTTGC	Forward	27	667	75.12	RACE
PjYUC-12867-Race-Nested-R	TTGCTTGGAGGGGTAGGTCGGGAAGGT	Reverse	27	408	74.78	RACE
Pj12867-Full-new-F	CGCAGGCTCGCAGGCTCTAT	Forward	20	1487	67.3	Amplify full length
Pj12867-Full-new-R	CCCCAAAAATTCTCTCTTCATTCC	Reverse	24	1487	63.25	Amplify full length
PjYUC-12867-qPCR-R	CAAGAACGGAGGGTTTTACG	Forward	20	99	59.2	qRT-PCR
PjYUC-12867-qPCR-F	GTGGCATGATGATTGCAGAG	Reverse	20	99	60.2	qRT-PCR

Table 9.6 Primer sequences target to *PjYUC2* gene

Primer	Sequence	Direction	length	Amplicon (bp)	TM	Experiment
PjYUC-16808-Race-F	TCCGGCGAGAGTTACAGGGGGAAAAGG	Forward	27	873	75.04	RACE
PjYUC-16808-Race-R	GCTTTTCGCCGGTAACAAGCTCCACCA	Reverse	27	990	75	RACE
PjYUC-16808-Race-Nested-F	CACCGGATATCGCAGCAACGTGCCTTA	Forward	27	418	75	RACE
PjYUC-16808-Race-Nested-R	TCCCATGAGCGTTCTTCAGCTCCAACG	Reverse	27	873	75.14	RACE
Pj16808-Full-new-F	CAATTCACGCAAAACAAAAGAAAATG	Forward	26	1433	64.62	Amplify full length
Pj12867-Full-new-F	CGCAGGCTCGCAGGCTCTAT	Forward	20	1433	67.3	Amplify full length
PjYUC-16808-qPCR-F	TAGCTCGGTTCATGTGTTGC	Forward	20	144	59.9	qRT-PCR
PjYUC-16808-qPCR-R	TTATCCGTGCTTCCGAGAAC	Reverse	20	144	60.2	qRT-PCR

Table 9.7 Primer sequences target to *PjYUC3* gene

Primer	Sequence	Direction	length	Amplicon (bp)	TM	Experiment
YUC24658-Race-F	GGTTTGGGTAGTGGACCGGGTACTGCT	Forward	775	27	72.01	RACE
YUC24658-Race-nested-F*	GGTGGAAAGGCAAAGACGGGCTTTATG	Forward	425	27	72.13	RACE and qRT-PCR
YUC24658-Race-R	CAATACATGAACCGAGCTTCGCACCAC	Reverse	851	27	71.85	RACE
YUC24658-Race-nested-R	GACGAACTGGCTCCTGGAGGGGTATTC	Reverse	473	27	72.04	RACE
PjYUC-Full-F	CCCTATTTAAACATCTCCAACTCC	Forward	1688	24	59.32	Amplify full length
PjYUC-Full-New-R	CTAACCAAAACATTTGATCTGGCAATG	Reverse	1688	27	65	Amplify full length and Overexpression cloning
PjYUC-pUB-GW-F	caccATGAATCATTGCGTGCCG	Forward	1518	18	61.64	Overexpression cloning
PjYUC-pro-TOPO-R	CGAGAAGAAGGCCTCTTGGTCGAG	Reverse		24	68.65	Primer walking to confirm the promoter sequence
PjYUC-pro-TOPO-R	CGAGAAGAAGGCCTCTTGGTCGAG	Reverse		24	68.65	Primer walking to confirm the promoter sequence
PjYUC-pro-2nd-F	TCAATGCTTTCTAGGTGTTCG	Forward		21	58.04	Primer walking to confirm the promoter sequence
PjYUC-pro-3nd-F	TGCTCATCTTGTTGGTCCTG	Forward		20	59.83	Primer walking to confirm the promoter sequence
PjYUC-pro-3nd-R	GGGCTTGTGACTCAAACTAGC	Reverse		21	59	Primer walking to confirm the promoter sequence
ProPjYUC-GW-attB4-F	GGGGACAACTTTGTATAGAAAAGTTGT GGACTCACAACCGGACAGATG		3166	49		Promoter cloning
ProPjYUC-GW-attB1-R	GGGGACTGCTTTTTTGTACAAACTTGGA GCAAGGTGTGAGTGAGTGA		3166	48		Promoter cloning
Pj2-24658qPCR-YUC-R*	CATATTCAAGCGAGGCCAAC	Reverse	160	20	60.6	qRT-PCR
MA-Pj_52606-qPCR-F	AATCAACCGCGCTGCTGCCA	Forward	203	20	60.2	qRT-PCR
MA-Pj_52606-qPCR-R	TCGGCTGTGGGAATTCGGGGAT	Reverse		22	59.9	qRT-PCR

***** These primers form pair for qRT-PCR experiments

Table 9.8 Primer sequences target to *PjYUC4* gene

Primer	Sequence	Direction	length	Amplicon (bp)	TM	Experiment
PjYUC-2021-Race-F	TTCGGGCATGGAGGTTTCGCTTGATCT	Forward	27	890	74.98	RACE
PjYUC-2021-Race-R	GAACGACCATGGCGGGCTTAGCATCAT	Reverse	27	643	75	RACE
PjYUC-2021-Race-Nested-F	TGAGAGGATCCGGTCCAGGGACATTCA	Forward	27	599	74.84	RACE
PjYUC-2021-Race-Nested-R	GTACTTGGCGGACTGCACGCACTCGTT	Reverse	27	357	75	RACE
PjYUC-2021-Full-F	ATGTTTAGTTTCACAGACACTGACATATTC	Forward	30	1471	61.18	Amplify full length
PjYUC-2021-Full-R	CAAGGAAAATAAGCCTCGTCATCC	Reverse	24	1471	64.18	Amplify full length
PjYUC-2021-qPCR-F	CTTTATGCGGTCGGGTTTAC	Forward	20	127	59.5	qRT-PCR
PjYUC-2021-qPCR-R	GCGTAGGCACTTTTTGCTTC	Reverse	20	127	60	qRT-PCR

Table 9.9 Primer sequences target to *PjSBT2* gene

Primer	Sequence	Direction	length	Amplicon (bp)	TM	Experiment
PjSBT2-3'end-R	CACTCTTCTTTTCACTGCCCTTATTCCATT	Reverse	30	890	65	RACE
PjSBT2-nested-R	CTCAAAATCGCCCTCATCATCACACAAAAC	Reverse	30	1344	65	RACE
PjSBT3-3'end-R	CACTCTTCTTTTCACTGCCCTTATTCCATT	Reverse	30	2148	65	RACE
PjSBT3-nested-R	CTCAAAATCGCCCTCATCATCACACAAAAC	Reverse	30	1344	65	RACE
Pj_3988-Race-5`end	CAGTCCAATTGCACCTTGTTCCAGCA	Reverse	27	1395	71	RACE
Pj_3988-Race-3`end	AGGCATCATTGCGGCATATCCGTCTT	Forward	26	890	72.5	RACE
Pj_41594-Race-5`end	CACGGATGCGTTGTAGCCCATGTTG	Reverse	25	2004	72.8	RACE
PjSBT2-full-2R	GCTCTTGATCCCGTAAGTGC	Reverse	20	1877		RACE
PjSBT2-full-2F	TGCAACAGGAAATTGATTGG	Forward	20	1894	59.52	RACE
Pj_41594-Race-3`end	GCCCATTTGCTATTAGCCATCCAGCA	Forward	26	177	71.1	RACE
Pj_3988_fullcDNA-F	agagcataagcttaacgttgaggaa	Forward	25	2530	61.81	Amplify full length
Pj_3988_fullcDNA-R	tcagaattacgcgcatcaaa	Reverse	20	2530	60.36	Amplify full length
Seq-Pj_3988-qPCR-F	CTTCTCAGCAAGAGGCCCTA	Forward	20	155	60	qRT-PCR
Seq-Pj_3988-qPCR-R	GCGACGTACCCGACAGTATT	Reverse	20	155	60	qRT-PCR

Table 9.10 Primer sequences target to *PjSBT4* gene

Primer	Sequence	Direction	length	Amplicon (bp)	TM	Experiment
PjSBT4-Race-GSP1-R	CCCGACCGCGCCTTGTTTCATT	Reverse	22	1202	68	RACE
PjSBT4-Race-nested-R	TTTGGTTGGAGCTGGCTGAGCGT	Reverse	23	1194	68	RACE
PjSBT4-Race-GSP1-F	CCTCTCTGTCCGGGGCTCGTCT	Forward	22	564	67.7	RACE
PjSBT4-Race-nested-F	GGCTCCGGCTGGTATTGAGGTT	Forward	22	292	65.9	RACE
Pj2620-3'end-R	CTGTCCTGCTTCCATTTGCCTTTAAT	Reverse	26	2082	63	RACE
Pj2620-nested-R	TACATTGGTTTGGTTGGAGCTGGCTG	Reverse	26	1202	65.3	RACE
PjSBT4-full-F	CACTGACATGGACTGAAGGAG	Forward	21	2432	58.29	Amplify full length
PjSBT4-full-R	AAGCGTAAGTAATAGTTTTTATTGAT	Reverse	26	2432	53.84	Amplify full length
PjSBT4-qPCR-2620_F	CCATTGGCGACAATCTTACC	Forward	20	81	60.3	qRT-PCR
PjSBT4-qPCR-2620_R	GCCCCTTGATGAAGAGTTTG	Reverse	20	81	59.7	qRT-PCR

Table 9.11 Primer sequences target to *PjSBT7* gene

Primer	Sequence	Direction	length	Amplicon (bp)	TM	Experiment
PjSBT7-Race-GSP1-R	CCCCCCCAACAAAACATCAGTCTCA	Reverse	25	1115	66	RACE
PjSBT7-Race-nested-R	GCAATCATCAGCTCCACAAACACG	Reverse	24	839	63.7	RACE
PjSBT7-Race-GSP1-F	GCTCCCGTCGCTGCATTATTTTC	Forward	23	951	63.52	RACE
PjSBT7-Race-nested-F	AATTCAACCTACTACGCCACGG	Forward	22	333	62.9	RACE
PjSBT7-full-new-F	AATGTGAAAGTGTGAACTCGTGA	Forward	23	2480	59.7	Amplify full length
PjSBT7-full-R	GAAAAATGCCATGCTAAATTCAT	Reverse	23	2480	59.4	Amplify full length
PjSBT7-qPCR-6068_F	AAGGCAAATGGAAGGAGGAC	Forward	20	97	60	qRT-PCR
PjSBT7-qPCR-6068_R	GCATGCCAATAGGAGTTTCC	Reverse	20	97	59.5	qRT-PCR

Table 9.12 Primer sequences target to *PjSBT8* gene

Primer	Sequence	Direction	length	Amplicon (bp)	TM	Experiment
PjSBT8-Race-1R	GTGTGTGTTCCATGGCC**G**TCT**GT**ATCT	Reverse	27	785	68	RACE
PjSBT8-Race-2R	CCTTATTCCAC**A**GAGAGCTGGCTCCAA	Reverse	27	523	68	RACE
PjSBT8-Race-1F	GCTGCTTACAAGGTTTGCTGGCCTAA**A**	Forward	26	1742	67	RACE
PjSBT8-Race-2F	CCGTTTCACAAAGAC**G**GTATTGCCATT	Forward	27	1622	65	RACE
PjSBT8-exon2-R	TCTGTCATAATCAGTGGGTGT	Reverse	21	255	57	RACE
PjSBT8-full-F	CTTACACCAACGTTTTCC	Forward	18	2621	53	Amplify full length
PjSBT8-Full-1R-new	TATGAGGTTTGCACGA	Reverse	17	2621	51	Amplify full length
PjSBT8-qPCR-35234_F	CAAGGTTTGCTGGCCTAA**A**G	Forward	20	86	60	qRT-PCR
PjSBT8-qPCR-35234_R	ACA**A**CACATCGACTCCATCG	Reverse	20	86	60	qRT-PCR
proPjSBT8-3F	TCTTACTCGGCCTCATCTCG	Forward	20		61	Primer walking to confirm the promoter sequence
proPjSBT8-2F	CCACGTTTGTGATGATCT	Forward	18		55	Primer walking to confirm the promoter sequence
Pj35234_pro_F	TATAGAAAAGTTGGTCCCCATCAGCATCAGT	Forward	32		59	Promoter cloning
Pj35234_pro_R	TTGTACAAACTTGGCTCTTCCTTTGCTAAATTTCTCTC	Reverse	39		59	Promoter cloning

* in red the nucleotides which are in conflict with the Sanger-confirmed target sequence

Table 9.13 Primer sequences target to *PjSBT11* gene

Primer	Sequence	Direction	length	Amplicon (bp)	TM	Experiment
PjSBT5-full-F	CAAAATGGTGCTGGAAA	Forward	17	2351	51	RACE
PjSBT11-Seq1-R	ATAGCAGCCTCCCAACTTCC	Reverse	20	865	59.9	RACE
PjSBT11-Seq2-F	TGCTGCTGTCAGTGAAGAGG	Forward	20	1101	60	RACE
PjSBT11-Seq3-R	TTGTTTGGCACCACTTTAACC	Reverse	21	2242	59.89	RACE
PjSBT5-Seq1-R	GCAGCAATTGGAAAAACTCC	Reverse	20	1450	59.69	RACE
PjSBT5-full-3F	CCTTGCATTCTTTTCCTCCA	Forward	20	981	60.18	RACE
PjSBT5-full-3R	GTTAGGATCCAAGGGGCAAG	Reverse	20	1084	60.82	RACE
PjSBT5-full-2F	GGTTCTGGGTGCATCGTACT	Forward	20	1800	60	RACE
PjSBT11-full-2F	TGACACTATCATTGGCGTCAT	Forward	21	2046	59.02	RACE
PjSBT11-full-2R	TTATTTCCCCTGATCCGATG	Reverse	20	1920	59.72	RACE
PjSBT11-exon1-R	TCACTTTTTCCAGCACCATT	Reverse	20	217	56.45	RACE
PjSBT11-full-F	CGCTGTGGAAGTGTGAA	Forward	17	2545	53.8	Amplify full length
PjSBT11-full-R	AGAGATCAAGTCACAGA	Reverse	17	2545	50	Amplify full length
PjSBT5-qPCR-19312_F	ACTCGCAGCATATCGTGTTG	Forward	20	100	59.9	qRT-PCR
PjSBT5-qPCR-19312_R	TAACACATCGACCCCATCAG	Reverse	20	100	59.4	qRT-PCR
PjSBT11-qPCR-57900_F	GATCATGCTCAGCTTCTGAGTG	Reverse	22	88	60.2	qRT-PCR
PjSBT11-qPCR-57900_R	CCGCAAATCCCAAGAAAC	Forward	18	88	59	qRT-PCR

* in red the nucleotides which are in conflict with the Sanger-confirmed target sequence

Table 9.14 Primer sequences target to *PjSBT3*gene

Primer	Sequence	Direction	length	Amplicon (bp)	TM	Experiment
PjSBT22231-Full-F	TACAATGCGCAATCCCTTTT	Forward	20	2339	60.46	Amplify full length
PjSBT22231-Full-R	GAAGCTGTTGCAGAAAATGC	Reverse	20	2339	58.65	Amplify full length
Pj22232-group3-F	ACACGTCACAGGATGTCCCGCTA	Reverse	23	118	59.69	qRT-PCR
Pj22232-group3-R	TGCGTCATGTCCCGGGTTGA	Forward	20	118	58.35	qRT-PCR

Table 9.15 Primer sequences target to *PjSBT6* gene

Primer	Sequence	Direction	length	Amplicon (bp)	TM	Experiment
Pj39971-group6-F	TGCGGACATGACTCGAGGTTTGA	Forward	23	119	58.35	qRT-PCR
Pj39971-group6-R	TTTAAACCCCCGCCGCCCAC	Reverse	20	119	59.90	qRT-PCR
PjSBT39971-Full-F	ATGATCCGCCATCTCAACTC	Forward	20	3239	60.04	Amplify full length
PjSBT39971-Full-R	CCTGGACAAACGTTGCACTA	Reverse	20	3239	59.76	Amplify full length

Table 9.16 Primer sequences target to *ShSBT* genes

Primer	Sequence	Direction	length	Amplicon (bp)	TM	Experiment
Sh_146-nested-F	GTGATACACCACAGCCTGCTCTGCTTC	Reverse	27		71	RACE
Sh_146-nested-R	CCGTTGAATCTGATTGAGTTCGAGCAA	Forward	27		70	RACE
ShSBT1-Full-F	ATGACAGGGTTGCATTACATTTC	Forward	23	2410	60	Amplify full length
ShSBT1-Full-R	CGCCCGTGCTACATAGAGAT	Reverse	20		60	Amplify full length
ShSBT4-Full-F	TCATGAGAGGCATCCATCTTT	Forward	21	2647	60	Amplify full length
ShSBT4-Full-R	CATCGCACGTGTCCCTTACT	Reverse	20		61	Amplify full length
ShSBT8-Full-F	GCAAGGATGACACGCATAAA	Forward	20	2487	60	Amplify full length
ShSBT8-Full-R	AATTGGCTCCAAAAAGCAAA	Reverse	20		60	Amplify full length
Sh9Contig_146-F	GTGTCGATAAGCCCAACGAT	Forward	20	99	59	qRT-PCR
Sh9Contig_146-R	CACCACAAGAAGCTGGGATT	Reverse	20	99	60	qRT-PCR
ShSBT4-qPCR-F	TCAAACTGCTCAGCAACGTC	Forward	20	112	60	qRT-PCR
ShSBT4-qPCR-R	AGCGTGAAGGGATTTTACCC	Reverse	20	112	60	qRT-PCR
ShSBT8-qPCR-F	AGCACAAGGACCCACAAATC	Forward	20	111	60	qRT-PCR
ShSBT8-qPCR-R	TTCAAGAGGCCGACAATACC	Reverse	20	111	60	qRT-PCR

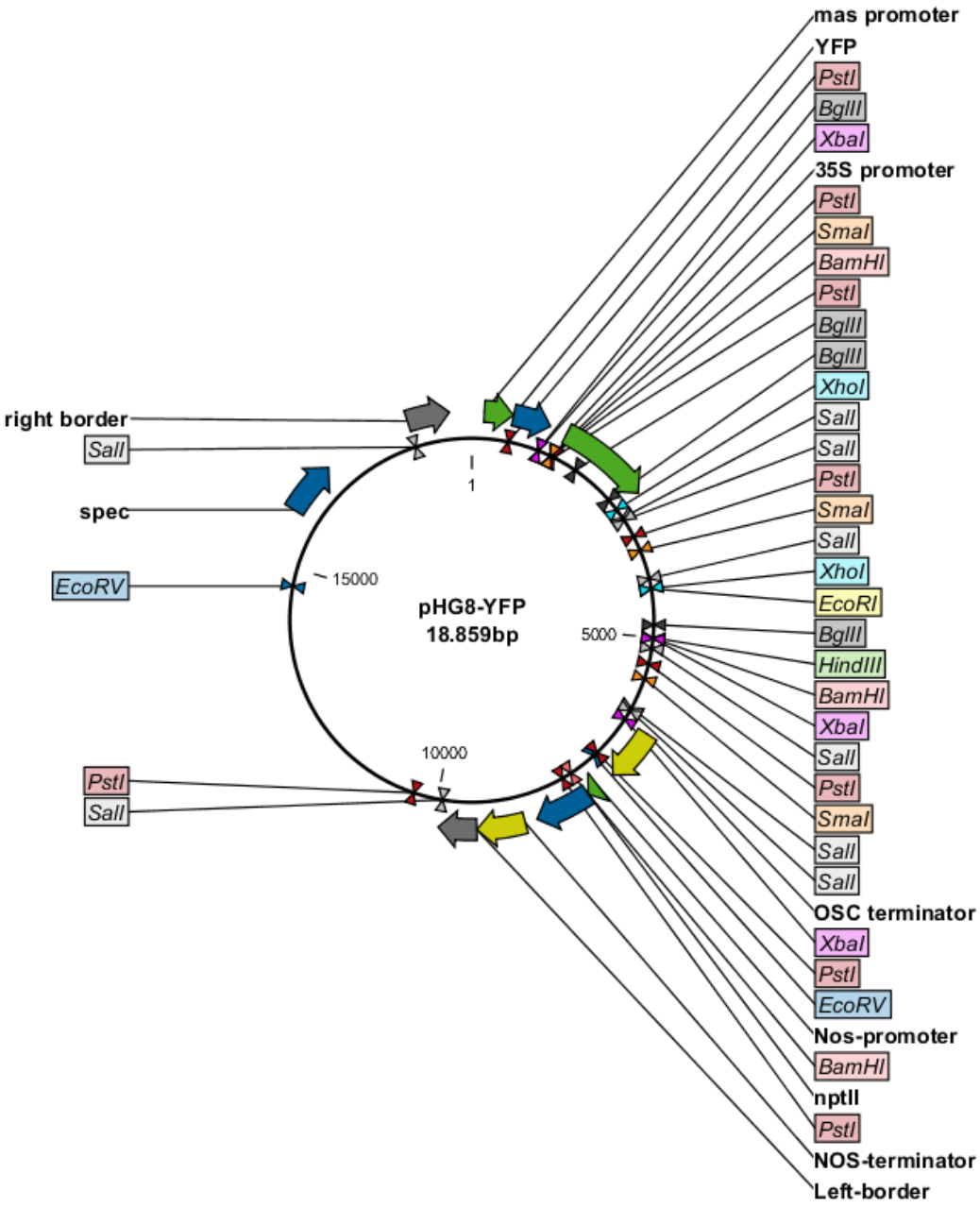
Table 9.17 Other primer sequences used in this thesis

Primer	Sequence	Direction	Target	length	Amplicon (bp)	TM	Experiment	Reference
Pj_UBC27-F	ACGCACAGCTCGTTACTGGACAG	Forward	UBC27ubiquitin-protein ligase	23	140	59	qRT-PCR	This thesis
Pj_UBC27-R	TTTCGTCACCGCCAACGCTCTC	Reverse	UBC27ubiquitin-protein ligase	22	140	59	qRT-PCR	This thesis
Pj_PTB-F	TCCGATGCAACAAGCTCCTGGG	Forward	Polypyrimidinetract-binding protein (PTB)	22	116	59	qRT-PCR	This thesis
Pj_PTB-R	ATGTGCCAGGAGCGGACACAAA	Reverse	Polypyrimidinetract-binding protein (PTB)	22	116	59	qRT-PCR	This thesis
Sh-UB1-qPCR-F	CATCCAGAAAGAGTCGACTTTG	Forward	Ubiquitin	23	122	52	qRT-PCR	(Fernández-Aparicio et al., 2013)
Sh-UB1-qPCR-R	CATAACATTTGCGGCAAATCA	Reverse	Ubiquitin	22	122	50	qRT-PCR	(Fernández-Aparicio et al., 2013)
Lj ATPsyn-F	CAATGTCGCCAAGGCCCATGGTG	Forward	Lotus ATP synthase	24		63	qRT-PCR	(Lopez-Gomez et al., 2012)
Lj ATPsyn-R	AACACCACTCTCGATCATTTCTCTG	Reverse	Lotus ATP synthase	26		55	qRT-PCR	(Lopez-Gomez et al., 2012)
Os_b-Actin-F	CCTGACGGAGCGTGGTTAC	Forward	Rice_Actin	20		58	qRT-PCR	(Li et al., 2012)
Os_b-Actin-R	CCAGGGCGATGTAGGAAAGC	Reverse	Rice_Actin	21		58	qRT-PCR	(Li et al., 2012)
R4L1pGWB532-F	CTTTATGCTTCCGGCTCGTA	Forward	R4L1pGWB532 vector	20		60	sequencing	This thesis
R4L1pGWB532-R	TTGTGCCCATTAACATCACC	Reverse	R4L1pGWB532 vector	20		60	sequencing	This thesis
pUB-NosT-F	ATTGCCAAATGTTTGAACGA	Forward	pUB-GW-GFP	20		59	sequencing	This thesis
pUB-5'UTR-R	TGTTTCAATTTCGTGTGTTGG	Reverse	pUB-GW-GFP	21		59	sequencing	This thesis
35SPro-pHGT-F	CCACTATCCTTCGCAAGACC	Forward	pHG8YFP vector	21		54	sequencing	This thesis
OscTer-pHGT-R	GGCGGTAAGGATCTGAGCTA	Reverse	pHG8YFP vector	21		55	sequencing	This thesis

8.2 VECTOR MAPS

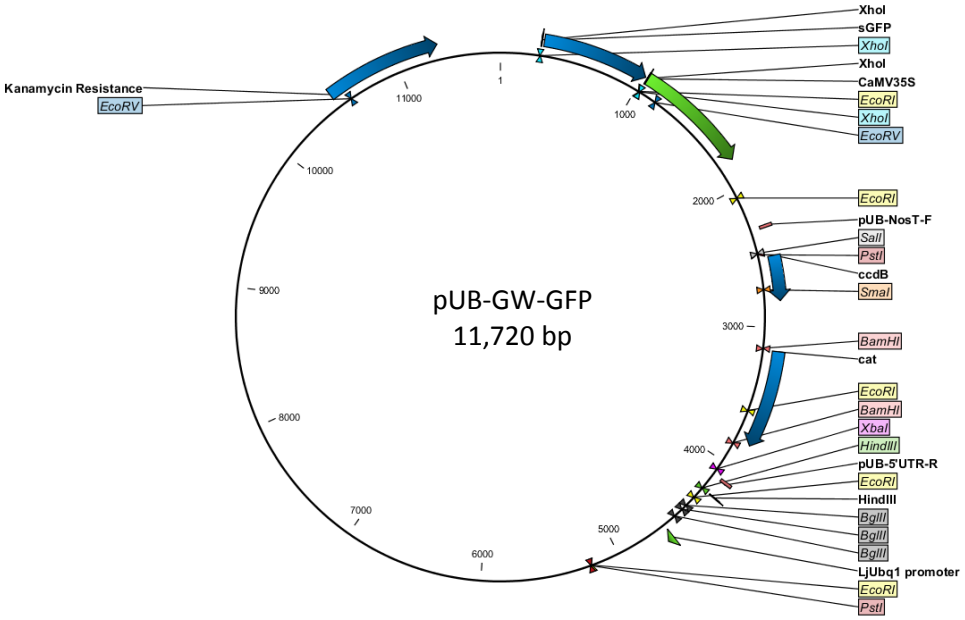
8.2.1 SILENCING VECTORS

pHG8YFP

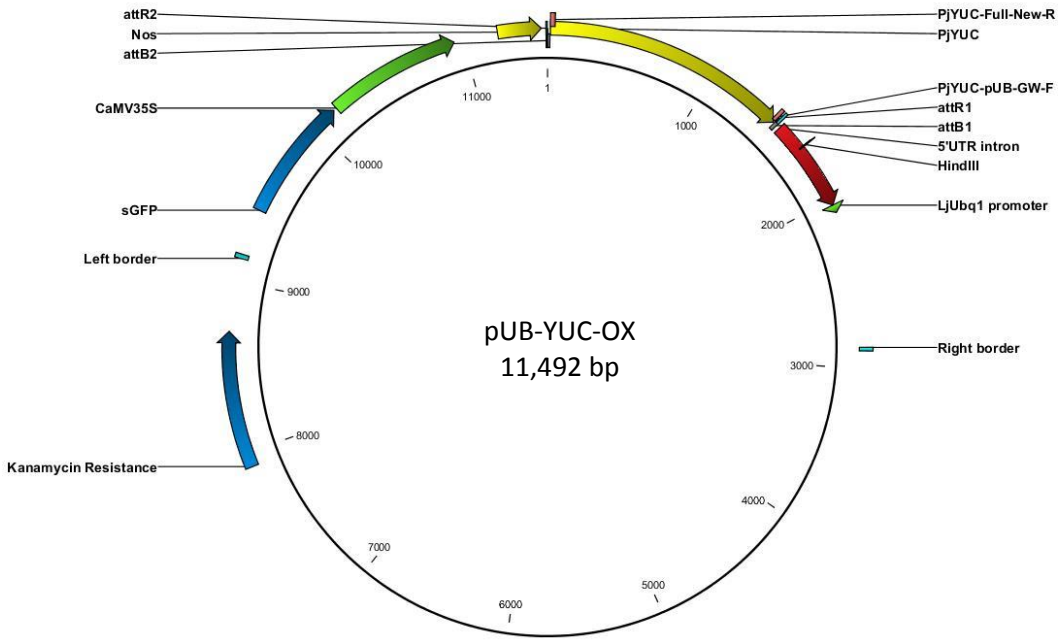


8.2.2 OVEREXPRESSING VECTORS

pUB-GW-GFP

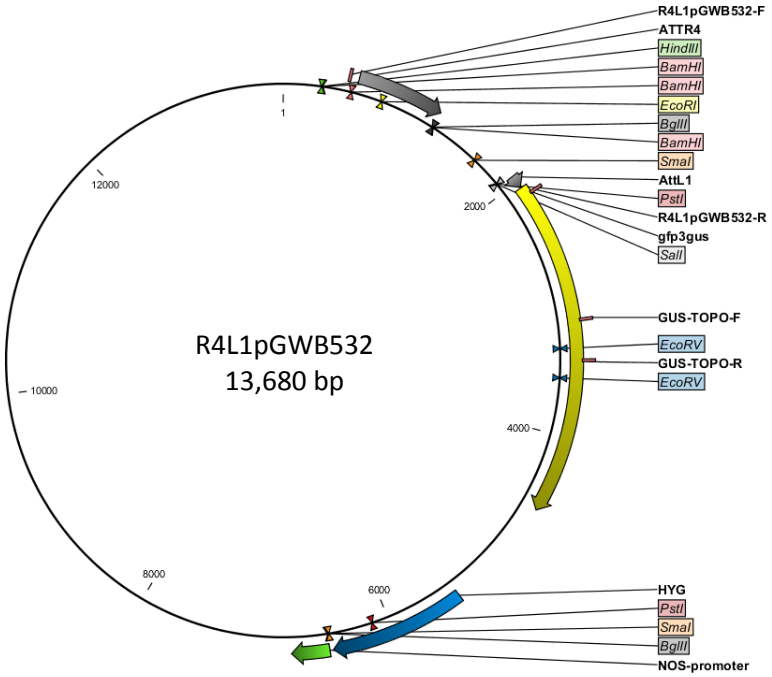


pUB-YUC-OX

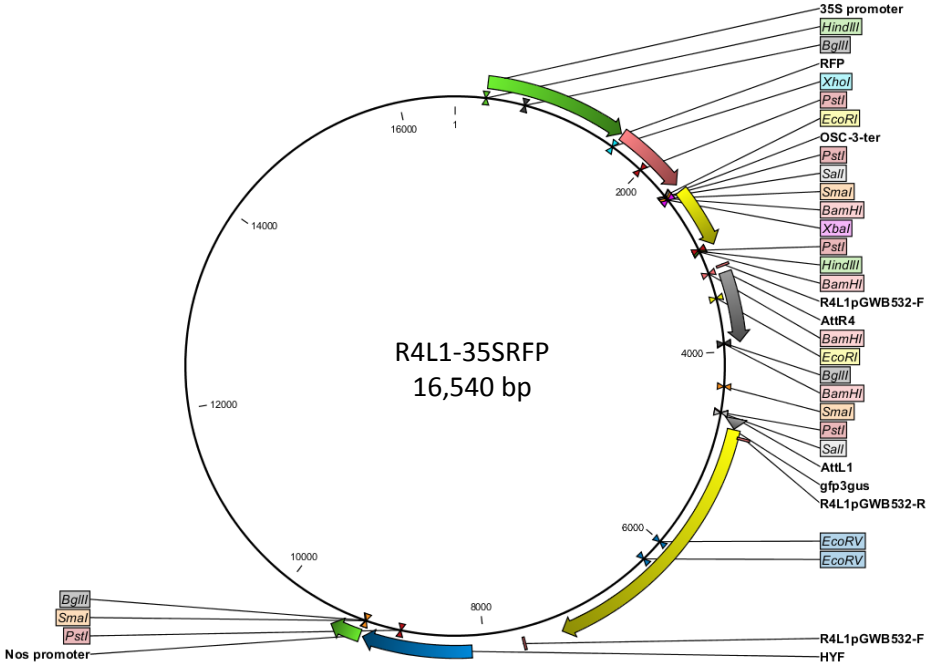


8.2.3 VECTOR FOR PROMOTER ANALYSIS

R4L1pWB532



R4L1-35SRFP



9

REFERENCES

Afzal, A.J., Natarajan, A., Saini, N., Iqbal, M.J., Geisler, M., El Shemy, H.A., Mungur, R., Willmitzer, L., and Lightfoot, D.A. (2009). The nematode resistance allele at the rhg1 locus alters the proteome and primary metabolism of soybean roots. Plant Physiol. **151**: 1264–80.

Agusti, J., Herold, S., Schwarz, M., Sanchez, P., Ljung, K., Dun, E.A., Brewer, P.B., Beveridge, C.A., Sieberer, T., Sehr, E.M., and Greb, T. (2011). Strigolactone signaling is required for auxin-dependent stimulation of secondary growth in plants. Proc. Natl. Acad. Sci. U. S. A. **108**: 20242–7.

Ahonsi, M.O., Berner, D.K., Emechebe, A.M., Lagoke, S.T., and Sanginga, N. (2003). Potential of ethylene-producing pseudomonads in combination with effective N2-fixing bradyrhizobial strains as supplements to legume rotation for Striga hermonthica control. Biol. Control **28**: 1–10.

Akiyama, K., Matsuzaki, K., and Hayashi, H. (2005). Plant sesquiterpenes induce hyphal branching in arbuscular mycorrhizal fungi. Nature **435**: 824–7.

Akiyama, K., Ogasawara, S., Ito, S., and Hayashi, H. (2010). Structural requirements of strigolactones for hyphal branching in AM fungi. Plant Cell Physiol. **51**: 1104–17.

Alakonya, A., Kumar, R., Koenig, D., Kimura, S., Townsley, B., Runo, S., Garces, H.M., Kang, J., Yanez, A., David-Schwartz, R., Machuka, J., and Sinha, N. (2012). Interspecific RNA Interference of SHOOT MERISTEMLESS-Like Disrupts Cuscuta pentagona Plant Parasitism. Plant Cell **24**: 3153–3166.

Albersheim, P., Darvill, A., Roberts, K., Sederoff, R., and Staehelin, A. (2011). Plant cell wall first. P. Albersheim, A. Darvill, K. Roberts, R. Sederoff, and A. Staehelin, eds (Garland Science).

Albrecht, H. (1999). Flavonoids Promote Haustoria Formation in the Root Parasite Triphysaria versicolor. PLANT Physiol. **119**: 585–592.

Alejandro, P.-L. (2013). Haustorium Invasion into Host Tissues. In Parasitic Orobanchaceae, D.M. Joel, J. Gressel, and L.J. Musselman, eds (Springer: New York), pp. 75–86.

Alonso, L., Fernández-Escobar, J, López, G., Rodríguez-Ojeda, M., and Sallago, F. (1996). New highly virulent sunflower broomrape (Orobanche cernua Loefl.) pathotype in Spain. In Advances in parasitic plant research, Proceedings of the 6th International Symposium in Parasitic Weed, Córdoba, 16–18 April, Spain, M. Moreno, J. Cubero, D. Berner, D. Joel, L. Musselman, and C. Parker, eds (Consejería de Agricultura y Pesca: Sevilla, Spain), pp. 639–644.

Aly, R., Cholakh, H., Joel, D.M., Leibman, D., Steinitz, B., Zelcer, A., Naglis, A., Yarden, O., and Gal-On, A. (2009). Gene silencing of mannose 6-phosphate reductase in the parasitic weed Orobanche aegyptiaca through the production of homologous dsRNA sequences in the host plant. Plant Biotechnol. J. **7**: 487–498.

Aly, R., Hamamouch, N., Abu-Nassar, J., Wolf, S., Joel, D.M., Eizenberg, H., Kaisler, E., Cramer, C., Gal-On, A., and Westwood, J.H. (2011). Movement of protein and macromolecules between host plants and the parasitic weed Phelipanche aegyptiaca Pers. Plant Cell Rep. **30**: 2233–41.

APHIS (2011). Witchweed: A Parasitic Pest. Plant Prot. Quar.

Apweiler, R. (2001). The InterPro database, an integrated documentation resource for protein families, domains and functional sites. Nucleic Acids Res. **29**: 37–40.

Arite, T., Umehara, M., Ishikawa, S., Hanada, A., Maekawa, M., Yamaguchi, S., and Kyozuka, J. (2009). d14, a strigolactone-insensitive mutant of rice, shows an accelerated outgrowth of tillers. Plant Cell Physiol. **50**: 1416–24.

Atera, E.A., Itoj, K., and Onyango, J.C. (2011). Evaluation of ecologies and severity of Striga weed on rice in sub-Saharan Africa. Agric. Biol. J. NORTH Am. **2**: 752–760.

BABIKER, A.G.T., HAMDOUN, A.M., RUDWAN, A., MANSI, N.G., and FAKI, H.H. (1987). Influence of soil moisture on activity and persistence of the strigol analogue GR 24. Weed Res. **27**: 173–178.

Bandaranayake, P.C.G., Filappova, T., Tomilov, A., Tomilova, N.B., Jamison-McClung, D., Ngo, Q., Inoue, K., and Yoder, J.I. (2010). A single-electron reducing quinone oxidoreductase is necessary to induce haustorium development in the root parasitic plant Triphysaria. Plant Cell **22**: 1404–19.

Bandaranayake, P.C.G., Tomilov, A., Tomilova, N.B., Ngo, Q.A., Wickett, N., DePamphilis, C.W., and Yoder, J.I. (2012). The TvPirin gene is necessary for haustorium development in the parasitic plant Triphysaria versicolor. Plant Physiol. **158**: 1046–53.

Bandaranayake, P.C.G. and Yoder, J.I. (2013a). Haustorium Initiation and Early Development. In D.M. Joel, J. Gressel, and L.J. Musselman, eds (Springer: New York), pp. 61–71.

Bandaranayake, P.C.G. and Yoder, J.I. (2013b). Trans-Specific Gene Silencing of Acetyl-CoA Carboxylase in a Root-Parasitic Plant.

Bar-Nun, N., Sachs, T., and Mayer, A.M. (2008). A role for IAA in the infection of Arabidopsis thaliana by Orobanche aegyptiaca. Ann. Bot. **101**: 261–5.

Bennett, J.R. and Mathews, S. (2006). Phylogeny of the parasitic plant family Orobanchaceae inferred from phytochrome A. Am. J. Bot. **93**: 1039–51.

Bleischwitz, M., Albert, M., Fuchsbauer, H.-L., and Kaldenhoff, R. (2010). Significance of Cuscutain, a cysteine protease from Cuscuta reflexa, in host-parasite interactions. BMC Plant Biol. **10**: 227.

Boerjan, W. (1995). superroot, a Recessive Mutation in Arabidopsis, Confers Auxin Overproduction. PLANT CELL ONLINE **7**: 1405–1419.

Booker, J., Auldridge, M., Wills, S., McCarty, D., Klee, H., and Leyser, O. (2004). MAX3/CCD7 is a carotenoid cleavage dioxygenase required for the synthesis of a novel plant signaling molecule. Curr. Biol. **14**: 1232–8.

Boyer, F.-D., de Saint Germain, A., Pillot, J.-P., Pouvreau, J.-B., Chen, V.X., Ramos, S., Stévenin, A., Simier, P., Delavault, P., Beau, J.-M., and Rameau, C. (2012). Structure-activity relationship studies of strigolactone-related molecules for branching inhibition in garden pea: molecule design for shoot branching. Plant Physiol. **159**: 1524–44.

Brunoud, G., Wells, D.M., Oliva, M., Larrieu, A., Mirabet, V., Burrow, A.H., Beeckman, T., Kepinski, S., Traas, J., Bennett, M.J., and Vernoux, T. (2012). A novel sensor to map auxin response and distribution at high spatio-temporal resolution. Nature **482**: 103–6.

Budič, M., Sabotič, J., Meglič, V., Kos, J., and Kidrič, M. (2013). Characterization of two novel subtilases from common bean (Phaseolus vulgaris L.) and their responses to drought. Plant Physiol. Biochem. **62**: 79–87.

Burlat, V., Kwon, M., Davin, L.B., and Lewis, N.G. (2001). Dirigent proteins and dirigent sites in lignifying tissues. Phytochemistry **57**: 883–897.

Chang, M. and Lynn, D.G. (1986). The haustorium and the chemistry of host recognition in parasitic angiosperms. J. Chem. Ecol. **12**: 561–579.

Cheng, Y., Dai, X., and Zhao, Y. (2006). Auxin biosynthesis by the YUCCA flavin monooxygenases controls the formation of floral organs and vascular tissues in Arabidopsis. Genes Dev. **20**: 1790–9.

Chichkova, N. V, Shaw, J., Galiullina, R.A., Drury, G.E., Tuzhikov, A.I., Kim, S.H., Kalkum, M., Hong, T.B., Gorshkova, E.N., Torrance, L., Vartapetian, A.B., and Taliansky, M. (2010). Phytaspase , a relocalisable cell death promoting. EMBO J. **29**: 1149–1161.

Cissoko, M., Boisnard, A., Rodenburg, J., Press, M.C., and Scholes, J.D. (2011). New Rice for Africa (NERICA) cultivars exhibit different levels of post-attachment resistance against the parasitic weeds Striga hermonthica and Striga asiatica. New Phytol. **192**: 952–63.

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics **21**: 3674–6.

Cosgrove, D.J. (2005). Growth of the plant cell wall. Nat. Rev. Mol. Cell Biol. **6**: 850–61.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.-R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiol. **139**: 5–17.

Dai, M., Hu, Y., Ma, Q., Zhao, Y., and Zhou, D.-X. (2008). Functional analysis of rice HOMEOBOX4 (Oshox4) gene reveals a negative function in gibberellin responses. Plant Mol. Biol. **66**: 289–301.

Das, M.K., Sharma, R.S., and Mishra, V. (2011). A novel cationic peroxidase (VanPrx) from a hemi-parasitic plant (Viscum angulatum) of Western Ghats (India): Purification, characterization and kinetic properties. J. Mol. Catal. B Enzym. **71**: 63–70.

Delaux, P.-M., Xie, X., Timme, R.E., Puech-Pages, V., Dunand, C., Lecompte, E., Delwiche, C.F., Yoneyama, K., Bécard, G., and Séjalon-Delmas, N. (2012). Origin of strigolactones in the green lineage. New Phytol. **195**: 857–71.

Dodson, G. and Wlodawer, A. (1998). Catalytic triads and their relatives. Trends Biochem. Sci. **23**: 347–352.

Duarte, J.M., Wall, P.K., Edger, P.P., Landherr, L.L., Ma, H., Pires, J.C., Leebens-Mack, J., and dePamphilis, C.W. (2010). Identification of shared single copy nuclear genes in Arabidopsis, Populus, Vitis and Oryza and their phylogenetic utility across various taxonomic levels. BMC Evol. Biol. **10**: 61.

Duroux, L. and Welinder, K.G. (2003). The Peroxidase Gene Family in Plants : A Phylogenetic Overview. 397–407.

Eda, H.U. and Ugimoto, Y.S. (2010). Vestitol as a Chemical Barrier against Intrusion of Parasitic Plant Striga hermonthica into Lotus japonicus Roots. **74**: 1662–1667.

Eizenberg, H., Hershenhorn, J., Ephrath Jhonathan H., and Kanampiu, F. (2013). Chemical Control. In Parasitic Orobanchaceae, D.M. Joel, G. Jonathan., and L.J. Musselman, eds (Springer: New York), pp. 415–432.

Ekblom, R. and Galindo, J. (2011). Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity (Edinb). **107**: 1–15.

Enalik, D.O.S., Eldin, E.R.I.C.Z., Own, B.R.M.C.C., Arbut, R.E.H., and Imon, P.H.S. (2012). U SING NEXT - GENERATION SEQUENCING APPROACHES TO ISOLATE SIMPLE SEQUENCE REPEAT (SSR) LOCI IN THE PLANT SCIENCES 1. **99**: 193–208.

Fernández-Aparicio, M., Huang, K., Wafula, E.K., Honaas, L.A., Wickett, N.J., Timko, M.P., Depamphilis, C.W., Yoder, J.I., and Westwood, J.H. (2013). Application of qRT-PCR and RNA-Seq analysis for the identification of housekeeping genes useful for normalization of gene expression values during Striga hermonthica development. Mol. Biol. Rep. **40**: 3395–407.

Fry, S.C., Smith, R.C., Renwick, K.F., Martin, D.J., Hodge, S.K., and Matthews, K.J. (1992). Xyloglucan endotransglycosylase, a new wall-loosening enzyme activity from plants. Biochem. J. **282 (Pt 3**: 821–8.

Fukui, K., Ito, S., and Asami, T. (2013). Selective mimics of strigolactone actions and their potential use for controlling damage caused by root parasitic weeds. Mol. Plant **6**: 88–99.

Gal-On, A., Naglis, A., Leibman, D., Ziadna, H., Kathiravan, K., Papayiannis, L., Holdengreber, V., Guenoune-Gelbert, D., Lapidot, M., and Aly, R. (2009). Broomrape Can Acquire Viruses from Its Hosts. Phytopathology **99**: 1321–1329.

Goldwasser, Y. and Rodenburg, J. (2013). Integrated Agronomic Management of Parasitic Weed Seed Banks. In Parasitic Orobanchaceae, D.M. Joel, Gressel Jonathan, and L.J. Musselman, eds (Springer: New York), pp. 393–413.

González-Verdejo, C.I., Barandiaran, X., Moreno, M.T., Cubero, J.I., and Di Pietro, A. (2006). A peroxidase gene expressed during early developmental stages of the parasitic plant Orobanche ramosa. J. Exp. Bot. **57**: 185–92.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. **29**: 644–52.

Von Groll, U. (2002). The Subtilisin-Like Serine Protease SDD1 Mediates Cell-to-Cell Signaling during Arabidopsis Stomatal Development. PLANT CELL ONLINE **14**: 1527–1539.

De Groote, H., Wangare, L., Kanampiu, F., Odendo, M., Diallo, A., Karaya, H., and Friesen, D. (2008). The potential of a herbicide resistant maize technology for Striga control in Africa. Agric. Syst. **97**: 83–94.

Grunewald, W., Karimi, M., Wieczorek, K., Van de Cappelle, E., Wischnitzki, E., Grundler, F., Inzé, D., Beeckman, T., and Gheysen, G. (2008). A role for AtWRKY23 in feeding site establishment of plant-parasitic nematodes. Plant Physiol. **148**: 358–68.

Gu, Y.-Q. (2002). Tomato Transcription Factors Pti4, Pti5, and Pti6 Activate Defense Responses When Expressed in Arabidopsis. PLANT CELL ONLINE **14**: 817–831.

Guan, X., Zhao, H., Xu, Y., and Wang, Y. (2011). Transient expression of glyoxal oxidase from the Chinese wild grape Vitis pseudoreticulata can suppress powdery mildew in a susceptible genotype. Protoplasma **248**: 415–23.

Hamiaux, C., Drummond, R.S.M., Janssen, B.J., Ledger, S.E., Cooney, J.M., Newcomb, R.D., and Snowden, K.C. (2012). DAD2 Is an a / b Hydrolase Likely to Be Involved in the Perception of the Plant Branching Hormone , Strigolactone. Curr. Biol. **22**: 2032–2036.

Haussmann, B.I.G., Hess, D.E., Omanya, G.O., Folkertsma, R.T., Reddy, B.V.S., Kayentao, M., Welz, H.G., and Geiger, H.H. (2004). Genomic regions influencing resistance to the parasitic weed Striga hermonthica in two recombinant inbred populations of sorghum. Theor. Appl. Genet. **109**: 1005–16.

Haussmann, B.I.G., Hess, D.E., Reddy, B.V.S., Welz, H.G., and Geiger, H.H. (2000). Analysis of resistance to Striga hermonthica in diallel crosses of sorghum. 33–40.

Heide-Jørgensen, H.S and KUIJT, J. (1995). The haustorium of the root parasite Triphysaria (Scrophulariaceae), with special reference to Xylem bridge ultrastructure. Am. J. Bot. **82**: 782–797.

Heim, M.A., Jakoby, M., Werber, M., Martin, C., Weisshaar, B., and Bailey, P.C. (2003). The basic helix-loop-helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. Mol. Biol. Evol. **20**: 735–47.

Hoegger, P.J., Kilaru, S., James, T.Y., Thacker, J.R., and Kües, U. (2006). Phylogenetic comparison and classification of laccase and related multicopper oxidase protein sequences. FEBS J. **273**: 2308–26.

Honaas, L.A., Wafula, E.K., Yang, Z., Der, J.P., Wickett, N.J., Altman, N.S., Taylor, C.G., Yoder, J.I., Timko, M.P., Westwood, J.H., and dePamphilis, C.W. (2013). Functional genomics of a generalist parasitic plant: laser microdissection of host-parasite interface reveals host-specific patterns of parasite gene expression. BMC Plant Biol. **13**: 9.

Huang, X. and Madan, a (1999). CAP3: A DNA sequence assembly program. Genome Res. **9**: 868–77.

Hwang, S.-G., Kim, D.S., and Jang, C.S. (2011). Comparative analysis of evolutionary dynamics of genes encoding leucine-rich repeat receptor-like kinase between rice and Arabidopsis. Genetica **139**: 1023–32.

Iseli, C., Jongeneel, C. V, and Bucher, P. (1999). ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. Proc. Int. Conf. Intell. Syst. Mol. Biol.: 138–48.

Ishida, J.K., Yoshida, S., Ito, M., Namba, S., and Shirasu, K. (2011). Agrobacterium rhizogenes-mediated transformation of the parasitic plant Phtheirospermum japonicum. PLoS One **6**: e25802.

Janssen, B.J. and Snowden, K.C. (2012). Strigolactone and karrikin signal perception: receptors, enzymes, or both? Front. Plant Sci. **3**: 296.

Jehle, A.K., Lipschis, M., Albert, M., Fallahzadeh-Mamaghani, V., Fürst, U., Mueller, K., and Felix, G. (2013). The receptor-like protein ReMAX of Arabidopsis detects the microbe-associated molecular pattern eMax from Xanthomonas. Plant Cell **25**: 2330–40.

Joel, D.M., Gressel, J., and Musselman, L.J. eds. (2013). Parasitic Orobanchaceae first. (Springer: New York).

Johnson, A.W., Gowada, G., Hassanali, A., Knox, J., Monaco, S., Razavi, Z., and Rosebery, G. (1981). The preparation of synthetic analogues of strigol. J. Chem. Soc. Perkin Trans. 1: 1734.

Jones, A.R., Kramer, E.M., Knox, K., Swarup, R., Bennett, M.J., Lazarus, C.M., Leyser, H.M.O., and Grierson, C.S. (2009). Auxin transport through non-hair cells sustains root-hair development. Nat. Cell Biol. **11**: 78–84.

Jones, M.P., Dingkuhn, M., Aluko/snm>, G.K., and Semon, M. (1997). Interspecific Oryza Sativa L. X O. Glaberrima Steud. progenies in upland rice improvement. Euphytica **94**: 237–246.

Jordá, L. and Vera, P. (2000). Local and systemic induction of two defense-related subtilisin-like protease promoters in transgenic Arabidopsis plants. Luciferin induction of PR gene expression. Plant Physiol. **124**: 1049–58.

Kapulnik, Y., Delaux, P.-M., Resnick, N., Mayzlish-Gati, E., Wininger, S., Bhattacharya, C., Séjalon-Delmas, N., Combier, J.-P., Bécard, G., Belausov, E., Beeckman, T., Dor, E., Hershenhorn, J., and Koltai, H. (2011). Strigolactones affect lateral root formation and root-hair elongation in Arabidopsis. Planta **233**: 209–16.

Kersten, P.J. and Kirk, T.K. (1987). Involvement of a new enzyme, glyoxal oxidase, in extracellular H2O2 production by Phanerochaete chrysosporium. J. Bacteriol. **169**: 2195–2201.

KGOSI, R.L., ZWANENBURG, B., MWAKABOKO, A.S., and MURDOCH, A.J. (2012). Strigolactone analogues induce suicidal seed germination of Striga spp. in soil. Weed Res. **52**: 197–203.

Kim, D., Kocz, R., Boone, L., Keyes, W.J., and Lynn, D.G. On becoming a parasite: evaluating the role of wall oxidases in parasitic plant development Dongjin Kim, Remigiusz Kocz, Laural Boone, W John Keyes and David G Lynn. 103–117.

Kistner, C., Winzer, T., Pitzschke, A., Mulder, L., Sato, S., Kaneko, T., Tabata, S., Sandal, N., Stougaard, J., Webb, K.J., Szczyglowski, K., and Parniske, M. (2005). Seven Lotus japonicus genes required for transcriptional reprogramming of the root during fungal and bacterial symbiosis. Plant Cell **17**: 2217–29.

Koizumi, K., Hayashi, T., and Gallagher, K.L. (2012). SCARECROW reinforces SHORT-ROOT signaling and inhibits periclinal cell divisions in the ground tissue by maintaining SHR at high levels in the endodermis. 1573–1577.

Koren, S., Schatz, M.C., Walenz, B.P., Martin, J., Howard, J.T., Ganapathy, G., Wang, Z., Rasko, D.A., McCombie, W.R., Jarvis, E.D., and Adam M Phillippy (2012). Hybrid error correction and de novo assembly of single-molecule sequencing reads. Nat. Biotechnol. **30**: 693–700.

Kottapalli, K.R., Kottapalli, P., Agrawal, G.K., Kikuchi, S., and Rakwal, R. (2007). Recessive bacterial leaf blight resistance in rice: Complexity, challenges and strategy. Biochem. Biophys. Res. Commun. **355**: 295–301.

Langfelder, P., Mischel, P.S., and Horvath, S. (2013). When is hub gene selection better than standard meta-analysis? PLoS One **8**: e61505.

Langston M. and English TJ. (1990). witchweed research and control in the United States. In Vegetative control of witchweed and herbicide evaluation of techniques, P. Sand, R. Eplee, and R. Westbrooks, eds (WSSA: Champaign, IL), pp. 108–113.

Lati, R., Aly, R., Eizenberg, H., and Lande, T. (2013). First Report of the Parasitic Plant Phelipanche aegyptiaca Infecting Kenaf in Israel. Plant Dis. **97**: 695–695.

Leuthner, B., Aichinger, C., Oehmen, E., Koopmann, E., Müller, O., Müller, P., Kahmann, R., Bölker, M., and Schreier, P.H. (2005). A H2O2-producing glyoxal oxidase is required for filamentous growth and pathogenicity in Ustilago maydis. Mol. Genet. Genomics **272**: 639–50.

Lewis, J.D., Wu, R., Guttman, D.S., and Desveaux, D. (2010). Allele-specific virulence attenuation of the Pseudomonas syringae HopZ1a type III effector via the Arabidopsis ZAR1 resistance protein. PLoS Genet. **6**: e1000894.

Li, C., Wang, Y., Huang, X., Li, J., Wang, H., and Li, J. (2013). De novo assembly and characterization of fruit transcriptome in Litchi chinensis Sonn and analysis of differentially regulated genes in fruit in response to shading. BMC Genomics **14**: 552.

Li, J. and Timko, M.P. (2009a). Gene-for-gene resistance in Striga-cowpea associations. Science **325**: 1094.

Li, J. and Timko, M.P. (2009b). Gene-for-gene resistance in Striga-cowpea associations. Science **325**: 1094.

Li, Y., Feng, D., Zhang, D., Su, J., Zhang, Y., Li, Z., Mu, P., Liu, B., Wang, H., and Wang, J. (2012). Rice MAPK phosphatase IBR5 negatively regulates drought stress tolerance in transgenic Nicotiana tabacum. Plant Sci. **188-189**: 10–8.

Liu, H., Wang, X., Zhang, H., Yang, Y., Ge, X., and Song, F. (2008). A rice serine carboxypeptidase-like gene OsBISCPL1 is involved in regulation of defense responses against biotic and oxidative stress. Gene **420**: 57–65.

Liu, J., Novero, M., Charnikhova, T., Ferrandino, A., Schubert, A., Ruyter-Spira, C., Bonfante, P., Lovisolo, C., Bouwmeester, H.J., and Cardinale, F. (2013). Carotenoid cleavage dioxygenase 7 modulates plant growth, reproduction, senescence, and determinate nodulation in the model legume Lotus japonicus. J. Exp. Bot. **64**: 1967–81.

Lopez-Gomez, M., Sandal, N., Stougaard, J., and Boller, T. (2012). Interplay of flg22-induced defence responses and nodulation in Lotus japonicus. J. Exp. Bot. **63**: 393–401.

Losner-Goshen, D. (1998). Pectolytic Activity by the Haustorium of the Parasitic PlantOrobancheL. (Orobanchaceae) in Host Roots. Ann. Bot. **81**: 319–326.

Maekawa, T., Kusakabe, M., Shimoda, Y., Sato, S., Tabata, S., Murooka, Y., and Hayashi, M. (2008). Polyubiquitin promoter-based binary vectors for overexpression and gene silencing in Lotus japonicus. Mol. Plant. Microbe. Interact. **21**: 375–82.

Mashiguchi, K., Tanaka, K., Sakai, T., Sugawara, S., Kawaide, H., and Natsume, M. (2011). The main auxin biosynthesis pathway in Arabidopsis. 1–6.

Matvienko, M. (2001). Transcriptional Responses in the Hemiparasitic Plant Triphysaria versicolor to Host Plant Signals. PLANT Physiol. **127**: 272–282.

Matvienko, M., Wojtowicz, A., Wrobel, R., Jamison, D., Goldwasser, Y., and Yoder, J.I. (2001). Quinone oxidoreductase message levels are differentially regulated in parasitic and non-parasitic plants exposed to allelopathic quinones. Plant J. **25**: 375–387.

McIntyre, L.M., Lopiano, K.K., Morse, A.M., Amin, V., Oberg, A.L., Young, L.J., and Nuzhdin, S. V (2011). RNA-seq: technical variability and sampling. BMC Genomics **12**: 293.

Meijer, A.H., Scarpella, E., Dijk, E.L., Qin, L., Taal, A.J.C., Rueb, S., Harrington, S.E., McCouch, S.R., Schilperoort, R.A., and Hoge, J.H.C. (1997). Transcriptional repression by Oshox1, a novel homeodomain leucine zipper protein from rice. Plant J. **11**: 263–276.

Mohamed, A.H. and Ejeta, G. (1998). Moisture content and dormancy in Striga asiatica seeds. **38**: 257–265.

Moreno, M.T. and Rubiales, D. (2008). Host plant resistance against broomrapes (Orobanche spp .): defence reactions and mechanisms of resistance. **152**: 131–141.

Moreno-Risueno, M.A., Van Norman, J.M., Moreno, A., Zhang, J., Ahnert, S.E., and Benfey, P.N. (2010). Oscillating gene expression determines competence for periodic Arabidopsis root branching. Science **329**: 1306–11.

Mower, J.P., Stefanović, S., Hao, W., Gummow, J.S., Jain, K., Ahmed, D., and Palmer, J.D. (2010). Horizontal acquisition of multiple mitochondrial genes from a parasitic plant followed by gene conversion with host mitochondrial genes. BMC Biol. **8**: 150.

MUSSELMAN, L.J. and DICKISON, W.C. (1975). The structure and development of the haustorium in parasitic Scrophulariaceae*. Bot. J. Linn. Soc. **70**: 183–212.

Nakagawa, M., Ueyama, M., Tsuruta, H., Uno, T., Kanamaru, K., Mikami, B., and Yamagata, H. (2010). Functional analysis of the cucumisin propeptide as a potent inhibitor of its mature enzyme. J. Biol. Chem. **285**: 29797–807.

Nakamura, S., Nakao, A., Kawamukai, M., Kimura, T., Ishiguro, S., and Nakagawa, T. (2009). Development of Gateway binary vectors, R4L1pGWBs, for promoter analysis in higher plants. Biosci. Biotechnol. Biochem. **73**: 2556–9.

Napoli, C.A. and Ruehle, J. (1996). New Mutations Affecting Meristem Growth and Potential in Petunia hybridaVilm. J. Hered. **87**: 371–377.

Neuteboom, L.W. (1999). A Novel Subtilisin-like Protease Gene from Arabidopsis thaliana is Expressed at Sites of Lateral Root Emergence. DNA Res. **6**: 13–19.

Nishitani, K. and Tominaga, R. (1992). Endo-xyloglucan transferase, a novel class of glycosyltransferase that catalyzes transfer of a segment of xyloglucan molecule to another xyloglucan molecule. J. Biol. Chem. **267**: 21058–21064.

Nun, N.B. and Mayer, A.M. (1993). Preconditioning and germination of Orobanche seeds: Respiration and protein synthesis. Phytochemistry **34**: 39–45.

Oláh, B., Brière, C., Bécard, G., Dénarié, J., and Gough, C. (2005). Nod factors and a diffusible factor from arbuscular mycorrhizal fungi stimulate lateral root formation in Medicago truncatula via the DMI1/DMI2 signalling pathway. Plant J. **44**: 195–207.

Oshima, Y., Mitsuda, N., Nakata, M., Nakagawa, T., Nagaya, S., Kato, K., and Ohme-Takagi, M. (2011). Novel vector systems to accelerate functional analysis of transcription factors using chimeric repressor gene-silencing technology (CRES-T). Plant Biotechnol. **28**: 201–210.

Ottmann, C., Rose, R., Huttenlocher, F., Cedzich, A., Hauske, P., Kaiser, M., Huber, R., and Schaller, A. (2009). Structural basis for Ca2+-independence and activation by homodimerization of tomato subtilase 3. Proc. Natl. Acad. Sci. U. S. A. **106**: 17223–8.

Pandey, S.P., Roccaro, M., Schön, M., Logemann, E., and Somssich, I.E. (2010). Transcriptional reprogramming regulated by WRKY18 and WRKY40 facilitates powdery mildew infection of Arabidopsis. Plant J. **64**: 912–23.

Park, J.M. (2001). Overexpression of the Tobacco Tsi1 Gene Encoding an EREBP/AP2-Type Transcription Factor Enhances Resistance against Pathogen Attack and Osmotic Stress in Tobacco. PLANT CELL ONLINE **13**: 1035–1046.

Park, J.-M., Manen, J.-F., and Schneeweiss, G.M. (2007). Horizontal gene transfer of a plastid gene in the non-photosynthetic flowering plants Orobanche and Phelipanche (Orobanchaceae). Mol. Phylogenet. Evol. **43**: 974–985.

Parker, C. (2009). Observations on the current status of Orobanche and Striga problems worldwide. Pest Manag. Sci. **65**: 453–9.

Parker, C. (2013). The Parasitic Weeds of the Orobanchaceae. In Parasitic Orobanchaceae, D.M. Joel, J. Gressel, and L.J. Musselman, eds (Springer: New York), pp. 313–344.

Passardi, F., Penel, C., and Dunand, C. (2004). Performing the paradoxical: how plant peroxidases modify the cell wall. Trends Plant Sci. **9**: 534–40.

Pearce, G., Yamaguchi, Y., Barona, G., and Ryan, C.A. (2010). A subtilisin-like protein from soybean contains an embedded, cryptic signal that activates defense-related genes. Proc. Natl. Acad. Sci. U. S. A. **107**: 14921–5.

Pennisi, E. (2010). Armed and Dangerous. Science (80-.). **327**: 804–805.

Prandi, C. and Franca, J.M. (2011). New Potent Fluorescent Analogues of Strigolactones : Synthesis and Biological Activity in Parasitic Weed Germination and Hyphal Branching in AM Fungi Plant hormones. 7–12.

Proust, H., Hoffmann, B., Xie, X., Yoneyama, K., Schaefer, D.G., Yoneyama, K., Nogué, F., and Rameau, C. (2011). Strigolactones regulate protonema branching and act as a quorum sensing-like signal in the moss Physcomitrella patens. Development **138**: 1531–9.

Ralph, S., Park, J.-Y., Bohlmann, J., and Mansfield, S.D. (2006). Dirigent proteins in conifer defense: gene discovery, phylogeny, and differential wound- and insect-induced

expression of a family of DIR and DIR-like genes in spruce (Picea spp.). Plant Mol. Biol. **60**: 21–40.

Ramírez, V., López, A., Mauch-Mani, B., Gil, M.J., and Vera, P. (2013). An extracellular subtilase switch for immune priming in Arabidopsis. PLoS Pathog. **9**: e1003445.

Ransom, J., Kanampiu, F., Gressel, J., De Groote, H., Burnet, M., and Odhiambo, G. (2012). Herbicide Applied to Imidazolinone Resistant-Maize Seed as a Striga Control Option for Small-Scale African Farmers. Weed Sci. **60**: 283–289.

Rashotte, A.M. and Goertzen, L.R. (2010). The CRF domain defines cytokinin response factor proteins in plants. BMC Plant Biol. **10**: 74.

Rashotte, A.M., Mason, M.G., Hutchison, C.E., Ferreira, F.J., Schaller, G.E., and Kieber, J.J. (2006). A subset of Arabidopsis AP2 transcription factors mediates cytokinin responses in concert with a two-component pathway. Proc. Natl. Acad. Sci. U. S. A. **103**: 11081–5.

Rautengarten, C., Steinhauser, D., Büssis, D., Stintzi, A., Schaller, A., Kopka, J., and Altmann, T. (2005). Inferring hypotheses on functional relationships of genes: Analysis of the Arabidopsis thaliana subtilase gene family. PLoS Comput. Biol. **1**: e40.

Rawlings, N.D., Barrett, A.J., and Bateman, A. (2012). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. **40**: D343–50.

Rehker, J., Lachnit, M., and Kaldenhoff, R. (2012). Molecular convergence of the parasitic plant species Cuscuta reflexa and Phelipanche aegyptiaca. Planta **236**: 557–66.

Reich, M., Liefeld, T., Gould, J., Lerner, J., Tamayo, P., and Mesirov, J.P. (2006). GenePattern 2.0. Nat. Genet. **38**: 500–1.

Riopel, J.L. and Musselman, L.J. (1979). Experimental initiation of haustoria in agalinis purpura (scrophulariaceae). Am. J. Bot. **66**: 570–575.

Riopel, J.L.. and Timko, M.P. (1995). Haustorial initiation and differentiation. In Parasitic Plant, M.C. Press and J.D. Graves, eds (Chapman & Hall: London), pp. 40–79.

Román, B., Torres, A.M., Rubiales, D., Cubero, J.I., and Satovic, Z. (2002). Mapping of quantitative trait loci controlling broomrape (Orobanche crenata Forsk.) resistance in faba bean (Vicia faba L.). Genome **45**: 1057–1063.

Rubiales, D. and Fernández-Aparicio, M. (2011). Innovations in parasitic weeds management in legume crops. A review. Agron. Sustain. Dev. **32**: 433–449.

Rubiales, D. and Heide-Jørgensen, H.. (2011). Parasitic plants. In Encyclopedia of Life Sciences (John Wiley & Sons, Ltd: Chichester, UK).

Sabatini, S., Heidstra, R., Wildwater, M., and Scheres, B. (2003). SCARECROW is involved in positioning the stem cell niche in the Arabidopsis root meristem SCARECROW is involved in positioning the stem cell niche in the Arabidopsis root meristem. 354–358.

De Saint Germain, A., Bonhomme, S., Boyer, F.-D., and Rameau, C. (2013). Novel insights into strigolactone distribution and signalling. Curr. Opin. Plant Biol. **16**: 583–9.

De Saint Germain, A., Ligerot, Y., Dun, E.A., Pillot, J.-P., Ross, J.J., Beveridge, C.A., and Rameau, C. (2013). Strigolactones stimulate internode elongation independently of gibberellins. Plant Physiol. **163**: 1012–1025.

Sampedro, J. and Cosgrove, D.J. (2005). The expansin superfamily. Genome Biol. **6**: 242.

Schaller, A., Stintzi, A., and Graff, L. (2012). Subtilases - versatile tools for protein turnover, plant development, and interactions with the environment. Physiol. Plant. **145**: 52–66.

Schweikert, C., Liszkay, A., and Schopfer, P. (2000). Scission of polysaccharides by peroxidase-generated hydroxyl radicals. Phytochemistry **53**: 565–570.

Scurfield, G. (1973). Reaction Wood: Its Structure and Function: Lignification may generate the force active in restoring the trunks of leaning trees to the vertical. Science **179**: 647–55.

Shi, C.-Y., Yang, H., Wei, C.-L., Yu, O., Zhang, Z.-Z., Jiang, C.-J., Sun, J., Li, Y.-Y., Chen, Q., Xia, T., and Wan, X.-C. (2011). Deep sequencing of the Camellia sinensis transcriptome revealed candidate genes for major metabolic pathways of tea-specific compounds. BMC Genomics **12**: 131.

Smaling, E.M.A., Stein, A., and Sloot, P.H.M. (1991). A statistical analysis of the influence of Striga hermonthica on maize yields in fertilizer trials in Southwestern Kenya. Plant Soil **138**: 1–8.

Smith, C.E., Dudley, M.W., and Lynn, D.G. (1990). Vegetative/Parasitic Transition: Control and Plasticity in Striga Development. PLANT Physiol. **93**: 208–215.

Spallek, T., Mutuku, J.M., and Shirasu, K. (2013). The genus Striga: A witch profile. Mol. Plant Pathol.

Stanke, M., Steinkamp, R., Waack, S., and Morgenstern, B. (2004). AUGUSTUS: a web server for gene finding in eukaryotes. Nucleic Acids Res. **32**: W309–12.

Sugimoto, Y., Wigchert, S.C.M., Thuring, J.W.J.F., and Zwanenburg, B. (1998). Synthesis of All Eight Stereoisomers of the Germination Stimulant Sorgolactone. J. Org. Chem. **63**: 1259–1267.

Suh, S.J., Bowman, B.C., Jeong, N., Yang, K., Kastl, C., Tolin, S.A., Maroof, M.A.S., and Jeong, S.-C. (2011). The Rsv3 Locus Conferring Resistance to Soybean Mosaic Virus is Associated with a Cluster of Coiled-Coil Nucleotide-Binding Leucine-Rich Repeat Genes. Plant Genome J. **4**: 55.

Swarbrick, P.J., Huang, K., Liu, G., Slate, J., Press, M.C., and Scholes, J.D. (2008). Global patterns of gene expression in rice cultivars undergoing a susceptible or resistant interaction with the parasitic plant Striga hermonthica. New Phytol. **179**: 515–529.

Takeda, N., Sato, S., Asamizu, E., Tabata, S., and Parniske, M. (2009). Apoplastic plant subtilases support arbuscular mycorrhiza development in Lotus japonicus. Plant J. **58**: 766–77.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28**: 2731–9.

Tanaka, H., Onouchi, H., Kondo, M., Hara-Nishimura, I., Nishimura, M., Machida, C., and Machida, Y. (2001). A subtilisin-like serine protease is required for epidermal surface formation in Arabidopsis embryos and juvenile plants. Development **128**: 4681–4689.

Tobeña-Santamaria, R., Bliek, M., Ljung, K., Sandberg, G., Mol, J.N.M., Souer, E., and Koes, R. (2002). FLOOZY of petunia is a flavin mono-oxygenase-like protein required for the specification of leaf and flower architecture. Genes Dev. **16**: 753–63.

Tomilov, A.A., Tomilova, N.B., Abdallah, I., and Yoder, J.I. (2005). Localized hormone fluxes and early haustorium development in the hemiparasitic plant Triphysaria versicolor. Plant Physiol. **138**: 1469–80.

Tomilov, A.A., Tomilova, N.B., Wroblewski, T., Michelmore, R., and Yoder, J.I. (2008). Trans-specific gene silencing between host and parasitic plants. Plant J. **56**: 389–397.

Tominaga, M., Yokota, E., Vidali, L., Sonobe, S., Hepler, P.K., and Shimmen, T. (2000). The role of plant villin in the organization of the actin cytoskeleton, cytoplasmic streaming and the architecture of the transvacuolar strand in root hair cells of Hydrocharis. Planta **210**: 836–843.

Valderrama, M.R., Roman, B., Satovic, Z., Rubiales, D., Cubero, J.I., and Torres, A.M. (2004). Locating quantitative trait loci associated with Orobanche crenata resistance in pea. Weed Res. **44**: 323–328.

Vance, C.P., Kirk, T.K., and Sherwood, R.T. (1980). Lignification as a Mechanism of Disease Resistance. Annu. Rev. Phytopathol. **18**: 259–288.

Vartapetian, a B., Tuzhikov, a I., Chichkova, N. V, Taliansky, M., and Wolpert, T.J. (2011). A plant alternative to animal caspases: subtilisin-like proteases. Cell Death Differ. **18**: 1289–97.

Vitha, S., Benes, K., Phillips, J.P., and Gartland, K.M. (1995). Histochemical GUS analysis. Methods Mol. Biol. **44**: 185–93.

Wall, P.K., Leebens-Mack, J., Müller, K.F., Field, D., Altman, N.S., and dePamphilis, C.W. (2008). PlantTribes: a gene and gene family resource for comparative genomics in plants. Nucleic Acids Res. **36**: D970–6.

Wang, L., Si, Y., Dedow, L.K., Shao, Y., Liu, P., and Brutnell, T.P. (2011). A low-cost library construction protocol and data analysis pipeline for Illumina-based strand-specific multiplex RNA-seq. PLoS One **6**: e26426.

Waruru, M. (2013). Deadly Striga weed spreading across Eastern Africa.

Waters, M.T., Nelson, D.C., Scaffidi, A., Flematti, G.R., Sun, Y.K., Dixon, K.W., and Smith, S.M. (2012). Specialisation within the DWARF14 protein family confers distinct responses to karrikins and strigolactones in Arabidopsis. Development **139**: 1285–95.

Westwood, J.H., Roney, J.K., Khatibi, P.A., and Stromberg, V.K. (2009). RNA translocation between parasitic plants and their hosts. 533–539.

Westwood, J.H., Yoder, J.I., Timko, M.P., and dePamphilis, C.W. (2010). The evolution of parasitism in plants. Trends Plant Sci. **15**: 227–35.

Wickett, N.J., Honaas, L.A., Wafula, E.K., Das, M., Huang, K., Wu, B., Landherr, L., Timko, M.P., Yoder, J., Westwood, J.H., and dePamphilis, C.W. (2011). Transcriptomes of the parasitic plant family Orobanchaceae reveal surprising conservation of chlorophyll synthesis. Curr. Biol. **21**: 2098–104.

Wm. Vance Baird and James L. Riopel (1984). Experimental studies of haustorium initiation and early development in Agalinis purpurea (L.0 Raf. (Scrophulariaceae). Bot. Soc. Am. **71**: 803–814.

Xi, Z., Bradley, R.K., Wurdack, K.J., Wong, K., Sugumaran, M., Bomblies, K., Rest, J.S., and Davis, C.C. (2012). Horizontal transfer of expressed genes in a parasitic flowering plant. BMC Genomics **13**: 227.

Xie, X., Yoneyama, K., and Yoneyama, K. (2010). The strigolactone story. Annu. Rev. Phytopathol. **48**: 93–117.

Xu, X., Chen, C., Fan, B., and Chen, Z. (2006). Physical and functional interactions between pathogen-induced Arabidopsis WRKY18, WRKY40, and WRKY60 transcription factors. Plant Cell **18**: 1310–26.

Yalpani, N. (1991). Salicylic Acid Is a Systemic Signal and an Inducer of Pathogenesis-Related Proteins in Virus-Infected Tobacco. PLANT CELL ONLINE **3**: 809–818.

Yamamoto, Y., Kamiya, N., Morinaka, Y., Matsuoka, M., and Sazuka, T. (2007). Auxin Biosynthesis by the YUCCA Genes in Rice 1 [W][OA]. **143**: 1362–1371.

Yoder, J.I., Gunathilake, P., Wu, B., Tomilova, N., and Tomilov, A.A. (2009). Engineering host resistance against parasitic weeds with RNA interference. Pest Manag. Sci. **65**: 460–466.

Yoneyama, K., Takeuchi, Y., and Sekimoto, H. (2007). Phosphorus deficiency in red clover promotes exudation of orobanchol, the signal for mycorrhizal symbionts and germination stimulant for root parasites. Planta **225**: 1031–8.

Yoneyama, K., Xie, X., Kusumoto, D., Sekimoto, H., Sugimoto, Y., Takeuchi, Y., and Yoneyama, K. (2007). Nitrogen deficiency as well as phosphorus deficiency in sorghum promotes the production and exudation of 5-deoxystrigol, the host recognition signal for arbuscular mycorrhizal fungi and root parasites. Planta **227**: 125–32.

Yoshida, S., Ishida, J.K., Kamal, N.M., Ali, A.M., Namba, S., and Shirasu, K. (2010). A full-length enriched cDNA library and expressed sequence tag analysis of the parasitic weed, Striga hermonthica. BMC Plant Biol. **10**: 55.

Yoshida, S., Maruyama, S., Nozaki, H., and Shirasu, K. (2010). Horizontal Gene Transfer by the Parasitic Plant Striga hermonthica . Sci. **328** : 1128.

Yoshida, S. and Shirasu, K. (2009). Multiple layers of incompatibility to the parasitic witchweed, Striga hermonthica. New Phytol. **183**: 180–9.

Yoshimoto, K., Hanaoka, H., Sato, S., Kato, T., Tabata, S., Noda, T., and Ohsumi, Y. (2004). Processing of ATG8s, ubiquitin-like proteins, and their deconjugation by ATG4s are essential for plant autophagy. Plant Cell **16**: 2967–83.

Yue, J., Hu, X., Sun, H., Yang, Y., and Huang, J. (2012). Widespread impact of horizontal gene transfer on plant colonization of land. Nat. Commun. **3**: 1152.

Zhao, C., Craig, J.C., Petzold, H.E., Dickerman, A.W., and Beers, E.P. (2005). The xylem and phloem transcriptomes from secondary tissues of the Arabidopsis root-hypocotyl. Plant Physiol. **138**: 803–18.

Zhao, L.-H., Zhou, X.E., Wu, Z.-S., Yi, W., Xu, Y., Li, S., Xu, T.-H., Liu, Y., Chen, R.-Z., Kovach, A., Kang, Y., Hou, L., He, Y., Xie, C., Song, W., Zhong, D., Xu, Y., Wang, Y., Li, J., Zhang, C., et al. (2013). Crystal structures of two phytohormone signal-transducing α/β hydrolases: karrikin-signaling KAI2 and strigolactone-signaling DWARF14. Cell Res. **23**: 436–9.

Zhao, Y., Christensen, S.K., Fankhauser, C., Cashman, J.R., Cohen, J.D., Weigel, D., and Chory, J. (2001). A role for flavin monooxygenase-like enzymes in auxin biosynthesis. Science **291**: 306–9.

Zheng, Z., Qamar, S.A., Chen, Z., and Mengiste, T. (2006). Arabidopsis WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. Plant J. **48**: 592–605.

Zhou, J., Tang, X., and Martin, G.B. (1997). The Pto kinase conferring resistance to tomato bacterial speck disease interacts with proteins that bind a cis-element of pathogenesis-related genes. EMBO J. **16**: 3207–18.

Zwanenburg, B., Malik, H., Mwakaboko, A.S., and Rutjes, F.P.J.T. (2011). New Strigolactone Analogues , Design , Synthesis and Bioactivity Parasitic Weeds : Beautiful flowers but a real pest.