Genes Expressed during the Biotrophic Phase of the Rust Fungus *Uromyces fabae*

Dissertation

zur Erlangung des akademischen Grades des Doktors der Naturwissenschaften (Dr. rer. nat.)

Vorgelegt von

Uta Hempel

Konstanz, im September 2005

Universität Konstanz
Fachbereich Biologie
Lehrstuhl für Phytopathologie

Tag der mündlichen Prüfung: 12.01.2006
Referent: Prof. Dr. K. W. Mendgen
Referentin: Prof. Dr. I. Adamska

For Markus

Acknowledgements

This thesis would not have been possible without the help of many brilliant and immensely supportive people.

I would especially like to thank my doctoral advisor Prof. Dr. Kurt Mendgen for providing the opportunity to work in the Department of Phytopathology at the University of Konstanz, for suggesting the topic of this thesis, for the excellent scientific equipment made available, the constant academic support and the cooperative atmosphere experienced within his department. I would also like to thank Prof. Dr. Iwona Adamska that she kindly agreed to contribute her scientific expertise and valuable time to act as second reviewer.

Special thanks go to Prof. Dr. Matthias Hahn for having provided invaluable spadework, supplemented by many fruitful discussions, which together have provided a solid basis for this thesis. I am also deeply indebted to Dr. Ralf T. Voegele for his guidance and patient assistance regarding the analysis of the studied *PIG*s. His extraordinary competence in respect to proteins, his professionally organized laboratory, and his outstanding knowledge of scientific methods have greatly contributed to the success of this thesis.

In addition, I am immensely grateful to all my colleagues in the scientific community in- and outside the University of Konstanz who have contributed their time, effort, and expertise to the success of this thesis. These include especially: Dr.-Ing. Ulrike Damm for the help with classification of phytopathogenic fungi and the many discussions about life in the lab and life in general; Dr. Christiane Runge-Froböse for her comprehensive knowledge of PCR methods and sequencing, which she kindly shared with me; Dr. Christine Struck PD, Dr. Stefan Wirsel, Dr. Nicole Jenczmionka, my fellow doctorate students Dipl. Biol. Karin Neubert, Dipl. Biol. Michael Ernst, Dipl. Biol. Ariane C. Kemen and Dipl. Biol. Eric Kemen for providing a lively and invigorating forum for mutual discussion of our respective projects; Carolyn M. Komar, PhD for an animal physiologist's review of this thesis and many helpful suggestions; Christine Giele, Dipl. Biol. Sebastian Bunk, Dipl. Biol. Julia Heinzmann and Anette Schmid, for their valuable assistance with my lab work; Dipl.-Ing. (FH) Heinz Vahlenkamp, who carried out the necessary immuno-fluorescent microscopy and advised on all issued related to microscopy; the experienced and immensely supportive staff of the University of Konstanz Botanical Garden, especially Ulrike Peters and Otmar Ficht, who provided the plants used for my experiments and assisted in all things related to plant care and gardening techniques; and finally, the dedicated and professional staff at the University of Konstanz Animal Research Facility (TFA) for tending and immunizing the animals involved in this research.

My heartfelt thanks also go to Shirley A. McDonald-Link, BA (Hons.) for taking the time and trouble to review sections of this thesis and eradicate improper uses of her native tongue. These acknowledgements would not be complete without mention of my parents, who have always supported me in every thinkable way. I would also like to use this opportunity to thank my family and friends, in particular Dipl. Päd. Barbara Ocklenburg for constantly motivation to keep going.

I thank my son Paul for starting to sleep five hours per night at the tender age of four weeks and thus giving me the time to give this thesis the finishing touches. Last but not least, I am deeply indebted to my husband Dipl.-Ing. Markus Teltscher, MBA, for his unfailing support, night –long discussions, an engineer's view of biological science, and all the work he invested in proofreading the English texts in this thesis. Thank you, Markus, without you I never would have managed!

1 Summary

1.1 English Summary

The subject of this thesis is the study of genes expressed during the biotrophic phase of the rust fungus *Uromyces fabae*, an obligate biotrophic pathogen of *Vicia faba* (broad bean).

As a first step, a previously initiated partial Expressed Sequence Tag (EST) sequencing project was completed and the results were analyzed. The aim of this project was to sequence at least 1'000 ESTs and to compare them to publicly available sequences. 58% of the 1'000 plus sequences analyzed using the BLASTX algorithm showed no significant homology to genes published in the NCBI database. The remaining 42% could be classified by functional groups, whereby a surprising number of ESTs bearing similarities to viral sequences were also found.

In a second step, three previously identified "*in Planta* Induced Genes" (*PIG*s) were to be characterized in more detail. This entailed: the internet-based analysis of unusual sequence characteristics; the determination of the number of gene copies; the localization of gene product in infected plant tissue; the verification of *in planta* gene expression; and the verification of the secretion signal.

To this end, six promising *PIGs* from the above group of ESTs were selected based on two important common characteristics: not only had an internet-based analysis revealed them to possess a secretion signal, but they also displayed no or very minor homology to publicly available sequences at the NCBI database.

The number of gene copies for each of the six *PIG*s was determined by Southern blotting, with the following results: *PIG14, PIG15* and *PIG23*: one copy each; *PIG5*: two copies*; PIG7* (*Rust Transferred Protein 1, RTP1*): presumably two copies with a single nucleotide polymorphism (SNP); *PIG9*: one to two copies.

Subsequently, an attempt was made to localize the gene products within the infected plant tissue by using antibodies against the PIG-proteins being studied. In order to produce antibodies, fusion proteins of the six *PIG*s were over-expressed in *E. coli*. During the subsequent protein purification process, the fusion proteins of *PIG9* and *PIG23* proved to be unstable. Thus, *PIG9* and *PIG23* were not studied any further. Using immuno-fluorescence microscopy, it was shown that the gene product of *PIG5* is located inside the haustoria and the gene products of *PIG14* and *PIG15* are located in the extrahaustorial matrix. Most intriguingly, the gene product of *PIG7* (*RTP1*) was found in the extrahaustorial matrix as well as in the plant(!) nucleus. This result correlated very well with the nuclear localization signals (NLS) found in PIG7p through an internet-based analysis.

Finally, the *in planta* expression of *PIG5*, *PIG7* (*RTP1*), *PIG14* and *PIG15* was verified by Northern blotting. The PIG-proteins were additionally expressed in transformed yeasts (*S. cerevisiae*), thus allowing the secretion signals to be positively verified for PIG7p and PIG14p using Western blotting.

The localization of PIG14p and PIG15p in the extrahaustorial matrix as well as the localization of PIG7p (RTP1p) in the extrahaustorial matrix, parts of the host cytoplasm and the plant nucleus provides supporting evidence that these proteins, especially RTP1p, play a (significant?) role in the interactions between *U. fabae* and *V. faba* during the biotrophic phase. Furthermore, the fact that PIG7p (RTP1p) can be found in the plant nucleus could be an indication that this protein is involved in the establishing of plant fungal interactions.

1.2 Zusammenfassung (German Summary)

Thema dieser Arbeit ist die Untersuchung von Genen des obligat biotrophen Rostpilzes *Uromyces fabae* (Erreger des Ackerbohnenrostes), die während der biotrophen Phase exprimiert werden.

Als Erstes wurde im Rahmen dieser Arbeit ein Teilsequenzierungsprojekt von *Uromyces fabae* zu Ende geführt und ausgewertet. Ziel war, mindesten 1000 Expressed Sequence Tags (ESTs) zu sequenzieren. Von den über 1000 Sequenzen, die mit dem BLASTX Algorithmus analysiert wurden, zeigten 58 % keine signifikante Ähnlichkeit zu schon bekannten Genen in der öffentlichen Datenbank des NCBI Servers. Die restlichen 42% der Sequenzen konnten funktionalen Gruppen zugeordnet werden, wobei überraschenderweise auch ESTs mit Ähnlichkeiten zu viralen Sequenzen gefunden wurden.

Des Weiteren sollten drei schon vorher als „*in planta* induced genes" (*PIG*s) beschriebene Gene näher charakterisiert werden. Dazu gehörte: eine Analyse bezüglich besonderer Sequenzcharakteristika, die Bestimmung der Anzahl von Genkopien, die Lokalisierung der Genprodukte in infiziertem Pflanzengewebe, die Verifizierung der *in planta* Gen Expression und die Verifizierung der theoretisch ermittelten Signalsequenz.

Zu diesem Zweck wurden sechs viel versprechende *PIG*s aus der oben genannten Gruppe von ESTs anhand von zwei Eigenschaften ausgewählt: Erstens zeigte eine Internet basierte Analyse, dass diese *PIG*s eine Signalsequenz haben; zweitens haben diese Gene auch keine (oder nur sehr geringe) Homologie zu öffentlich bekannten Sequenzen der NCBI Datenbank.

Die Anzahl der Genkopien wurde durch Southern Blotting mit folgenden Resultaten bestimmt: *PIG14*, *PIG15* und *PIG23*: jeweils eine Kopie; *PIG5*: zwei Kopien; *PIG7* (*Rust Transferred Protein 1, RTP1*): zwei Kopien vermutlich mit einem Einzel-Nucleotid-Polymorphismus (SNP); *PIG9*: ein oder zwei Kopien.

Als nächstes wurden die Genprodukte im infizierten Pflanzengewebe mithilfe von Antikörpern gegen die jeweiligen PIG-Proteine lokalisiert. Zur Herstellung der Antikörper wurden Fusionsproteine der sechs *PIG*s in *E. coli* überexprimiert. Während der folgenden Proteinreinigung erwiesen sich die Fusionsproteine von *PIG9* und *PIG23* als unstabil. Daher wurden *PIG9* und *PIG23* nicht mehr weiter untersucht. Durch den Einsatz der Immuno-Fluoreszenz Mikroskopie wurde gezeigt, dass das Genprodukt von *PIG5*, im Haustorium und die Genprodukte von *PIG14* und *PIG15* in der extra-haustoriellen Matrix lokalisiert sind. Besonders interessant ist, dass das Genprodukt von *PIG7* (*RTP1*) nicht nur in der extra-haustoriellen Matrix sondern auch im pflanzlichen(!) Zellkern gefunden wurde. Dieses Ergebnis korrespondiert sehr gut mit dem NLS („Nuclear Localizations Signal") welches in PIG7p bei einer Sequenzanalyse gefunden wurde.

Die *in planta* Induktion von *PIG5*, *PIG7* (*RTP1*), *PIG14* und *PIG15* wurde durch Northern blotting verifiziert. Die PIG-Proteine wurden außerdem in transformierten Hefen (*S. cerevisiae*) exprimiert, dadurch konnten die Sekretionssignale eindeutig für PIG7p and PIG14p mittels Western Blotting nachgewiesen werden.

Die Lokalisation von PIG14p und PIG15p in der extra-haustoriellen Matrix sowie die Lokalisation von PIG7p (RTP1p) sowohl in der extra-haustoriellen Matrix, in Teilen des Wirts- Zytoplasmas und im Zellkern der Pflanze unterstützt die These, dass die hier beschriebenen PIGps, insbesondere RTP1p, eine (signifikante?) Rolle in der Interaktion zwischen *U. fabae* und *V. faba* während der biotrophen Phase, spielt. Des Weiteren kann die Tatsache, dass PIG7p (RTP1p) im Zellkern der Pflanze gefunden wird, ein Hinweis darauf sein, dass dieses Protein an der Etablierung von der Pilz/Pflanzen Interaktion beteiligt ist.

2 Introduction

2.1 Plant Fungal Interactions

Fungi are eukaryotic, heterotrophic organisms, which reproduce by means of spores. The Kingdom of Fungi is divided into three divisions: the *Eumycota* (true fungi), *Chromista* and *Protozoa*. This thesis deals only with the *Eumycota*, which can be differentiated from other fungi by the fact that they form chitinous hyphal walls. The *Eumycota* are classified into four phyla: the *Chytridiomycota*, the *Zygomycota*, the *Ascomycota* and the *Basidiomycota* (Alexopoulos et al. 1996; Carlile et al. 2001). A phytopathogenic example for the *Chytridiomycota* is *Olpidium brassicae,* the agent for cabbage damping-off disease. Examples for the *Zygomycota* are *Mucor mucedo* and *Rhizopus stolonifer,* both of which cause fruit rot. Examples for the *Ascomycota* are *Saccharomyces cerevisiae*, commonly known as "Baker's Yeast", *Fusarium graminearum,* which causes head blight in barley and wheat, *Venturia inaequalis,* the causative agent of apple scab, and truffles (*Tuber spec.*). Some members of the *Zygomycota* and *Ascomycota* divisions form mycorrhizal associations, which are discussed below. Most mushrooms are classified within the *Basidiomycota* (Bruns et al. 1991; Bruns et al. 1993). *Uromyces fabae*, the agent for bean rust and the object of this study, belongs to the order *Uredinales* (rust fungi) within the *Basidiomycota* (Alexopoulos et al. 1996)**.**

The *Eumycota* includes numerous fungi that interact with plants. When studying these fungi, it becomes evident that they have evolved various different strategies to enable them to live off their hosts (Heath 1991). We differentiate between saprophytic, necrotrophic, symbiotic, and biotrophic fungi (Saikkonen et al. 1998; Prell and Day 2001). Those which cause harm to their hosts (necrotrophic and biotrophic fungi) are generally known as phytopathogenic fungi (Greek: phyto = plant, pathogenic = causing disease). The four above-mentioned groups are explained in more detail below.

Saprophytes live only on dead organic matter and are unable to attack and colonize living plants. Necrophytes, on the other hand, cause the death of the infected plant cells (Prell and Day 2001). Many of these fungi produce extracellular enzymes and toxins to help kill the host cells. After they have killed their host are they able to colonize the dead substrate. Necrotophic fungi live or at least overwinter on the dead plant material (Prell and Day 2001).

Symbiotic fungi are, literally speaking, fungi that "live together" with other species. A symbiosis is often, but not always, a mutually beneficial relationship. The most prominent group of symbiotic fungi is that of the mycorrhizae. A mycorrhiza is a fungal mycelium that is attached to the roots of a plant, whereby the fungus and the plant exchange nutrients. In general, mycorrhizal associations vary widely in structure and function. The fungus is assumed to mostly gather water, phosphor and trace minerals from an area greater than that covered by the plant's root system. In return, the plant supplies mainly sugars to the fungus (Strack et al. 2003).

As mentioned above, symbiotic relationships are not necessarily mutualistic. One example of such a potentially non-mutualistic coexistence is provided by the endophytic fungi, which live within their host plant, either inter- or intracellularly, and cause no visible phytopathogenic symptoms in their host plants (Saikkonen et al. 1998). Despite this fact, an advantage for the host is not necessarily given (Saikkonen et al. 1998). Endophytes receive nutrition from the plant, while the host plant may or may not benefit from the fungus in return (e.g. increased resistance to herbivores, pathogens and abiotic stresses). Endophytic fungi are usually found in aboveground tissue and only occasionally in roots. In contrast to mycorrhiza, they lack external hyphae (Saikkonen et al. 1998).

A phytopathogenic fungus that needs to feed on host cells that are still alive is labeled biotrophic (Greek: bios = life, trophy = feeding). Biotrophic fungi have developed a very specialized life style, whereby many of them establish a long-term feeding relationship with the living cells of their hosts, rather than killing the host cells as part of the infection process. Typically, these fungi grow between host cells and invade only a few cells to produce nutrient-absorbing structures termed haustoria (Latin: hauriere = to drink or to scoop). Through their feeding activities, they create a nutrient sink at the infection site, whereby the host plant is disadvantaged but not killed immediately (Mendgen and Hahn 2002).

The obligate biotrophic rust fungi, which include *Uromyces fabae*, the main topic of this thesis, are described separately, and in detail, in chapters 2.2 and 2.3.

Biotrophic fungi apply diverse methods for gaining entry into host plant tissue. These range from entry through natural plant openings (e.g. stomata) to various mechanisms for direct penetration through the outer surface (Howard and Valent 1996). The direct penetration mechanisms can include the use of enzymes for the breakdown of cell walls and cuticles and/or physical high-pressure penetration (Howard et al. 1991). These and other previously studied mechanisms for host cell infection by phytopathogenic fungi, together with the corresponding plant defense strategies are described below.

The most straightforward method for biotrophic fungi to gain entry into the host plant tissue is via the natural plant openings. A well-documented example hereof is the rust fungus infection process. For example, the germ tube of the rust fungus *Uromyces fabae* grows – presumably caused by a pH gradient – towards the stomata of the host plant, before forming an appressorium and entering through the stoma (Edwards and Bowling 1986; Heath and Skalamera 1997).

Examples of direct physical penetration are provided by the rice pathogen *Magnaporthe grisea* and the cereal pathogen *Colletotrichum graminicola* (Howard et al. 1991; Bechinger et al. 1999). In these cases the fungi differentiate highly specialized appressoria, which send a penetration peg into the adjacent plant cell and creates an opening through which the fungi can then enter their hosts. Penetration of the host requires an elevated appressorial turgor pressure (Howard et al. 1991; Bechinger et al. 1999). The infection apparatus of *M. grisea* can generate turgor pressures in excess of 8.0 MPa (80 bars) (Howard et al. 1991). The force exerted by the appressoria of *C. graminicola* was found to be about 17 μN (Bechinger et al. 1999).

Biochemical methods of host cell penetration have been the object of numerous scientific studies. Fungal pathogens have long been known to secrete enzymes such as cellulases and xylanases which are able to degrade plant cell structures (Kolattukudy 1985; Gomez-Gomez et al. 2002). One example is the penetration of apples by the post-harvest pathogen *Penicillium expansum*. The fungus first acidifies the ambient environment of the apple fruit before macerating the host tissue by the secretion of hydrolytic enzymes, especially poly-galacturonase (Yao et al. 1996; Prusky et al. 2004).

However, not in all cases do the secreted enzymes enhance the ability of a fungus to attack its host. This fact has been demonstrated by the targeted inactivation of two genes from the soil-borne tomato vascular wilt fungus *Fusarium oxysporum* f. sp. *Lycopersici,* whose predicted amino acid sequences show significant homology with xylanases and subtilisin-like fungal proteinases. The inactivation of these enzymes had no discernable effect on fungal virulence in tomato plants. This indicates that both genes may not be essential for pathogenicity (Di Pietro et al. 2001; Gomez-Gomez et al. 2002). Furthermore, the general role of cutinases as pathogenesis factors has also been the subject of controversial debate (Dickman et al. 1989; Stahl and Schafer 1992; Rogers et al. 1994; Kolattukudy et al. 1995).

There are several ways in which a phytopathogenic fungus can influence the metabolism of the host plant. Some plant-pathogenic fungi are able to produce substances that either reduce the activity of the host cells or inhibit their resistance completely. This is shown in the following example, which also illustrates the intricate biochemical processes that characterize plant-pathogen interactions: an isolate of *Fusarium solani* is able to produce a fructose-analog, which is bio-activated by enzymes of the host plant. Thereby a major metabolic pathway is inhibited, which results in the inhibition of root growth (Dayan et al. 2002).

Plant-pathogenic fungi can further produce plant-specific substances, which disturb the normal equilibrium of the plant and cause disruptions of growth and differentiation of both cells and tissue. For example, studies of the fungus *Gibberella fujikuroi*, have led to the discovery of gibberellins (Hori 1898; Kurosawa 1926). The gibberellins are secreted by the fungus and influence the growth of the host plant rice: As a result of the infection, the plants grow to several times their normal height and then die prematurely, a phenomenon the Japanese call "bakanae" (foolish seedling) disease (Hori 1898; Kurosawa 1926; Hooley 1994).

Having seen how fungal phytopathogens attack their host, some examples of how plants defend themselves against these diverse threats are given below. Generally, it is known that most plants remain free of serious infection by microbial pathogens (Hadwiger and Culley 1993; Mysore and Ryu 2004). This is possible based on the fact that plants have a number of defensive mechanisms at their disposal. These defenses are either of a permanent nature (constitutive), or are activated only when phytopathogenic attack is underway.

One common permanent defense mechanism is the constitutive expression of defensive barriers, such as a thick cuticle or bark. These are physical or mechanical barriers, mostly composed of lignin, suberine or waxes (Mendgen 1996; Heath 2000; Kim et al. 2000a; Kim et al. 2000b). A second permanent defense mechanism is the constitutive synthesis of poisonous substances, which many plants produce as anti-fungal secondary metabolites. These substances, for example resident anti-microbial compounds or phytoanticipins, are already present in healthy plants and represent in-built chemical barriers to infection by potential pathogens (Osbourn 1999). Another example is that of saponins, which many plant species synthesize as part of their normal program of growth and development. Saponins are antimicrobially active agents and are likely to play a role as determinants of plant disease resistance (Haralampidis et al. 2002).

A prominent example of non-permanent plant defense mechanisms against pathogen attack are the phytoalexins (Osbourn 1999). Phytoalexins are substances that are synthesized by plants to limit infection after being wounded or infected by pathogens. Phytoalexins were first discovered by Müller and Börger (1940). Müller and Börger published evidence that after infection by *Phytophthora infestans,* potato plants seemed to have some acquired resistance, and speculated that the plants produced an antifungal substance after infection – called a phytoalexin. The term phytoalexin stems from the medical term Alexin (Greek: alexein = to defend) (Müller and Börger 1940). In addition, Sbaghi et al. (1996) showed that the ability to degrade phytoalexins correlates to the pathogenicity of *B. cinerea* in grapevines.

Having seen that plants have numerous defensive mechanisms to fend off pathogen attack, the question raises itself as to how then do plants select the proper defense against a given pathogen? For this purpose, most plants have developed resistance genes (R-genes) against pathogenic attack. One the one hand, R-genes provide a competitive advantage by acting as a kind of "immune system" in plants, recognizing pathogens and inducing a defense response. On the other hand, the maintenance of outdated R-gene polymorphisms can negatively affect the fitness of the plant (Tian et al. 2003).

Just as plants have developed R-genes, fungi have evolved genes to counteract this resistance (Gene for Gene Hypothesis) (Flor 1956; Flor 1971). To verify this hypothesis, Flor carried out cross-breeding experiments between different races of the flax rust *Melampsora lini,* as well as between different races of its host, *Linum usitatissimum,* which were either susceptible or resistant to rust infection (Flor 1942; Flor 1946; Flor 1955). He concluded that an incompatible plant-fungal interaction requires the presence of one particular resistance gene *R* in the host plant and another particular avirulence gene *Avr* in the pathogen (Flor 1956).

Interestingly, plants also seem to have developed pathogenesis related genes (PR-genes) to counter non-host specific pathogen attack. For example, when studying rust interactions with the non-host model plant, *Arabidopsis thaliana,* seventeen accessions of *A. thaliana* inoculated with the cowpea rust fungus *Uromyces vignae* exhibited a variety of expressions of non-host resistance, whereby infection hyphae growth typically ceased before the formation of the first haustorium. These results suggest that the non-host resistance to rust fungi displayed by the *Arabidopsis* primarily involves the restriction of infection hyphae growth as a result of defense gene expression (Mellersh and Heath 2003).

The ecological and economic impact of plant fungal interactions on a worldwide scale is considerable. Plant pathogens can cause great damage in agricultural monocultures due to their high host specificity (Schumann 1991). One example is the coffee rust. When the coffee rust fungus, *Hemileia vastatrix*, reached Ceylon (now Sri Lanka) in 1875, coffee trees *(Coffea arabica)* were grown on about 160'000 hectares. Around that time, Ceylon was exporting ca. 45 million kg of coffee per year. By 1889, due to the defoliation of the infected coffee trees, production was reduced to 2.3 million kg (Schumann 1991). Today, tea instead of coffee is produced in Sri Lanka, and the cup of tea, rather than coffee, has become a familiar part of the United Kingdom's culture (Schumann 1991).

However, the fact that biotrophic fungi can reduce the competitive abilities of the host in a natural environment also has a positive side. A small number of biotrophic pathogens have been used as biological control agents for agricultural weeds (Scott and Delfose 1996). In Australia, blackberry leaf rust (*Phragmidium violaceum*) has been released in an effort to control of the weed *Rubus fruticosus L. aggregate* (Agriculture & Resource Management Council of Australia & New Zealand 2001). The rust fungus weakens the blackberry plants sufficiently to allow the herbicides to work more effectively (Bruzzese et al. 2000; Agriculture & Resource Management Council of Australia & New Zealand 2001).

Having outlined a number of examples demonstrating the importance of the rust fungi within the framework of plant fungal interactions, the life cycle of this type of fungus will be examined below.

2.2 Life Cycle of Rust Fungi

Rust fungi have been explicitly described in literature dating back to as early as the first century AD namely in the treatise "naturalis historia", where Plinius the Elder wrote about "*rubigo*", the rust fungus (Plinius ca. 77). In addition, more than two centuries prior, Cato, in his farming manual, advised planting beans in fields that are strong and not prone to fail. He also advised farmers to sow bread wheat and durum wheat in open, high fields where the sun shines longest (Cato ca. 150 b. C.). Considering today's knowledge that rust fungi infect best in darkness when the weather is misty and cool, we can assume that this advises will have helped to prevent severe rust infections in Roman times.

Today, more than 100 genera and some 7'000 species, of the rust fungus family (*Uredinales*) are known to exist (Maier et al. 2003). The study from which this estimation was made focused mainly on European rust and did not include any tropical taxa. Hennen & McCaine (1993) estimated between 25'000 and 125'000 species of the rust fungus family (*Uredinales*) to exist worldwide. Within this large and ecologically and economically important family of phytopathogenic fungi, there are species specialized on important agricultural crops such as coffee plants (e.g. *Hemileia vastatrix)*, wheat (e.g. *Puccinia graminis*), cowpea (e.g. *Uromyces vignae*), and beans (e.g. *Uromyces fabae, Uromyces appendiculatus)*.

In addition, at least 17 different life cycles of rust fungi have been identified (Figueiredo 2000). For example, some rust fungi have a heteroecious life style, meaning that they depend on an alternation of hosts. One stage of the pathogen's development occurs in one host, and the following in a second host, which is not necessarily phylogenetically closely related to the first host (Agrios 1997). Possibly the best-known example is black stem rust on wheat caused by *Puccinia graminis*. This fungus is coexistent with wheat throughout the world, although the most damage is caused in areas with a moderate climate and low average rainfall (Roelfs 1982). Since the alternate host of *P. graminis* is *Berberis* (barberry) or *Mahonia*, the eradication of barberry is used to control black stem rust in these regions (Roelfs 1982).

Uromyces fabae is the causative agent of rust on *Vicia faba* (Broad Bean) and has an autocious life style, meaning that it does not require an alternation of hosts. As it is the case with all biotrophic fungi, *U. fabae* must walk a very fine line between getting enough nutrition from its host for reproduction on the one hand, and not killing that very same host on the other - at least not before reproduction has occurred.

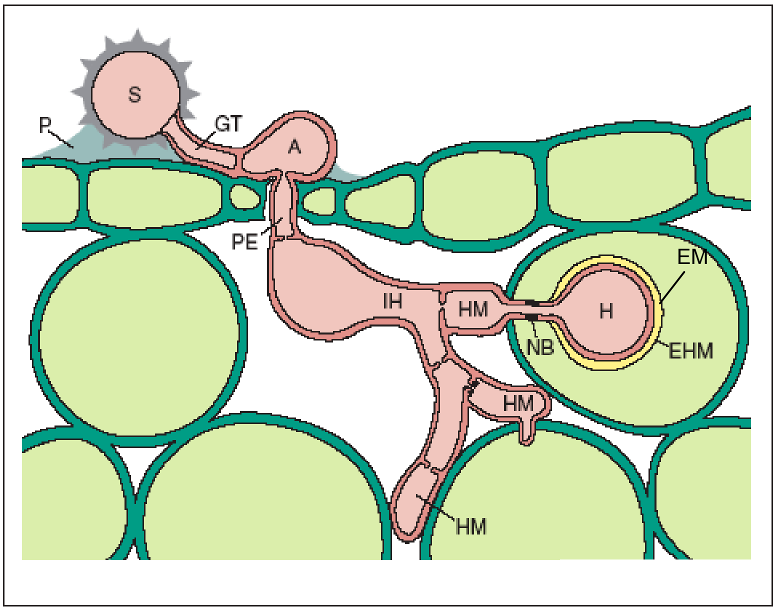


Figure 2-1 Schematic representation of early rust infection structures by *Uromyces fabae* modified according to Mendgen and Hahn (2002). (S): uredospore; (P): adhesion pad; (GT): germtube; (A): appressorium; (PE): penetration hypha; (IH): infection hypha; (HM): haustorial mother cell; (NB): neckband;(H): haustorium, (EHM): extrahaustorial membrane surrounding the (EM): extrahaustorial matrix.

The infection by *U. fabae* is divided into three phases: firstly, the penetration phase, during which the fungus enters its host, followed by the biotrophic phase, during which the fungus lives parasitically in its host, and finally the sporulation phase, during which the fungus reproduces (Mendgen 1996; Mendgen and Hahn 2002). All of the three phases are described below.

U. fabae has developed a complex strategy to achieve non-destructive host tissue penetration. In the case of (summertime) infection by uredospores, the penetration phase begins with the landing of the uredospore on its host. The uredospore proceeds to form an adhesion pad, and, depending on the amount of water available, initiates with the germination process (Deising et al. 1992; Clement et al. 1997). Upon contact with an aqueous environment, a cutinase and two nonspecific serine-esterases, located on the surface of the spores are rapidly released (Deising et al. 1992). Assumedly steered by a pH gradient, the germ tube grows in the direction of stomata until it reaches the stomatal guard cell lips (Edwards and Bowling 1986). Upon contact with these cells, the appressorium is formed (Hoch and Staples 1987; Hoch et al. 1987). During the formation of the appressoria, and before entering the host through the stoma, the rust fungus produces and secretes several enzymes (Deising et al. 2000). These include proteases (Rauscher et al. 1995) and isoformes of cellulases (Heiler et al. 1993). After having grown through the stoma, the fungus forms the substomatal vesicle and establishes primary hyphae in the plant. Eventually the fungus extends into the leaf mesophyll underneath the epidermis and forms the haustorial mother cell (Heath 1989; Xu and Mendgen 1991; Xu and Heath 1998). During the penetration phase of the infection, and until the formation of the haustorial mother cell, the rust fungus is metabolically self-sustaining (Freytag et al. 1988; Deising and Siegrist 1995).

The biotrophic phase commences with the formation of the most specialized hypha, the haustorium. Gene expression- and immuno-cytological studies of rust fungi support the idea that the haustorium is a transfer apparatus for the long-term absorption of host nutrients (Mendgen and Hahn 2002). Therefore, the metabolism of the haustorium has been studied to a great extent. Examples of these studies and a short explanation of the results are given in chapter 2.3.

Directly after the formation of the haustorium, and for reasons not fully understood, the plant cell's nucleus briefly moves into its vicinity (Heath and Skalamera 1997). The fungus proceeds to absorb nutrients and continues growing through the host tissue. Eventually, as the fungus matures, it develops sporogenous tissue and the sporulation phase begins.

During the sporulation phase, new spores are formed in the sporogenous tissue and become visible as blisters on the host's surface. Upon maturity, the blisters burst open and the spores are released to recommence the infection process. During their life cycle, rust fungi form different types of spores (aecio-, basidio-, pykno-, teleuto-, and uredospores) for infection and reproduction purposes (Agrios 1997). Important for the infection process are the hardy teleutospores, which are responsible for the initial infection in springtime, and the more numerous uredospores, which are responsible for epidemic infection during summer (Agrios 1997).

2.3 Molecular Studies of *Uromyces fabae*

Various *in planta* induced genes (*PIG*s) of *U. fabae* have been detected using a haustorium-specific cDNA library and have been partially described (Hahn and Mendgen 1997). In addition, a number of *Uromyces fabae* genes have been characterized. The following is a brief outline of selected, previously published *Uromyces fabae* gene studies related to amino acid transport, sugar uptake and vitamin synthesis, based on the pioneering work of K. Mendgen, M. Hahn, C. Struck, S. Wirsel and R. Voegele:

Amino acid transport

> Over the past years, several amino acid transporters of *U. fabae* (*AAT1, AAT2, AAT3*) have been described (Hahn and Mendgen 1997; Hahn et al. 1997; Martin 2001; Müller 2001; Struck et al. 2002; Voegele and Mendgen 2003; Struck et al. 2004). The genes for fungal amino acid transporters showed dissimilar regulation. Transcripts of the first and the third transporter were detected during all developmental stages, whereas those of the second transporter appeared to be developmental stage-dependent (Wirsel et al. 2001; Struck et al. 2004).

Sugar uptake and nutrition

In order to extract nutrition from its host, *Vicia faba*, the rust fungus *U. fabae* has developed an ingenious system by which hexoses are cleaved from sucrose within the plant and transferred into the fungal haustorium (Voegele et al. 2001). To cleave the sucrose – the only easily available sugar - within the host, the fungus evidently secretes an invertase (INV1p) that most likely cleaves sucrose into fructose and glucose (Möll 2001). The hexoses are then transferred from the host plant into the fungal haustorium by the hexose transporter HXT1p (Voegele et al. 2001).

The presence of the hexose transporter HXT1p appears to be dependent on the development stage of the fungus (Voegele et al. 2001; Wirsel et al. 2001). Given that INV1p is not only found during the same development stages as HXT1p, but also much earlier, INV1p potentially has additional functions besides just cleaving sucrose. Possibly, INV1p is secreted into the plant to condition the infected tissue into becoming a hexose sink and therefore a source of nutrition for the fungus (Möll 2001; Voegele and Mendgen 2003).

Efficient nutrient mobilization is a key success factor for biotrophic plant parasites such as the rust fungi. In the course of a cDNA library screening for elements involved in sugar utilization in *Uromyces fabae*, a sequence was identified as a fungal beta-glucosidase (*BGL1*) on the basis of sequence homology. *BGL1* seems to be expressed in all stages of growth, including the haustorial stage. In the course of immuno-localization studies, the gene product BGL1p was located in the periphery of intercellular hyphae and haustoria (Haerter and Voegele 2004).

Nutrient uptake by *U. fabae* from host cells is promoted by an H^+-ATPase that generates an electrochemical gradient across the haustorial membrane, thus playing an important role in the haustorial function (Struck et al. 1996). The mRNA for such plasma membrane ATPases was detected in spores and all infection structures (Struck et al. 1998; Wirsel et al. 2001).

Vitamin synthesis

Haustoria of *U. fabae* produce notable amounts of THI1p and THI2p. THI1p seems to be involved in the synthesis of Hydroxymethylpirimidin-pyrophosphate, whereas THI2p is involved in the synthesis of Hydroxyethylthiazol-phosphate. The condensation product of these two molecules is thiamin-monophosphate. Both of the thiamin biosynthesis genes are regulated in a development stage-dependent manner (Sohn 2000; Sohn et al. 2000).

The research above shows that rust haustoria are not mere feeding structures but are involved in redirecting or reprogramming the metabolic flow of its host plant (Voegele and Mendgen 2003). In addition, a large percentage of the *U. fabae PIG*s extracted from the above-mentioned haustorium-specific cDNA library remains uncharacterized. Thus, M. Hahn initiated a small-scale Expressed Sequence Tag (EST) Project (explained below) based on haustorial cDNA in order to gain more insight into the events in the haustoria of *U. fabae* (M. Hahn et al, unpublished).

2.4 EST Projects as a Method for Studying Plant Fungal Interactions

There is great interest in genes expressed specifically during the biotrophic stage of the fungal life cycle because of the notion that such genes may play a significant role during development of pathogenicity (Kahmann and Basse 2001). One common method used to study these genes is the execution of an EST project.

An Expressed Sequence Tag or EST is a single pass read from a cDNA clone, randomly selected from a cDNA library. The resulting sequence information can then be used for homology searches in public databases in order to retrieve information about the putative function of the gene product of an EST. EST projects are also helpful in obtaining an overview of the expressed genes in certain tissues or time-dependent differentiation stages.

An EST often represents only a small part of an entire gene. ESTs therefore have the advantage of being quickly and economically obtainable, allowing for mass sequencing with limited expense both in time and resources.

To date, only a small number of EST projects related to phytopathogenic fungi have been publicized. A list detailing some of these projects can be found at http://cogeme.ex.ac.uk (Soanes et al. 2002). As of December 2004, EST projects of thirteen plant pathogens were briefly described on this Internet site. In addition to this list, Soanes et al. (2002) briefly reviewed the genomic and bioinformatic resources available to date and the status of fungal genomics.

Once an EST is found to have a significant homologue in the public databases, a first assumption can be made regarding the function of the corresponding gene product. In case of further interest, the full-length clone of this gene is obtained or reconstructed. Thereafter, an in-depth functional analysis of the gene respectively gene product is carried out. In the case of *U. fabae* this approach led to the results described in chapter 2.3.

If a transformation system for the studied organism is already available, an additional analysis using knockout mutants can be carried out. If a transformation system for the studied organism does not exist, the above-mentioned functional analysis is performed directly by cloning the gene into expression vectors and transforming a suitable organism. Unfortunately, no stable transformation system for *Uromyces fabae*, the subject of this thesis, exists to date.

ESTs can also be used to assemble DNA arrays, a powerful tool for the analysis of the entire expression pattern of a tissue, a specific differentiation stage or mutants (DeRisi et al. 1997). This method allows the identification of genes involved in the same metabolic pathway as well as the study of interaction between regulatory genes respectively their products. When studying plant-fungal interactions, the time- and infection stage- dependent regulation of fungal- and plant genes can be monitored using DNA arrays. This kind of monitoring has the advantage that those genes relevant for fungal pathogenesis and plant resistance can be found more easily.

There are also a number of readily accessible IT-based tools available that allow for a theoretical study of the data provided by an EST project (http://au.expasy.org/, as of October 2004). For example, publicly available online genetic information on intercellular sorting and secretion (e.g. PSORT II) can provide valuable hints concerning the potential function of a gene product.

As described above, an EST project can also uncover genes that have no homologies to sequences in public databases. To gain knowledge about these genes (resp. their gene products), and the possible role they play in the interaction between parasite and host, other aspects besides known homologies must be studied. In case of *U. fabae*, these aspects include gene product (i.e. protein) export and localization, as well as development stage-dependent expression profiles. To this end, some theoretical background information regarding protein transport is provided in chapter 2.5.

2.5 The Logistics of Protein Transport

2.5.1 Protein Sorting and Secretion

A common feature in cells of prokaryotic and eukaryotic origin is the transport of proteins from their site of synthesis, mostly the cytoplasm, to other destinations either inside or outside the cell. This combination of protein sorting, intracellular transport, and secretion to the surrounding environment, is a complex process, which includes numerous steps and requires sophisticated logistics.

In eukaryotic cells, protein synthesis takes place to a large extend on ribosomes that are initially freely mobile in the cytosol. Some proteins are synthesized as immature pre-proteins with an amino-terminal, transient targeting sequence (also termed signal sequence or leader sequence), which serve to direct the protein to its proper destination. The subsequent transport of these proteins to their final destination is then realized using one of two very distinct pathways (Alberts 1994).

The journey on the **first pathway** begins with the release of the finished protein into the cytosol. There the protein remains, unless it contains a sorting signal for further transport to the mitochondria, peroxisomes, chloroplasts (in plants) or nucleus (Kelly 1985; Alberts 1994). Because during this thesis, evidence was found, that one of the studied proteins (PIG7p) could be located in the nuclei of the host cells, the general protein transport into the nucleus is further described below.

Approximately 17% of all eukaryotic proteins may be imported into the cell nucleus (Cokol et al. 2000). Transport into the nucleus is thus a common occurrence and has been described and reviewed frequently (Yoneda et al. 1999; Hodel et al. 2001; Macara 2001; Kosugi and Ohashi 2002). To facilitate their transport into the nucleus, these proteins contain a nuclear localization signal (NLS), which is explained in more detail in chapter 2.5.2. Transportation of macromolecules, such as proteins, from the cytoplasm to the nucleus occurs through nuclear pore complexes (NPCs) (Görlich and Kutay 1999). The NPCs create aqueous channels across the nuclear envelope through which macromolecular transport between nucleus and cytoplasm occurs (Davis 1995). Small molecules diffuse passively through the NPCs, while nuclear import and export of larger proteins is an active, signal-mediated process. These larger proteins must bind to receptors at the nuclear pores and be actively transported through the NPCs into the nucleus (Bonner 1978; Lang et al. 1986; Fried and Kutay 2003).

Traveling along the **second pathway**, (pre-)proteins destined for secretion, the cell organelles, the plasma membrane, or the cell membrane, are first transported to the endoplasmic reticulum (ER). These proteins enter the ER during their synthesis, facilitated by their (mostly N-terminal) leader sequence (Kelly 1985). The proteins are then either retained in the ER or transported to the Golgi apparatus after proper folding and modification (Alberts 1994; Kostova and Wolf 2003). Upon arrival in the Golgi apparatus, the proteins may be retained anew, sorted to the lysosomes and endosomes or, upon fusion of the secretory vesicles membranes with the plasma membrane, sorted to the extracellular space (Alberts 1994).

The importance and function of targeting sequences such as the N-terminal leader sequences and Nuclear Localization Signals are described in more detail in chapter 2.5.2 below.

2.5.2 Targeting Sequences

For different protein destinations, varying types of targeting sequences are known to exist. During this project, attention was focused on two specific types of sequences: N-terminal leader sequences and Nuclear Localization Signals (NLS).

The existence of an N-terminal leader sequence was a pre-requisite characteristic of the genes studied during this thesis. The interest in NLS stems from the fact that, such a sequence was found in PIG7p, which in turn could be located in the nucleus of the host plant.

Leader sequences of proteins are small N-terminal sequences (18 to 30 amino acids) containing information about the future localization of the respective protein. In most cases of protein export, the leader sequence is responsible for the passage into or through a membrane (Muesch et al. 1990; Alberts 1994; Izard and Kendall 1994). However, in most cases, the N-terminal leader sequence is not part of the mature protein and only in some cases the leader remains part of the mature protein (Muesch et al. 1990). A protein carrying this leader sequence is called a pre-protein, and is a transient precursor to the mature protein. Leader sequences of proteins designated for secretion are cleaved from the protein by a protease while passing through the ER (Burgess and Kelly 1987; Alberts 1994).

As described above, certain proteins, despite being larger than the passive diffusion limit, can accumulate within the nucleus (Kalderon et al. 1984; Adam et al. 1990; Görlich and Kutay 1999). These observations have led to the discovery of so-called Nuclear Localization Signals (NLS) within the protein, for example in the simian virus 40 (SV40) reviewed by Macara (2001). A NLS is usually a short internal sequence of amino acids located within the protein, which mediates the import of proteins into the nucleus (Cokol et al. 2000). NLS have basic residues and in rare cases hydrophobic residues. It has been shown that the deletion of a NLS disrupts nuclear import (Cokol et al. 2000).

Different types of nuclear localization signals have been described (Macara 2001). They are listed in the following table (Table 2-1). Macara also assumes that many more NLS must exist, but, so far, none have been identified (Macara 2001).

Table 2-1 Nuclear Localization Signal Sequences (Macara 2001)

Signal	**Proposed consensus sequence**	**Source of reference** cited in Macara (2001)
Classical monopartite NLS	B_4,P(B_3x),Pxx(B_3x),B_3(H/P)	PSORT II server
Classical bipartite NLS	BBx_{10}(B_3x_2)	PSORT II server
M9 NLS	(Y/F/W)x_2JxSxZG(P/K)(M/L/V)(K/R)	(Bogerd et al. 1999)
Viral NLS	RxxRR$x_{1,2}$RBR	(Palmeri and Malim; Truant and Cullen)
Ribosomal L23 NLS	VHSHKKKKIRTSPTFRRPKTLRLRRQPKYRRKSAP RRNK	(Jakel and Gorlich)

B, basic residue (K or R); J hydrophilic residue; Z, hydrophobic residue; x any residue; subscript numbers show number of residues; letters in parentheses can be any order; letters separated by slash are alternate permitted residues.

A NLS may operate using various possible mechanisms. The signal could either bind directly to the nuclear pore channel (NPC) and subsequently be transported through the pores, or the signal could be recognized by soluble receptors, which then transport it through the pores. For permeabilized mammalian cells, it was shown that nuclear protein import requires soluble cytoplasmic factors (Adam et al. 1990; Adam and Adam 1994), of which four different variants are known (see Table 2-2). Most of the proteins that carry other molecules through the NPC are members of the importin β (also termed karyopherin β) family (Macara 2001). Importin β is a complex carrier able to transport its cargo either via direct bonds (at defined sites) or with the help of an adapter protein (importin α) that binds to the NLS of the cargo.

Table 2-2 Soluble Import Factors (Macara 2001)

Name	**(Proposed) function**	**Source of reference** cited in Macara (2001)
Importin α (karyopherin α, Kapα, PTAC58)	Adapter protein 55kDa binds to NLS	(Adam et al. 1990; Adam and Adam 1994; Görlich et al. 1994; Moroianu et al. 1995)
Importin β, karyopherin β, p97, PTAC97)	Binds to nucleporins, directs cargo carrier or cargo carrier via adapter protein	(Chi et al. 1995; Görlich et al. 1995; Radu et al. 1995)
Ran (TC4)	GTPase	(Melchior et al. 1993; Moore and Blobel 1993)
NTF2 (p10, pp15)	Part of a multi-component system of cytosolic factors that assembles itself at the pore complex during nuclear import	(Paschal and Gerace 1995; Paschal et al. 1996)

One example of the importance of NLS in plant pathogen interactions is the avirulence protein, RRS1-R, of *Arabidopsis thaliana*. RRS1-R, which has a putative NLS, confers broad-spectrum resistance to several strains of the causative agent of bacterial wilt, *Ralstonia solanacearum* (Deslandes et al. 2003). This bacterium is a devastating plant pathogen with global distribution and a wide host range. A physical interaction between RRS1-R and PopP2, an avirulence protein of *R. solanacearum* and a type III effector targeting the plant nucleus, has been identified. The co-localization of the Avr protein PopP2 and the RRS1 protein, in the nucleus of *Arabidopsis* protoplasts, has been shown using GFP transformants (Deslandes et al. 2003).

The above example highlights another aspect of plant pathogen interactions and protein transport. As can be seen in the case of PopP2, mechanisms exist which allow for the export of alien pathogen proteins into the nucleus of the host cells, not just into the cell nucleus of the pathogen itself.

The results of this thesis indicate that one of the studied proteins, PIG7p (RTP1p) has an NLS targeted at the host cell nucleus. This protein can be found at the plant fungal interface as well as in the host nucleus itself, a phenomenon that has never been observed in plant fungal interactions before. Therefore, the process of alien protein export into host cells is described in chapter 2.5.3 below.

2.5.3 Alien Protein Transport into Host Cells

The results of this thesis indicate the presence of a fungal protein from *Uromyces fabae* in the nucleus of its host plant, *Vicia faba*. To date, very little is known about the transport of fungal proteins into plant nuclei. Thus it may be helpful to study known examples of alien protein transport into host cells by other pathogens.

For example, alien protein transport into host cells has been described in some detail with regard to the plant pathogenic soil bacterium, *Agrobacterium tumefaciens*, the gram-negative bacterial tomato pathogen, *Xanthomonas campestris* (Szurek et al. 2002) and the simian virus, SV40 (Herrera-Estrella et al. 1990; Citovsky et al. 1992; Citovsky et al. 1994; Ballas and Citovsky 1997). Despite the substantial differences between bacteria, viruses and fungi, the findings made in connection with these three pathogens are described below, as analogies with the phenomenon observed in *U. fabae* during this thesis may exist.

A. tumefaciens is capable of transferring its DNA and some of its proteins (VirD1 and VirD2) into the cell nucleus of the host plant, tobacco. Supporting evidence is provided by experimental findings that the amino-terminal portion of the *A. tumefaciens* VirD2 protein is able to direct a β-galactosidase fusion protein into the tobacco nucleus (Herrera-Estrella et al. 1990).

In case of gram-negative bacterial (plant) pathogens it is known that sophisticated strategies are employed to invade and colonize plants. Many of these pathogens use a specialized secretion system as a molecular syringe to inject effector proteins directly into the host cell reviewed by Büttner and Bonas (2002). These effector proteins modulate a variety of host cellular pathways, such as rearrangements of the cytoskeleton and defense responses (Vivian and Arnold 2000; Büttner and Bonas 2003). One example is an effector protein from the bacteria, *Xanthomonas campestris* pv. *Vesicatoria,* which can be located in the nuclei of infected pepper leaves (Minsavage et al. 1990; Vivian and Arnold 2000; Szurek et al. 2002). This gene belongs to an avirulence gene family, which encodes proteins targeted to plant cells by the abovementioned specialized bacterial secretion apparatus. A member of this gene family is also found in *Xanthomonas oryzae* pv. *Oryzae,* where it is targeted to the host cell nucleus and has the potential to interact with the host DNA (Yang et al. 2000).

Another well-studied pathogen is the simian virus, SV40. In order to infect monkey- and other mammalian cells, SV40 transfers DNA and the T-antigen into its host. In addition, virion particles and viral structural proteins of SV40 can be located in the host nucleus. It is assumed that these viral structural proteins facilitate virion import into the nucleus and viral gene expression (Clever et al. 1993; Yamada and Kasamatsu 1993).

During the infection process, SV40 binds to certain molecules on the cell surface and enters cells via a unique endocytic pathway. SV40 is then delivered to the endoplasmic reticulum (reviewed by (Norkin 1999). Subsequently, SV40 traverses towards the nucleus. Using the cell's own nuclear import machinery, the viral genome enter the nucleus through the nuclear pore complex (NPC) (Yamada and Kasamatsu 1993; Dean 1997; Kasamatsu and Nakanishi 1998). In this context, it has been shown that in order to accumulate in the host cell nucleus, the SV40 large T antigen requires an NLS (Kalderon et al. 1984).

As mentioned previously, in-depth knowledge about alien protein transport into host cell nuclei is limited. Moreover, the function of alien proteins in the plant nucleus can only be speculated upon. This is particularly true when discussing fungal protein export to plant nuclei, although, as will be shown later on in this thesis, fungal pathogen proteins can in fact be found in the host plant nucleus.

2.6 Goals of this Thesis

The goal of this thesis was to further deepen the knowledge currently available on plant fungal interactions based on a study of the genes expressed during the biotrophic phase of the life cycle of the rust fungus, *Uromyces fabae*.

As a first step, a previously initiated partial genome-sequencing project was to be completed and the results analyzed. The aim of this project was to sequence at least 1’000 ESTs (Expressed Sequence Tags) and to classify these ESTs into putative functional groups by comparing them to publicly available sequences.

As a follow-up step, an analysis was to be conducted to verify the presence of a signal sequence in those genes sequenced during the EST project which had been previously described as “*in planta* induced” genes (*PIG*s) (Hahn and Mendgen 1997). Finally, three of those *PIG*s identified as having a signal sequences, and thus potentially being secreted, were to be analyzed in more detail, including:

- unusual sequence characteristics,
- expression profiles,
- number of gene copies,
- verification of secretion signals, and, most importantly of all,
- final localization of the corresponding gene products *in planta*.

3 Materials and Methods

3.1 Use of Plants and Fungal Spores

3.1.1 Plant Cultivation

To insure the availability of suitable host plants for the experiments carried out during this thesis, Broad Bean seeds (*Vicia faba* cv. Con amore, Nickerson Zwaan, Edemissen, D) were planted in steam-treated soil and cultivated in a climate chamber (Lichtthermostatschrank, Laborgerätebau E. Schutt, Göttingen, D) for 21 days (21°C; 230 $\mu Em^{-2}S^{-1}$) with a 16 h photoperiod. Thereafter they were transferred to the greenhouse and subjected to a varying number of day/night cycles (16 h light / 8 h darkness).

3.1.2 Inoculation and Harvesting Methods

To generate the necessary amount of infected *Vicia faba* leaves, 21-day-old Broad Bean plants were spray-inoculated with *Uromyces fabae* uredospores (race I2) (Deising et al. 1991). The suspension was prepared by mixing 500 ml of deionized water with 350 mg fat-reduced milk powder and 250 mg uredospores. The suspension was evenly sprayed onto the plants. In case only individual leaves needed to be inoculated, the uredospores were applied with a brush.

The plants were then kept in the dark for 24 h at 100% humidity and allowed to dry for at least 12 h under ambient conditions without extra illumination. Thereafter the plants were subjected to a varying number of day/night cycles (16 h light / 8 h darkness). New shoots and buds were removed at least every three days. This prevented the plants from shedding their infected leaves. Those leaves required for further processing were subsequently harvested using a razor blade.

3.1.3 Isolation of Haustoria from *Uromyces fabae*

Isolation of haustoria from *Uromyces fabae* was performed according to Hahn and Mendgen (1992). Five to seven days after infection, 25 g of heavily infected leaves without visible uredospores on the leaf surface were gently washed in cold, deionized water to remove spores and mycelium from the surface. Because most of the haustoria were used for RNA preparation, the following steps were carried out in the cold storage room with pre-cooled solutions and, if possible, on ice.

After transfer to the cold storage room, leaves were first homogenized with 160 ml Homogenization buffer using a Warring blender (max. rpm, 15 sec.). To remove the debris, the leaf homogenate was filtered through nylon tissue with 20 μm pore size and rinsed once more with 20 ml Homogenization buffer. The filtrate was then centrifuged for 5 min (4°C, 5’000 g), and the resulting pellet, consisting mainly of chloroplasts and haustoria, was brought into suspension with 8 ml Suspension Buffer resulting in a concentration of approximately 1 to 1.5 x 10^9 haustoria/ml.

Accumulation of haustoria was carried out with Concanavalin A bound to a Sepharose 6MB column (Hahn and Mendgen 1992). To prepare the column, 5 g Sepharose 6 MB Megabeads (Amersham Bioscience Europe GmbH, Freiburg, D) were washed with 1 mM HCl for 15 min, before being rinsed with Coupling Buffer. 12 ml of this gel were mixed with 75 mg Concanavalin A in 12 ml Coupling Buffer and put on a roller over night at 4°C. Afterwards, the Sepharose gel was washed thoroughly with Coupling buffer and the remaining active groups were blocked with Blocking Buffer (see Table 3-2) for 2 h.

After the column had been prepared, haustoria (and chloroplast) suspension was applied until it was dark green. The haustoria were then allowed to bind to Concanavalin A for about 10 minutes. The column was washed with Suspension Buffer until clear and the binding step was repeated with new haustoria and chloroplast suspension. After the entire suspension had been applied to the column and the excess chloroplasts had been washed away, the matrix was vigorously dispersed with 2 ml Suspension Buffer to remove the haustoria from the Sepharose beads and to bring them into suspension. The haustoria in the Suspension buffer were then transferred to an Eppendorf tube and centrifuged for 5 min at 12’000 rpm and 4°C. The supernatant was discarded, while the pellet was either processed directly or frozen in liquid nitrogen and stored at –70°C.

Table 3-1 Homogenization- and Suspension Buffer for Haustoria Isolation (Hahn and Mendgen 1992)

Homogenization Buffer:		**Suspension Buffer:**	
300.0 mM	Sorbitol	300.0 mM	Sorbitol
20.0 mM	3-morpholinopropansulfonic acid, pH 7.2	10.0 mM	3-morpholinopropansulfonic acid, pH 7.2
0.1%	BSA	0.2%	BSA
0.2%	β-Mercaptoethanol	1.0 mM	KCl
0.2%	PEG 6000	1.0 mM	$CaCl_2$
		1.0 mM	$MnCl_2$

Table 3-2 Coupling and Blocking Buffer for Haustoria Isolation (Hahn and Mendgen 1992)

Coupling Buffer pH 8.6		**Blocking Buffer pH 8.3**	
0.1 M	$NaHCO_3$	0.1 M	$NaHCO_3$
0.5 M	NaCl	0.5 M	Glycine

3.1.4 Preparation of *Uromyces fabae in vitro* Structures

Germ tube preparation:

Germ tubes are *U. fabae* infection structures that can be easily isolation in a pure fraction. After 4 g Uredospores of *U. fabae* had been washed for 30 min in 1.5 l of deionized water with 0.01% Tween20, they were allowed to grow in the dark in 4 l of deionized water with 0.01% Tween20 for 3 to 5 h. In order to ensure a permanent oxygen supply, the spores were mixed at high speed using a drilling machine with an attached flexible tube instead of a drill bit (Hahn and Mendgen 1997). The spores were checked for germination under a microscope and if germination was at least 80%, the spores and germ tubes were vacuum filtered through two layers of filter paper in a Buchner funnel. The nearly dry spores were then collected for further use.

Preparation of other infection structures (germ tube to haustorial mother cell):

It is possible to produce *Uromyces fabae* infection structures on artificial surfaces (Deising et al. 1991). This approach is useful for isolating specific infection structures (such as appressoria, penetration hyphae, substomatal vesicle, infection hyphae and haustorial mother cells). It must be noted that with this approach, individual infection structures are not isolated in pure fractions but in fractions were one of the infection structure stages is predominant.

In preparation, Polyethylene foil of the type Polystar® 300 x 0.10 mm (Rische + Herfurth GmbH, Hamburg, D) were heated to 35°C for a duration of 8 weeks and stored under ambient conditions for several months to allow for the dissipation of volatile additives. The thus prepared Polyethylene foil was then introduced into the base of an inoculation tower, where 400 mg of uredospores were subsequently dispersed using compressed air. After the spores were allowed to settle on the PE foil for 20 min, the foil was moistened with about 4 ml H_2O and incubated in a dark room (19°C, 100% humidity) for a defined period (4 h to 24 h). The subsequent formation and presence of the specific infection structures (appressoria, penetration hyphae, substomatal vesicle, infection hyphae and haustorial mother cells) was verified and quantified using light microscopy.

3.1.5 Bean Leaf Infiltration Experiments

Infiltration experiments were carried out to study what effect the introduction of antibodies against the studied PIGps into the plant has on the development of the infection. For this purpose, 3 to 4 week old broad bean plants (*Vicia faba*) in different stages of *U. fabae* infection were used. A solution containing antibodies against the PIGps studied during this thesis in a dilution of 1/10 and 1/100 was injected into one leaf per plant using a syringe with a fine needle. After about 12 h under ambient conditions without extra illumination, the plants were transferred to a green house. Thus treated plants were subjected to a varying number of day/night cycles (16 h light / 8 h darkness) before being studied and compared for variation in infection development.

3.2 Media, Strains and Plasmids

3.2.1 Media

Within the framework of this project, various media were used to cultivate both bacteria and fungi. During the medium preparation process, all of these were autoclaved for 20 min at 121°C.

For cultivation of bacteria, both LB- and NZY media were used:

Table 3-3 LB Medium pH 7.5* (Luria et al. 1960; Sambrook et al. 1989)

Contents	Weight/ Volume	Manufacturer
Bacto-Trypton	10.0 g/l	Difco Laboratories, Detroit, MI, USA
Yeast extract	5.0 g/l	Select Yeast Extract Sigma Aldrich Chemie GmbH, Steinheim, D
NaCl	10.0 g/l	C. Roth & Co. GmbH, Karlsruhe, D

*The pH of the medium was adjusted to pH 7.5 with 5 M NaOH, before being autoclaved.

LB medium was first described in the early days of phage genetics studies (Luria et al. 1960) and is still widely used by the molecular biologist today. All of the formulations contained 10 g tryptone and 5 g yeast extract/liter. Some also included 0.1% sugar, which is not used in this case. NaCl amounts can vary between 0.5 g and 10 g. For *E. coli* growth within the framework of this thesis, 10.0 g NaCl per liter were used.

To prepare the **LB agar** plates used during this thesis, 18 g/l Agar-Agar was added to freshly prepared (not autoclaved) LB medium and autoclaved. If a selection for antibiotic resistance of transformed bacteria was required, one of the following antibiotics was added to the LB agar after autoclaving:

- Ampicilline 100 μg/ml
- Kanamycine 50 μg/ml

The medium was then poured into Petri dishes and stored at 4°C.

Table 3-4 NZYM (NZY) pH 7.5 (Blattner et al. 1977; Sambrook et al. 1989)

Contents	Weight/ Volume	Manufacturer
NZ-Amine	10.0 g/l	Sigma Aldrich Chemie GmbH, Steinheim, D
Yeast extract	5.0 g/l	Select Yeast Extract Sigma Aldrich Chemie GmbH, Steinheim, D
NaCl	5.0 g/l	C. Roth & Co. GmbH, Karlsruhe, D
Above ingredients were autoclaved for 20 min at 121°C, before addition of		
2 M $MgSO_4$	4.0 ml/l	

Traditionally, the *E. coli* host for lambda phages is grown on a NZY medium. This medium is not as rich as LB medium, and *E. coli* growth is slower, making it less likely that *E. coli* will overgrow the phage.

During this thesis, NZY medium was used to prepare TopAgar for phage lysates. TopAgar was prepared as follows: 0.7%, agarose (Carl Roth GmbH + Co, Karlsruhe, D) was mixed with NZY-medium and autoclaved. If required, TopAgar was reheated several times in the microwave oven. For a description of the subsequent plating procedure, see chapter 3.5.

For cultivation of all fungi besides *Botrytis fabae* and *Ascochyta fabae*, malt extract was used:

Table 3-5 Malt Extract

Contents	Weight/ Volume	Manufacturer
Malt extract	30.0 g/l	Biomalt, Gesundprodukte, Kirn D

Malt extract is a rich medium for fungal growth and stock culture preparation. If not otherwise mentioned, all fungi used for further DNA preparation were grown using malt extract in liquid culture (20°C, 80 rpm, up to one week). Malt extract was also used for infiltration experiments with *Botrytis fabae* in a dilution of 1/1'000 to stimulate pathogenesis of this fungus.

The cultivation of *B. fabae* and *A. fabae* took place using oatmeal- and potato-dextrose liquid culture medium, as well as hay extract:

Table 3-6 Potato Dextrose Liquid Culture Medium (MacFaddin 1985)

Contents	Weight/ Volume	Manufacturer
Potato Dextrose Broth	24 g/l	DIFCO Laboratories, Detroit, MI, U.S.A

Table 3-7 Oatmeal Liquid Culture Medium

Contents	Weight/ Volume	Procedure
Oatmeal	30 g/l	Freshly ground, organically grown oatmeal was cooked for about 0.5 h in tap water. The liquid phase was then decanted and autoclaved for 20 min at 121°C

Table 3-8 Hay Extract

Contents	Weight/ Volume	Procedure
Hay	30 g/l	Dry hay was cooked for about 1 h, then autoclaved for 20 min at 121°C for a better extraction, filtered, replenished to 1'000 ml and again autoclaved for 20 min at 121°C.

Botrytis fabae was also cultivated on Oatmeal agar or Hay agar. To create Oatmeal agar, the above described oatmeal liquid culture medium was mixed with 12 g/l agar after cooling. The agar was then autoclaved for 20 min at 121°C. To make Hay agar, the prepared hay-extract was mixed with 15g/l agar after cooling. The agar was then also autoclaved for 20 min at 121°C.

Oatmeal- and Hay agar were used to stimulate the spore formation of *Ascochyta fabae* and *Botrytis fabae*. Whereas Hay agar is a minimal medium, Oatmeal agar provides more nutrition, especially in the form of starch.

The yeast *S. cerevisiae*, strain RE700A (see Table 3-12) used during this thesis was cultivated in an SC-Maltose medium with uracil. If the yeast had been transformed with the vector pDR195, the cultivation medium did not contain uracil:

Table 3-9 SC-Maltose Medium (Shermann 1991)

Contents	Weight/ Volume	Manufacturer
Yeast Nitrogen Base	6.8 g/l	Difco Laboratories, Detroit, MI, USA
Maltose	20.0 g/l	Merck, Darmstadt, D
Dropout mix	2.0 g/l	See Table 3-10

The SC-maltose medium was used to prepare SC-maltose agar plates. For this, 17 g Bacto-agar were added to 500 ml aqua bi-distilled in a one-liter flask. The other ingredients were put into a separate flask and filled up to 500 ml with aqua bi-distilled. Both flasks were autoclaved separately, and after having been allowed to cool to approximately 80°C, both solutions were mixed together and poured into Petri dishes.

Table 3-10 Dropout Amino Acid Mix for SC-Maltose Medium

Contents	Weight	Contents	Weight
Adenine	0.5 g	Alanine	2.0 g
Arginine	2.0 g	Asparagine	2.0 g
Aspartarte	2.0 g	Cysteine	2.0 g
Glutamine	2.0 g	Glutamate	2.0 g
Glycine	2.0 g	Histidine	2.0 g
Isoleucine	2.0 g	Leucine	2.0 g
Lysine	2.0 g	Methionine	2.0 g
Para-aminobezoate	0.2 g	Phenylalanine	2.0 g
Proline	2.0 g	Serine	2.0 g
Threonine	2.0 g	Tryptophane	2.0 g
Thyrosine	2.0 g	Valine	2.0 g

This dropout mix does not contain uracil. For the SC-Maltose + Uracil Medium, the dropout mix additionally contains 2.0 g uracil.

3.2.2 Strains

The following bacterial- and yeast strains were used during this thesis:

Table 3-11 Bacterial Strains

Name	Genotype	Reference
BL21 (DE3) plys *E* (*E. coli*)	*F^-, ompT, $hsdS_B$, (r_B^- m_B^-), gal, dcmΔ (mrcA,) 183Δ,(mrcCB-hsdSMR-mrr,) 173 endA1 supE44, thi-1, gyrA96, relA1, lac(P2lysogen); pLysE*	(Studier et al. 1990)
NM514 (*E. coli*)	*HsdR514, (rK^-mK^-), argH, galE, galX, lycB7, strA(Hfl^+)*	(Murray et al. 1977)
TG1 (*E. coli*)	*SupE hsdΔ thiΔ(lac-mo AB)[F'(traD36, pro AB, $^+lacI^q$, lacZ ΔM15)]*	(Sambrook et al. 1989)

For recombinant protein production, the *Escherichia coli* strain BL21 (DE3) plys *E* was used. This strain descends from *E. coli* B and is deficient in both *lon-* and the outer membrane *ompT* proteases. Cells deficient of these proteases accumulate recombinant proteins at a high rate and are less likely to degrade these proteins during purification.

In order to obtain the sequence information for the EST project (see chapter 3.5), the *E. coli* strain NM514 was infected with λGT10 containing the *Uromyces fabae* gene bank.

For the production of plasmids for further transformation, the *E. coli* strain TG1 was used (see chapter 3.7.1).

In addition to the bacteria strains, the following yeast strain was transformed with pDR195, which contained the sequences of the *PIG*s to be studied (see chapter 3.7.2 and Table 3-13), and used to produce the PIG-proteins.

Table 3-12 Yeast Strain

Name	Reference
RE700A (*S. cerevisiae*)	(Reifenberger et al. 1995)

3.2.3 Phages and Vectors

The basic data for the EST project was obtained from a cDNA library of haustorial genes (Hahn and Mendgen 1997) kindly provided for use within the framework of this thesis by M. Hahn. The library was created using a TimeSaver cDNA synthesis kit (Amersham Bioscience Europe GmbH, Freiburg, D), and the λGT10 phage.

The production of recombinant proteins for antibody production and the testing of PIGp signal sequences was realized using two different plasmid vectors, which are described in more detail in chapter 3.6:

Table 3-13 Plasmid Vectors

Vector	Source/ Reference	Selection	Promotor
pET28a$^+$	Novagen, CNBiosciences, Inc., San Diego, CA, USA (Studier and Moffatt 1986)	Kanamycine	T7ϕ
pDR195	(Rentsch et al. 1995)	Ampicilline	PMA1

3.3 PCR Applications

3.3.1 PCR-Cycler

A tool commonly applied during this thesis was the polymerase chain reaction (PCR). Several PCR Cyclers were used, based on machine availability:

- Eppendorf Mastercycler gradient and Eppendorf Mastercycler personal, (Eppendorf GmbH, Hamburg, D),
- PTC-100 ™ (MJ Research, Inc., Watertown, Mass., USA) and
- MiniCycler™ (MJ Research, Inc.), without top heating.

3.3.2 Chemicals for PCR Applications

For PCR applications, Millipore® water was used as diluting agent. The dNTPs (MBI Fermentas, St.Leon Rot, D) were mixed to obtain a 10-fold concentration. The buffer was mixed to obtain 5-fold concentration already containing $MgCl_2$. To prevent the mixture from evaporating, an oil layer was applied when using the MiniCycler ™ without top heating or when the reaction mixture was 20 μl or less (except for sequencing reactions). In this case, 20 μl mineral oil (Sigma Aldrich, Steinheim, D) was deposited atop the reaction mix.

A standard PCR reaction contained:

1x Reaction buffer	(50 mM KCl, 1.5 mM $MgCl_2$, 10 mM Tris, pH 9.0)
0.2 mM dNTPs	(Equal ratio blend of dATP, dGTP, dCTP, dTTP)
0.5 μM upper primer	
0.5 μM lower primer	
0.02 U/μl Taq polymerase	
20 ng Template DNA	
Millipore H_2O up to the final volume	

The elongation time was chosen depending on the length of the template DNA. For template DNA up to 1 kb, the program Universal 1.5 was chosen, and for DNA up to 2 kB (or if the length was unclear as for plaque PCR), Universal 2.0 was applied.

For sequencing reactions either an upper or a lower primer (0.2 mM) was used. The reaction was performed using 3 μl of the BigDye Terminator Cycle Sequencing Ready Reaction Mix 1.0 or 2.0 Kit (Perkin Elmer, Applied Biosystems Darmstadt, D) including: A-Dye Terminator labeled with dichloro[R6G]; C-Dye Terminator labeled with dichloro[ROX]; G-Dye Terminator labeled with dichloro[R110]; T-Dye Terminator labeled with dichloro[TAMRA]; Deoxynucleoside triphosphates (dATP, dCTP, dITP, dUTP); AmpliTaq® DNA Polymerase, FS, with thermally stable pyrophosphatase; magnesium chloride, and as Reaction buffer, Tris-HCl buffer, pH 9.0. The final volume was 10 μl, containing 20 ng template DNA.

3.3.3 PCR Programs

3 different PCR programs were used during this thesis. The parameters used for each program are listed below:

Table 3-14 PCR Program Parameters

	PCR Program		
Process steps	Universal 1.5	Universal 2.0	GATC Seq.
1. First denaturing step	2 min at 94°C	2 min at 94°C	1.5 min at 95°C
2. Elongation cycle	29 cycles	29 cycles	25 cycles
a. Denaturing	30 sec at 94°C	30 sec at 94°C	1.5 min at 95°C
b. Annealing	30 sec at 60°C	30 sec at 60°C	30 sec at 50°C
c. Elongation	1.5 min at 72°C	2 min at 72°C	4 min at 60°C
3. Last elongation	5 min at 72°C	5 min at 72°C	-
4. Cooling phase	∞ 4°C	∞ 4°C	∞ 2°C

3.3.4 Primers

The primers used during this thesis are described in the following two tables:

Table 3-15 List of Primers for General Use

Name	**Sequence (5´→ 3´)**	**Used for**
GT10F	GAG CAA GTT CAG CCT GG	λGT10
GT10R	CTT ATG AGT ATT TCT TCC AGG GTA	λGT10
pET T7n	ACC CCT CAA GAC CCG TTT AG	pET28a
pET term n	ACC CCT CAA GAC CCG TTT AG	pET28a
pDR195up	TCT CTT TCT TTC CTA TAA CAC CA	pDR195
pDR195rev	AAC CTT GAT TGG AGA CTT GAC	pDR195
PIG5-1	GTT GGT CTC CAT CAA AGC	*PIG5*
PIG5-2	CGA ATT TTT GAG CAC CAT CC	*PIG5*
PIG7-1	CGC AAC TCT ACA ACT CAA C	*PIG7*
PIG7-2	AAG TGC GAC CAA GGT AAG	*PIG7*
PIG7-3	TGC GGA AGT GTG ATT AGA AGG	*PIG7*
PIG7-10 5’	CAA ACC TTC GCT TAC TTT TCA C	*PIG7*
PIG7-11 3’	CGG GTA TGA TGA AAT CCC	*PIG7*
PIG9-2	CTG TGG CTT TCT TTG GTC	*PIG9*
PIG9-3	AGC TTG AAC AAC GCA GTC C	*PIG9*
PIG9-4	TGA GGA CTC TTG GCT ATG	*PIG9*
PIG23-1	GTG CTT GTT GTA TTC TTC C	*PIG23*

These primers were used for applications such as plaque PCR, amplification of cloned inserts and sequencing reactions. They were also applied to generate DIG-labeled probes for Northern and Southern blotting.

In order to affix a restriction site to cDNA fragments for cloning into vectors, special primers were designed. In addition to the inverse target sequence, recognition sites for chosen restriction enzymes with a number of protective bases flanking the 5’-end were manufactured. These extra nucleotides were required because some restriction enzymes cleave DNA poorly when their recognition sites are located at the ends of DNA fragments.

The above-mentioned special primers are described in the following table:

Table 3-16 List of Primers for Restriction Site Affixation to cDNA Fragments

Name	**Sequence (5´→ 3´)**	**Restriction site**	**Used for**
PIG5-Cleavage	GCT AGT ACA TAT GGC CAT CCC TAT GGT GGC CAA TC	*Nde*I	*PIG5*
PIG5-5’	GTC GCG GAT CCG TTA TTA GGC AAA TTG TGA C	*BamH*I	*PIG5*
PIG5-3’	GTC GAC GGA GCT CGC CCG TGG ACA ATC ATG	*Sac*I	*PIG5*
PIG7-Cleavage	CGA CCT CCA TAT GCA ACT AGT AGG CTC AGA TGT G	*Nde*I	*PIG7*
PIG9-Cleavage	GCT AGT CCA TAT GCC TCT GAT TAC AGC CGA AAC	*Nde*I	*PIG9*
PIG9Mitte	CAG ACC ACA TAT GGC CAC AGA GAA AGG CAC TAA G	*Nde*I	*PIG9*
PIG9-3’short	GCT GAC GGA GCT CCC ATC GTG CTT GCC TTG AG	*Sac*I	*PIG9*
PIG9-3’long	GTC GAC GGA GCT CCA AAT CAA GGC TTA GAC	*Sac*I	*PIG9*
PIG9-5’	GTC GCG GAT CCA TTA CAG CCG AAA CAG CAC	*BamH*I	*PIG9*
PIG14-5’	GTC GCG GAT CCG TAT CTC GCC TTG CTC CTC G	*BamH*I	*PIG14*
PIG14-3’	GTC GAC GGA GCT CGT CAT TTG GTG GAC TCA TAG	*Sac*I	*PIG14*
PIG23-5’	GAA TTC GAG CTC ATG GAT CAC CCG AAA CAA C	*Sac*I	*PIG23*
PIG23-3’	GGT GCT CGA GCA GGT TTG TTA GGC GTG GTC	*Xho*I	*PIG23*
PIG14-Cleavage	GAT CCT GCA TAT GAA CGC AAC TGG CCC AAG C	*Nde*I	*PIG14*
PIG15-Cleavage	GCC AGT CCA TAT GGT TAG ACC TTA CAA TCG TAT CCC	*Nde*I	*PIG15*
PIG23-Cleavage	GCC AGT CCA TAT GTT TTC TCA GGC TCC AGT GTT A	*Nde*I	*PIG23*

For reasons of convenience and economy, the abovementioned special primers were also applied for control PCRs, as well as to generate DIG-labeled probes for Northern and Southern blotting.

3.3.5 PCR Product Purification

PCR products were purified using the Qiaquick PCR Purification Kit (Qiagen GmbH, Hilden, D) or the E.Z.N.A® Cycle-Pure Kit (Peqlab, Biotechnologie GmbH, Erlangen, D). Both spin column purifications kits were used as specified in the manufacturers manuals. When using the Qiagen kit, the suggested optional washing step with 35% guanidine hydrochloride proved to be essential in removing all primers.

3.3.6 DIG Non-Radioactive Nucleic Acid Labeling

Digoxigenin-11-dUTP (DIG-dUTP) was incorporated into the DNA probe by Taq DNA polymerase during PCR in order to identify DNA or RNA on a Southern- respective Northern blot. Preferentially, the primers were chosen to replicate only cDNA sections and to exclude the λGT10 phage. In some cases (*PIG7*, *PIG14* and *PIG15*) it was unavoidable to include short sections of the phage in the DIG-labeled DNA probe in order to obtain a sufficiently large probe.

Table 3-17 10x DIG-dNTP Mix for DIG-Labeled Probes

Nucleotide	Final concentration	Manufacturer
dGTP	2.00 mM	Boehringer Mannheim, Mannheim, D
dCTP	2.00 mM	Boehringer Mannheim, Mannheim, D
dATP	2.00 mM	Boehringer Mannheim, Mannheim, D
dTTP	1.33 mM	Boehringer Mannheim, Mannheim, D
DIG-dUTP	0.67 mM	Boehringer Mannheim, Mannheim, D

Table 3-18 Primer Combinations for DIG-Labeled Probes

DIG-probe for	5'-primer	3'-primer	Template
PIG5	PIG5-5'	PIG5-3'	H322
PIG7	PIG7-10	PIG7-11	PDR195:*PIG7* (originating from H9)
PIG9	PIG9-5'	PIG9-3'	H62
PIG14	GT10R	GT10F	H1408
PIG15	GT10R	PIG15-2	H1309
PIG23	PIG23-5'	PIG23-3'	H1063

The template DNA for the DIG-labeled probes originated from the haustorial EST project.

3.4 Gels, Buffers and Sample Preparation

3.4.1 Agarose Gels and Buffers for DNA and RNA Analysis

DNA and RNA analysis by agarose gel electrophoresis was performed according to (Sharp et al. 1973). For DNA analysis, the following gels were used:

Table 3-19 Agarose gels for DNA Analysis

Application	Agarose conc.	Buffer	Running time
PCR product analysis	1.2 % (w/v)	TAE	20.0 V/cm for about 1 h
Plasmid analysis	0.8 % (w/v)	TAE	15.0 V/cm for about 1 h
Southern blotting	0.8 % (w/v)	TAE	1.7 V/cm for about 18 h

To prepare the agarose gels, agarose was melted in TAE buffer in a microwave oven, and then poured into gel moulds. The TAE buffer (see Table 3-20) was prepared as a 50-fold stock solution, stored under ambient conditions and diluted prior to use.

Table 3-20 TAE Buffer, (Maniatis et al. 1982)

Contents	Concentration
Tris base	40.0 mM
Na_2-EDTA	1.0 mM
Glacial acetic acid	5.7 %

For RNA analysis, the following gel was used:

Table 3-21 Agarose Gel for RNA Analysis (Maniatis et al. 1982)

Contents	Weight or Volume
Agarose	0.5 g
10 x RNA Running buffer	4.0 ml
H_2O (bi-distilled)	35.0 ml
37% Formaldehyde	2.0 ml

Before the formaldehyde solution was added, all other ingredients were microwaved for 1.5 min and cooled to approximately 55°C. The agarose gels for RNA applications were poured under a hood and allowed to solidify for approximately one hour.

For RNA analysis, the following Running buffer was used:

Table 3-22 10-fold RNA Running Buffer (Maniatis et al. 1982)

Contents	Final Concentration
MOPS	200 mM
1 M NaAcetate	50 mM
1 M EDTA pH 8.0	10 mM

The buffer was adjusted to pH 7.0 using NaOH and autoclaved at 121°C for 20 min. It displayed a light straw-yellow color after autoclaving and was always stored in darkness. If the color changed to a dark yellow, the buffer was discarded, because this color indicates an advanced state of oxidation (Maniatis et al. 1982).

For RNA analysis, the samples were prepared with an identical amount of RNA for each sample and then vacuum dried. The samples were resuspended in RNA Sample buffer (MBI Fermentas, St. Leon Rot, D) to a final volume of 10 μl, after which 1 μl ethidium bromide solution (1 mg/ml) was added. After applying the samples, the gel was run at 1 V/cm over night.

3.4.2 SDS PAGE Gels

SDS PAGE was performed according to Laemmli (1970). SDS PAGE gels for protein separation were run in a BioRad Protean II Mini Apparatus (BioRad, Munich, D). The following gels and buffers were used:

Table 3-23 12% Mini Gel

Contents	Separating Gel	Stacking Gel
H_2O	1.70 ml	1.50 ml
1.5 M Tris pH 8.8	1.25 ml	-
0.5 M Tris pH 6.8	-	0.63 ml
Acrylamide / Bisacrylamid (30 / 2,6)	2.00 ml	0.33 ml
10% SDS	50.0 µl	25.0 µl
TEMED	2.5 µl	2.5 µl
10% APS	25.0 µl	12.5 µl

Table 3-24 Sample Buffer (4-fold concentration) for SDS PAGE

Contents	Volume
Distilled water	4.0 ml
0.5 M Tris pH 6.8	1.0 ml
Glycerol (80%)	0.8 ml
SDS (10%)	1.6 ml
β-Mercaptoethanol	0.4 ml
Bromphenolblue	Tip of spatula

Table 3-25 Running Buffer for SDS PAGE

Contents	Concentration
Glycine	200.0 mM
Tris base	25.0 mM
SDS	0.1 %

Concentrated Sample buffer was added to the protein sample to create a one-fold Sample buffer solution. These probes were cooked at 95°C for 5 min, cooled on ice, shortly spun down and then directly applied on the 12% mini gel, which was suspended in the Running buffer.

3.5 Lambda Phage Cultivation

A fresh *E. coli* NM514 colony on a LB plate was transferred into 5 ml of LB medium plus 0.2% maltose and 10 mM $MgSO_4$ and placed on a shaker at 37°C for approximately 4 h until the OD_{600} was between 0.5 and 0.9. The lambda receptor on the bacterial cell's outer surface is part of the maltose-usage pathway (Boos 1976). Therefore, maltose was added to induce the receptor to high levels. In a next step, the bacteria were pelleted by centrifugation. After the medium had been poured off, the bacterial pellet was gently resuspended in 10 mM $MgSO_4$ to achieve a density of 0.5 at 680 nm. *E. coli* prepared in this way can be stored at 4°C and used for up to one week.

200 μl of the thus prepared *E. coli* NM514 was mixed with the desired volume of the not-amplified *Uromyces fabae* λGT10 library (Hahn and Mendgen 1997) and incubated (37°C, 150 rpm, 20 min). 3 ml of TopAgar were then blended with the phage/bacteria mixture. This combination was spread on pre-warmed NZY agar plates and incubated overnight at 37°C. During plating, the temperature did not exceed 48°C. The following day, the formation of plaques in the bacterial layer was controlled. Only plates with discrete plaques were used, in order to prevent a mixture of different plaques in subsequent assays.

The plaques were picked using shortened yellow pipette tips and transferred to 200 μl of SM buffer (see Table 3-26). To deactivate the bacteria, 2-3 drops of chloroform were added. If the mixture was to be stored at –70°C, 7 vol.% of DMSO were added.

2 μl of the plaques suspended in SM buffer were used as template for PCR amplification.

Table 3-26 SM-Buffer (Sambrook et al. 1989)

Contents	Concentration
Tris-HCl, pH 7.5	50.00 mM
Gelatin	0.01 %
$MgSO_4$	8.00 mM
NaCl	100.00 mM

3.6 Plasmid Vector Design

3.6.1 Cloning into the Plasmid Vector pET28a for Antigen Production

In order to obtain proteins of the *PIG*s to be studied for antibody production, these proteins were over-expressed using a suitable plasmid vector. The vector chosen was pET28a$^+$ (see chapter 3.2.3), which codes for a histidine tag, simplifying the isolation of the fusion proteins.

Starting with full-length cDNA clones, cDNA was amplified by PCR using specially designed primers (see Table 3-16), in order to affix an additional restriction site to the 5'-end and specially designed primers in order to affix an additional restriction site to the 3'-end or a λ-phage-specific primer at the 3'-end of the clone. The 5'-primers were also designed to exclude the potential secretion signals from the newly produced DNA to be cloned into the vector and to cause an in-frame insertion in the plasmid construct. The λ-phage-specific 3'-primers were chosen to include the cloning site (*Not*I, *EcoR*I) from the phage DNA. By including this cloning site or using the affixed restriction site at the 3'-end, the PCR product can be digested with the corresponding restriction enzymes as listed in the following table 3-27:

Table 3-28 List of Restriction Enzymes Used for Cloning into pET28a$^+$

PIG cDNA/ Name of Construct	Restriction 5'-end	Manufacturer	Restriction 3'-end	Manufacturer
PIG5	*Nde*I	NEB, Frankfurt/M, D	*Not*I	MBI, St.Leon-Rot, D
PIG5n	*Sac*I	Stratagene, Amsterdam, NL	*Bam*HI	Stratagene, Amsterdam, NL
PIG7	*Nde*I	NEB, Frankfurt/M, D	*Eco*RI	Stratagene, Amsterdam, NL
PIG9	*Nde*I	NEB, Frankfurt/M, D	*Eco*RI	Stratagene, Amsterdam, NL
PIG9s	*Sac*I	Stratagene, Amsterdam, NL	*Bam*HI	Stratagene, Amsterdam, NL
PIG9l	*Sac*I	Stratagene, Amsterdam, NL	*Bam*HI	Stratagene, Amsterdam, NL
PIG14	*Nde*I	NEB, Frankfurt/M, D	*Eco*RI	Stratagene, Amsterdam, NL
PIG15	*Nde*I	NEB, Frankfurt/M, D	*Eco*RI	Stratagene, Amsterdam, NL
PIG23	*Sac*I	Stratagene, Amsterdam, NL	*Xho*I	Boehringer, Mannheim, D

The plasmid vector pET28a$^+$ and the corresponding inserts were digested with the appropriate restriction enzymes according to the manufacturer's manual in order to generate compatible ends for cloning. Afterwards the enzymes were heat-inactivated for 15 min at 65°C. To remove the phosphate groups from the 5'-ends, the vector was dephoshorylated with shrimp alkaline phosphatase (USB™, Bad Homburg, D). This served to prevent self-ligation of insufficiently digested vector.

After the various above mentioned process steps, the inserts and vector were ligated. Ligation was performed according to Maniatis et al. (1982) using 10 ng plasmid, 30 ng insert DNA and 1 unit T4 ligase in Ligase buffer (USB™, Bad Homburg, D). The ligation mix was incubated over night in ice-cold water in a Styrofoam box. The resulting newly generated plasmid was transformed into *E. coli* TG1, multiplied and the modified plasmid vector isolated (see chapter 3.8.2.2).

The resulting modified vector pET28a:insert (see Figure 3-1) codes for the start codon and a histidine tag originating from the vector DNA, part of the PIG cDNA with the stop codon, some non-coding sequence and part of the phage DNA, including the restriction site.

Figure 3-1 Model of Plasmid Vector pET28a$^+$ with Insert for Antigen Production

3.6.2 Cloning into Yeast Shuttle Vector pDR195 for Heterologue Protein Production

The yeast shuttle vector pDR195 (see chapter 3.2.3) can be multiplied in *E. coli* and then be used to transform yeast cells. Because pDR195 offers only a few restriction sites in its cloning site, only one restriction site (*Not*I) was used to clone the insert into the yeast shuttle vector.

The plasmid Vector pDR195 and the corresponding inserts of the studied *PIG*s (*PIG5, PIG7, PIG14, PIG15*) were digested with the appropriate restriction enzyme *Not*I according to the manufacturer's manual (MBI Fermentas GmbH, St. Leon-Rot, D) in order to generate compatible ends for the cloning. In contrast to the pET28a cloning process, the inserts for cloning into pDR195 include the putative signal sequences of the studied PIGs. To remove the phosphate groups from the 5'-ends, the vector was treated with shrimp alkaline phosphatase (USB™, Bad Homburg, D). This served to prevent self-ligation of the vector. Afterwards the enzymes were heat inactivated for 15 min at 65°C.

After the various above mentioned process steps, the inserts and vector were ligated as described in 3.6.1. The resulting plasmid was transformed into *E. coli* TG1 and multiplied. After the right orientation of the insert was controlled by PCR, the plasmid was isolated and used to transform yeast cells (see chapter 3.7.2).

3.7 Transformation

3.7.1 Transformation of Bacterial Cells

The transformation of bacterial cells served either to multiply the plasmid vectors or to facilitate further over-expression of the fusion protein. Transformation was performed according to Mandel and Higa (1970). First, *E. coli* strains were grown in LB medium for about 3 h at 37°C and 275 rpm until an OD_{600} of 0.5 - 0.6 was reached. 5 ml aliquots of this culture were then put on ice before being centrifuged for 10 min at 4°C and 1'000 g. The supernatants were discarded, and each pellet was resuspended in 5 ml ice-cold 0.1 M $CaCl_2$ solution and put on ice for 30 min. The centrifuging step was repeated and the supernatants discarded. Each pellet was resuspended in 500 μl ice-cold 0.1 M $CaCl_2$ solution and put on ice for between 1 h and 18 h. In each case, 100 μl of the suspension were mixed with plasmid- or ligation make-up and incubated on ice for 30 min. The subsequent heat shock (90 sec, 42°C) was followed by another five minutes on ice. 900 μl LB medium were added to the transformation mixture and incubated for one hour at 37°C. Afterwards 100 μl respectively 300 μl of the mixture were plated on LB plates containing the required antibiotic. The plates were incubated at 37°C for 12 h to maximum 24 h, until bacterial colonies became visible. To verify whether the desired plasmid/insert construction was present, single colonies were then picked and controlled by PCR.

3.7.2 Transformation of Yeast Cells

Yeast transformation was used to over-express proteins in yeast, and was performed according to Elble (1992). Five to six colonies of the *Saccharomyces cerevisiae* strain SEY2102 (Emr et al. 1983) (grown on SC-maltose + uracil plate) were mixed with the pDR195 constructs before being mixed with 0.5 ml of PLATE-solution (see Table 3-29). The transformation make-up was incubated over night at room temperature (RT) before being subjected to heat shock (15 min, 42°C). The mixture was then briefly spun down and the pellet resuspended in 250 μl H_2O. 100 μl of each transformation make-up were plated on SC-Maltose-agar (without uracil) and incubated for about six days at 30°C, until yeast colonies became visible. To verify whether the desired plasmid/insert construction was present, single colonies were then picked and controlled by PCR.

Table 3-29 PLATE-Solution (Elble 1992)

Contents	Concentration
PEG 3350	40 %
Lithiumacetat	100 mM
Tris-HCl, pH 7.5	10 mM
EDTA	1 mM

3.8 Processing of Nucleic Acid

3.8.1 Nucleic Acid Quantification

In order to quantify DNA and RNA concentration and purity, the A_{260} value and the A_{260}/A_{280} ratio was determined using UV spectroscopy. Sample readings were made in quartz cuvettes, with the absorbance of diluted DNA or RNA samples being measured at 260 nm and 280 nm. The nucleic acid concentration was calculated using the Beer-Lambert law, which stipulates a linear relation between absorbance and concentration [$A = \log(I_0 / I) = \varepsilon * c * d$ with: A = extinction, I_0 = luminosity of entering light, I = luminosity of escaping light, c = molar concentration of the absorbing substance, d = layer thickness (mm), ε = molar absorption coefficient]. Using this equation, an A_{260} reading of 1.0 is equivalent to 50 μg/ml of DNA, respectively 40 μg/ml of RNA. For DNA (RNA) an A_{260}/A_{280} ratio of 1.8 (2.0) is indicative of highly purified nucleic acid. After purification of the DNA- respective RNA preparations, the A_{260}/A_{280} ratio was controlled. If the above-mentioned ratio was not achieved, the preparations were purified once again. If, after repeated purification, the desired purity could not be achieved, a lesser A_{260}/A_{280} ratio was accepted. The determined DNA, respective RNA concentration was always subsequently confirmed using an agarose gel.

3.8.2 DNA Isolation

3.8.2.1 DNA Isolation from Plants and Fungi

DNA isolation from *U. fabae* and *U. appendiculatus* germ tubes, bean leaves and fungal mycelia was carried out according to Hahn and Mendgen (1997). After the plant or fungal material was ground in liquid nitrogen using mortar, pestle and sea sand, it was suspended in 20 ml Extraction buffer and gently stirred for 30 min.

To remove proteins, the mixture was incubated with Proteinase K (150 μg/ml buffer) for 2 h at 50°C before being centrifuged at 3'000 g in an RC5C centrifuge (Sorvall® instruments, DU PONT, Wilmington DE, USA).

The supernatant was then extracted with 3/4 volume of phenol (equilibrated in Tris-HCl pH 7,5) spun down, and twice extracted with 3/4 volume phenol/chloroform (1:1). The resulting supernatant was precipitated with a 1/10 volume 3 M Na-acetate, pH 4.5 and an equal volume of 2-propanol for 1 h at 0°C, after which it was again centrifuged for 5 min at 12'000 g.

If the resulting pellet was still brownish, it was dissolved in TE buffer (see Table 3-31) and precipitated for a second time before being washed it in 75% ethanol and being dissolved again in 1 ml TE buffer. The DNA solution was then centrifuged for 2 h at 44'000 g in a RCM120 centrifuge (Sorvall® instruments, DU PONT, Wilmington DE, USA). The resulting supernatant contained the DNA and was further processed as required.

Table 3-30 DNA Extraction Buffer

Contents	**Concentration**
Tris-HCl, pH 8.0	10.0 mM
EDTA, pH 8.0	100.0 mM
SDS	0.5 %
RNase A	50.0 μg/ml

Table 3-31 TE Buffer

Contents	**Concentration**
Tris-HCl, pH 8.0	10 mM
EDTA, pH 8.0	1 mM

3.8.2.2 Plasmid Preparation

Plasmid preparation for transformation or sequencing was performed either by using kits supplied by various manufacturers (Plasmid Midi Kit, Qiagen, Hilden, D; Nucleobond AX 100 and AX 500, Macherey-Nagel, Düren, D; E.Z.N.A. Plasmid Miniprep Kit II, Peqlab, Erlangen, D); and according to the manufacturers manuals, or as described in (Birnboim and Doly 1979) for mini preparation. Cultures of *E. coli* TG1 with the desired plasmid were grown over night (LB-medium + antibiotic) and then centrifuged for five minutes at 10'000 g at RT. The supernatant was discarded and the pellet resuspended in 100 μl of Resuspending buffer (50 mM glucose; 25 mM Tris/HCl, pH 8.0; 10 mM EDTA, pH 8.0). An alkaline lysis was then performed using 200 μl of Lysis buffer (0.2 M NaOH, 1% SDS). This mixture was incubated at RT for 10 min, neutralized with 150 μl cold Neutralization buffer (3 M K-acetate, 5 M glacial acetic acid), and then centrifuged for five minutes at 10'000 g and 4°C. The supernatant was either used for transformation or mixed with 50 μl 1% RNAse solution and incubated for 10 min at RT. Afterwards, the supernatant was mixed with the same volume of phenol/chloroform (1:1) and centrifuged (2 min, 10'000 g, 4°C). The watery phase of the supernatant was then precipitated with a double volume of 100% ethanol. The resulting pellet, which contained the plasmid, was washed with 70% ethanol, dried and dissolved in H_2O.

3.8.3 DNA Sequencing

For DNA sequencing, 50-100 ng of column purified DNA (E.N.Z.A Cycle Pure Kit, Peqlab, Erlangen, D) was mixed with about 10 pmol of primer and 3 μl BigDye Kit 1.0 or 2.0 (Perkin Elmer, Überlingen, D). The mixture was diluted with Milli Pore ® water up to a final volume of 10 μl. Single-strand amplification was performed using the GATC-Seq program (3.3.3). The resulting product was sent to GATC GmbH, Konstanz, D, where it was sequenced using an ABI 3700 sequencer. The resulting sequence files were analyzed using the software "EditView".

3.8.4 RNA Isolation

3.8.4.1 Preparation and Cleansing of Materials

In preparation for RNA isolation, all materials, solutions and water were either baked for 4 h at 180°C or treated with DEPC (Diethylpyrocarbonate). Therefore, solutions, water, pipet tips, reaction tubes and other autoclaveable material which could not be exposed to dry heat were incubated overnight in 0.1% DEPC and then autoclaved. DEPC has the function of reacting with the histidine residues of proteins and inactivating RNases. However, DEPC can also react with RNA, so it needs to be removed by heat treatment before the materials are used for RNA isolation. Autoclaving inactivates DEPC via hydrolysis, whereby CO_2 and ethanol are released as reaction by-products.

Because compounds containing primary amine groups such as Tris (2-Amino-2-hydroxymethyl-1,3-propanediol) also react with DEPC, all solution containing Tris were prepared with autoclaved DEPC-treated water. For this purpose, water and DEPC were mixed to a final concentration of 0.1% with DEPC stirred over night and autoclaved. The thus treated water was used to dissolve Tris up to the desired concentration.

3.8.4.2 RNA Isolation from U. fabae Uredospores

RNA isolation from *U. fabae* uredospores was performed according to Hahn and Mendgen (1997). Uredospores were ground in liquid nitrogen using a mortar, pestle and sea sand for about 45 min. Portions of about 0.1 g each were mixed with a 20-fold volume of PeqGold RNAPure™ solution (Peqlab Biotechnologie GmbH, Erlangen, D). The solutions were allowed to incubate at RT for about 5 min before 80 μl chloroform were added and mixed vigorously for 15 sec. The suspensions were then incubated for another 5 min at RT and centrifuged for 5 min at 10’000 g. The clear aqueous phase was carefully removed without interphase and mixed into 200 μl 2-propanol before being incubated for 15 min at RT. In order to pellet the RNA, the suspension was then centrifuged for 10 min at 10’000 g and 4°C. The whitish pellet was then washed twice with 75% ethanol, dissolved in water and the RNA concentration determined with a photometer.

3.8.4.3 RNA Isolation U. fabae Germ Tubes

RNA isolation from *U. fabae* germ tubes was performed according to Hahn and Mendgen (1997). Germinated uredospores (3.1.4) were ground in liquid nitrogen using a mortar, pestle and sea sand for about 45 min. Portions of about 0.1 g were each mixed with at least the 20-fold volume of PeqGold RNAPure™ solution (Peqlab Biotechnologie GmbH, Erlangen, Germany). Further processing was performed as described in chapter 3.8.4.2.

3.8.4.4 RNA Isolation from Haustoria

RNA isolation was performed according to Hahn and Mendgen (1997). The frozen haustorial pellets were poured into a mortar with at least the 20-fold volume of PeqGold RNAPure™ solution (Peqlab Biotechnologie GmbH, Erlangen, D). After grinding the pellets for several minutes, the mixture was transferred to a microfuge tube. The isolation was performed analog to the RNA isolation from uredospores described in chapter 3.8.4.2.

3.8.4.5 RNA Isolation from Infected and Uninfected Vicia faba Leaves

RNA Isolation from infected and uninfected *Vicia faba* leaves was performed according to Hahn and Mendgen (1997). About 15 g *Vicia faba* leaves were ground in liquid nitrogen using a mortar and pestle. Portions (0.3 g) of the frozen powder was successively mixed with 1 ml of PeqGold RNAPure™ solution (Peqlab Biotechnologie GmbH, Erlangen, D). The mixture was vortexed for 15 sec and then allowed to incubate at RT for about one minute. It was then centrifuged for 5 minutes at 1’300 g. The resulting supernatant was added to 200 μl chloroform and mixed vigorously for 15 sec. The suspension was incubated for three minutes at RT and then centrifuged for five minutes at 1’300 g. The clear aqueous phase was carefully removed without interphase and mixed with 0.5 volume of 2-propanol and 0.25 volume of High-salt buffer (1.2 M Na-citrate, 0.8 M NaCl) before being incubated for 10 min at RT. After incubation, the aqueous phase was then centrifuged for ten minutes at 1’300g and 4°C. The resulting whitish pellet was washed two times with 75% ethanol, dissolved in water, and the RNA concentration determined with a photometer.

3.8.4.6 RNA Isolation from U. fabae in vitro Structures

RNA Isolation from *U. fabae in vitro* structures was performed according Hahn and Mendgen (1997). *In vitro* structures were prepared as described in section 3.1.4. After addition of 2.5 ml GTC-solution (see Table 3-32) per foil, the germinated uredospores were harvested using a scraper. The solution containing the *in vitro* structures was transferred to a glass potter and homogenized. Subsequently, the homogenate was mixed with 0.5 volume of phenol/chloroform/isoamylalcohol in an Erlenmeyer flask. All steps of this process were performed in a cold storage room with pre-cooled tools, and all liquids were kept on ice. The extracts were also kept on ice at all times. The emulsion was then centrifuged for 10 min at 19’700 g and 4°C. The hydrophilic upper phase was mixed with 0.1 volume of 2 M Na-acetate, pH 4.0 and 0.5 volume of phenol/chloroform/isoamylalcohol (see Table 3-34) and centrifuged again. These steps were performed on ice under a fume hood. This procedure was repeated once, after which the RNA in the hydrophilic upper phase was precipitated over night at –20°C with one volume of 2-propanol and centrifuged for 10 min at 19’700 g and 4°C. The still moist pellet was suspended in 0.1 volumes of SET buffer (see Table 3-33) (corresponding to the starting volume of hydrophilic upper phase) and mixed with 1 volume of chloroform/ isoamylalcohol (see Table 3-35). This mixture was centrifuged for 10 min at 19’700 g and 4°C. The hydrophilic upper phase was removed and the RNA precipitated with lithium-chloride (final concentration 2 M). The RNA pellet was washed with 70% ethanol and suspended in Millipore® water before being put into storage at –80°C.

Table 3-32 Guanodin-Thiocyanat-Citrat-solution (GTC-solution)

Contents	Concentration	Supplier
Guanidinium-Thiocyanat	4 M	Difco Laboratories, Detroit, MI, USA
Na-Citrat	25 mM	Select Yeast Extract Sigma Aldrich Chemie GmbH, Steinheim, D
Sarcosyl (sodium laurel sarcosinate)	5 %	C. Roth & Co. GmbH, Karlsruhe, D
β-Mercaptoethanol	100 mM	C. Roth & Co. GmbH, Karlsruhe, D

Table 3-33 SET Buffer

Contents	Concentration
EDTA	1.0 mM
SDS	0.1 %
Tris-HCl, pH 7.5	10.0 mM

Table 3-34 Phenol/Chloroform/Isoamylalcohol

Contents	Concentration
Phenol pH 4.0	50 %
Chloroform	48 %
Isoamylalcohol	2 %

Table 3-35 Chloroform/Isoamylalcohol

Contents	Concentration
Chloroform	96 %
Isoamylalcohol	4 %

3.8.5 Nucleic Acid Blotting and Detection (Southern and Northern Blotting)

3.8.5.1 DNA Blotting and Detection (Southern)

During this thesis, Southern blotting served to analyze the number of gene copies of the genes to be studied and to detect homologous genes of *PIG7* in other fungi. DNA blotting and detection was performed according to Southern (1975). Overnight, 1 µg DNA per sample was digested with the desired enzymes (*Bam*HI, *Eco*RI, *Eco*RV, *Hind*III and *Mun*I) in corresponding buffers. 0.1 volume of 3 M Na-acetate pH 5.5 and 2.5 volumes of 100% ethanol were added to the mixtures and the DNA precipitated for 30 min at -20°C. The DNA was spun down for 15 min at 14'000 rpm, the supernatant discarded, and the pellets washed with 70% ethanol. The dried pellets were then resuspended in 20 µl TE buffer and 4 µl of 6–fold concentrated Loading buffer added to the suspension.

The gel was run as described in section 3.4.1 and incubated in freshly prepared ethidium bromide solution (1 μg/ml) for about 15 min in order to make the DNA visible under UV illumination. To remove the ethidium bromide, the gel was then washed in a water bath on a shaker for 15 min at RT and low speed. In order to hydrolyze larger DNA fragments, the gel was incubated and shaken in 400 ml 0.25 M HCl for 10 min at RT. The HCl was then removed by aspiration, the gel rinsed briefly in H_2O and washed twice for 15 min each at RT on a shaker in 500 ml 0.5 M NaOH, 1.5 M NaCl. Afterwards the gel was rinsed briefly in H_2O, followed by washing twice for 15 min each with 500 ml 0.5 M Tris/HCl pH 7.5, 3 M NaCl at RT on a shaker and again briefly rinsed in H_2O.

Transfer of DNA: For an upward transfer with 20-fold SSC solution (see Table 3-50), the DNA gel was placed under an Amersham Hybond-N$^+$ nylon membrane (Amersham Bioscience Europe GmbH, Freiburg, D). The membrane had been previously soaked in bi-distilled water followed by a short incubation in two-fold concentrated SSC. The DNA was transferred over night under ambient conditions (about 21°C).

To verify the DNA transfer, the gel was removed from the membrane and stained in ethidium bromide solution for about 20 min, washed briefly in H_2O, and illuminated under an UV-transiluminator. The membrane was then dried between two pieces of Whatman 3MM paper, and baked at 120°C for 30 min.

Hybridization of DNA: Hybridization was carried out in a screw-capped tube in a preheated oven at 68°C. The dry membrane was transferred to the glass tube, 40 ml pre-hybridization solution (see Table 3-37) was added and tube was incubated rotating for two hours. The hybridization with a DIG-labeled probe (10 ml hybridization solution (

Table 3-38)) took place overnight under the same conditions as pre-hybridization. Afterwards the hybridization solution was collected und stored at –20°C for reuse.

Developing the blot: The blot was washed once in a glass tray with 240 ml Wash buffer I (see Table 3-39) for 10 min at RT, before being washed 4 times in Wash buffer I for 20 min each at hybridization temperature (68°C). The blot was then incubated in a glass tray on a shaker with 240 ml Maleic acid buffer (MAB) (see Table 3-46) for 5 minutes at RT, before being incubated in 185 ml Blocking buffer (see Table 3-52) for 60 min at RT. Then the membrane was transferred to a large rotating drum and incubated with 15 ml of anti-DIG antibody-solution (Roche Diagnostics GmbH, Mannheim, D) for 30 min at RT.

To remove unbound antibodies, the membrane was washed on a shaker four times with 240 ml MAB for 10 min each at RT, followed by washing with 190 ml Substrate buffer (see Table 3-42) at RT for 10 min. The membrane was then transferred to a plastic bag (open on three sides, DNA side facing up) where 3 ml of CSPD-solution (see Table 3-43) were added. The top of the plastic bag was moved up and down for 5 min to spread the CSPD-solution evenly over the membrane. After removing the CSPD-solution and sealing the bag, the membrane was incubated for 30 min at 37°C. This served to attain the plateau of the chemoluminescence and to stabilize the time dependency of the chemoluminescence. The membrane (wrapped in the plastic bag) was then placed in an auto-radiography cassette, where a film was added (Amersham Hyperfilm ECL, Amersham Biosciences Europe GmbH, Freiburg, D). Exposure times varied between 30 - 60 min.

Stripping the blot: For repeated hybridization with different probes or with the same probe at different temperatures, the probe, together with antibodies and CSPD, had to be removed from the membrane without damaging the bound DNA.

To achieve this, the blot was transferred to a glass tray on a shaker, where it was rinsed twice for 15 min in 250 ml stripping solution (see Table 3-51) at 42°C, followed by four washings of 10 min each in 250 ml 2 x SSC at RT. The blot was then either dried and stored for later use, or used directly for another round of hybridization (starting with the pre-hybridization).

3.8.5.2 RNA Blotting and Detection (Northern)

During this thesis, Northern blotting was used to characterize the expression profile of the genes to be studied. RNA transfer was performed according to Chomczynski (1992). If not noted otherwise, all steps were performed at RT.

For an alkaline downward transfer, the RNA gel was placed on a Hybond-N^+ nylon membrane (Boehringer Mannheim, Mannheim, D). The RNA was cross-linked to the membrane at 80°C for 20 min. The membrane was incubated by rolling in a glass tube with 40 ml pre-hybridization solution (see Table 3-37) for 2 h at 68°C, before being incubated overnight with 10 ml hybridization solution (see Table 3-38) at 68°C This allowed the DIG-labeled DNA probe to bind to mRNA on the membrane. Afterwards the membrane was washed 3 times for 20 min with 60 ml Wash buffer I at 68°C.

Prior to detection of the DIG-labeled DNA probe, the membrane was transferred to a shaker and washed for 5 min with 200 ml MAB I (see Table 3-46) and then incubated for 60 min with 190 ml Blocking buffer (freshly made from stock, see Table 3-52). Then the membrane was transferred to a large rotating drum together with 10 ml of Anti-DIG Antibody-solution (Roche, Diagnostics GmbH, Mannheim, D) where it was incubated under rotation for 30 min. The alkaline phosphatase-conjugated anti-DIG antibody served to detect the DIG-labeled probe. Afterwards, the membrane was transferred to a shaker and washed three times for 10 min each with 200 ml MAB I.

Before applying 3 ml CSPD-solution, the membrane was incubated for 5 min with 190 ml Substrate buffer. After removing the excess CSPD solution, the membrane was incubated in the sealed plastic bag for 30 min at 37°C. This served to attain the plateau of the chemoluminescence faster and to stabilize the time-dependency of the chemoluminescence. The plastic bag with the membrane was then transferred to an autoradiography cassette and exposed to a film (Amersham Hyperfilm ECL, Amersham Biosciences Europe GmbH, Freiburg, D) for between 10 and 60 min.

3.8.5.3 *Buffers and Solutions used for DNA and RNA Blotting and Detection*

Table 3-36 RNA Transfer Buffer

Contents	Final Concentration
NaCl	3 M
NaOH	8 mM
Na-laurylsarcosine	2 mM

Table 3-37 Prehybridization Solution for Southern and Northern Blots

Contents (Stock Solution)	Final Concentration
Na-Pi, pH 7.2	250.0 mM
Na-EDTA pH 8.0	1.0 mM
Blocking reagent stock	0.5 %
SDS	17.5 %

Table 3-38 Hybridization Solution for Southern and Northern Blots

Contents	Volume
DIG-labeled DNA probe (0.05 μg/μl)	2.5 µl
Prehybridization solution	10.0 ml

This solution can be stored at -20°C and used repeatedly.

Table 3-39 Wash Buffer I for Southern and Northern Blots

Contents (Stock Solution)	Final Concentration
Na-Pi pH 7.2	20 mM
Na-EDTA pH 8.0	1 mM
25% SDS	1 %

The Wash buffer I components were added to water in the above order to prevent precipitation.

Table 3-40 Antibody-solution for Northern Blots

Contents	Volume
Alkaline phosphatase conjugated anti-DIG-antibody (Boehringer Mannheim, Mannheim, D)	0.67 µl
Blocking buffer for Northern Blots	10 ml

The antibody-solution was always freshly made directly prior to usage.

Table 3-41 Antibody-Solution for Southern Blots

Contents	Volume
Alkaline phosphatase conjugated anti-DIG-antibody (Boehringer Mannheim, Mannheim, D)	1 µl
Blocking buffer for Northern Blots	15 ml

The antibody-solution was always freshly made directly prior to usage.

Table 3-42 Substrate Buffer for Southern and Northern Blots

Contents (Stock Solution)	Final Concentration
Tris-HCl pH 9.5	0.1 M
NaCl	0.1 M

Table 3-43 CSPD-Solution

Contents	Volume
CSPD stock solution (Roche Diagnostics GmbH, Mannheim, D)	100 µl
Substrate buffer	10 ml

Table 3-44 1 M Na-Pi Buffer, pH 7.2 (autoclaved, stored at RT)

Contents	Volume
1M Na_2HPO_4	288 ml
1M NaH_2PO_4	112 ml

The pH was adjusted by adding more Na_2HPO_4 (alkaline) or NaH_2PO_4 (acidic).

Table 3-45 1 M Na-Pi Buffer, pH 6.8 (autoclaved, stored at RT)

Contents	Volume
1M Na_2HPO_4	196 ml
1M NaH_2PO_4	204 ml

The pH was adjusted by adding more Na_2HPO_4 (alkaline) or NaH_2PO_4 (acidic).

Table 3-46 Maleic Acid Buffer I, pH 8.0 (MAB I)

Contents	Final Concentration
Maleic acid	0.1 M
NaCl	3.0 M
Tween20	0.3 %

The pH was adjusted to 8.0 using NaOH(s), or 10 M NaOH.

Table 3-47 Maleic Acid Buffer II, pH 7.5 (MAB II)

Contents	Final Concentration
Maleic acid	0.10 M
NaCl	0.15 M

The buffer was adjusted to pH 7.5 using 10 M NaOH.

Table 3-48 TE buffer, pH 8.0

Contents (stock solution)	Final Concentration
Tris/HCl pH 8.0	10 mM
Na-EDTA pH 8.0	1 mM

The buffer was autoclaved and stored at RT.

Table 3-49 1 x TAE

Contents	Final Concentration
Tris/acetate	40 mM
EDTA pH 8.0	1 mM

Table 3-50 20 x SSC

Contents	Final Concentration
NaCl	3.0 M
Na-citrate	0.3 M

The pH was adjusted to pH 7.0 using HCl_{conc}

Table 3-51 Stripping Solution for Southern Blots

Contents	Final concentration
NaOH	0.2 M
SDS	0.1 %

To reuse the membranes for Southern Blotting, the DIG-labeled probe can be removed from the membrane with the stripping solution.

Table 3-52 10% Blocking Reagent Stock for Southern and Northern Blots

Contents	Weight or Volume
Blocking reagent # 1096176 (Boehringer Mannheim, Mannheim, D)	10 g
MAB-II (see Table 3-47)	100 ml

The blocking reagent stock solution was carefully heated to 60°C in a water bath and dissolved as completely as possible, autoclaved and stored at 4°C. Blocking Buffer was always freshly made from the stock solution by mixing 190 MAB II with 10 ml 10% blocking reagent stock solution.

3.9 Protein Processing

3.9.1 His-tagged Protein Expression and Isolation

During this thesis, His-tagged proteins were used for antibody production and as positive control for Western blot. Expression and purification of His-tagged proteins was performed according to the Qiagen protein purification manual (QIAexpressionist 2000) under denaturing conditions. 2.5 ml of an overnight culture of *E. coli* BL21 *plys E* containing the desired plasmid (pET28a^{+}:insert) and suspended in LB medium containing 50 μg/ml kanamycin was used as inoculum for 250 ml LB medium also containing 50 μg/ml kanamycin. Under vigorous shaking at 37°C, the culture was allowed to grow until OD_{600} was 0.5 - 0.7. A 1 ml sample of the culture was taken before IPTG was added to attain a final concentration of 1 mM. The cultures were again grown for 2 to 4 hours with samples being taken hourly (same amount of cells as in the first sample). The cells were then harvested by centrifugation for 15 min at 4’000 g. The supernatant was removed and the pellet lysed in 5 ml Buffer B (for Buffers B–E, see Table 3-53). For this, the cells were stirred for one hour at RT. The resulting lysate was centrifuged for 20 min at 10’000 g and the supernatant incubated on a roller for one hour with five ml Ni-NTA resin (Qiagen, Hilden, D) under ambient conditions. Subsequently, the mixture was filled in a column and the flow-through collected. The filled column was washed twice with 4 ml Buffer C and the wash fractions collected. The recombinant protein was eluted four times each with 0.5 ml Buffer D, followed by four elutions with 0.5 ml Buffer E. The resulting fractions were collected. Finally, the flow through, wash fractions and eluted protein fractions were analyzed using a 12% SDS-PAGE gel.

Table 3-53 Buffer B, C, D, E

Contents	Final Concentration
NaH_2PO_4	100.0 mM
Tris-HCl	10.0 mM
Urea	8.0 M

Buffer B:	pH adjusted to 8.0 using HCl
Buffer C:	pH adjusted to 6.3 using HCl
Buffer D:	pH adjusted to 5.9 using HCl
Buffer E:	pH adjusted to 4.5 using HCl

3.9.2 Preparation of Culture Filtrate from Yeast Cells

For infiltration experiments and Western blot (see 3.9.4) culture filtrate from transformed yeast cells was required. Yeast cells transformed with a yeast shuttle vector (see chapter 3.7.2) were grown for three to four days until OD_{600} reached about 12. The cells were spun down and the supernatant was sterilized by filtration through NALGENE® 0.2 μm syringe filters (Nunc GmbH & Co. KG, Wiesbaden, D). For volumes of 0.5 ml and 2 ml, the filtrate was concentrated using an Amicon® Microcon® or Amicon® Centricon® Centrifugal Filter Device (Millipore, Billerica, MA, USA) respectively. The pH value was adjusted to 5.0 (Felle and Hanstein 2002) using 0.1 M Tris buffer, pH 7.4.

3.9.3 Determination of Protein Concentration

Numerous experiments carried out during this thesis required the application of a defined amount of protein. Determination of the protein concentration was performed using the BioRad protein assay (BioRad Laboratories, Munich, D). This is a simple colorimetric assay for measuring total protein concentration, based on the Bradford dye-binding procedure (Bradford 1976). This assay measures the color change at the OD_{595} of Coomassie Brilliant Blue G-250 dye when it binds to a protein (primarily to basic and aromatic amino acid residues) and quantifies most proteins and polypeptides with molecular mass > 3'000 – 5'000 kDa. As applied during this thesis, the protein samples or dilutions thereof were topped up to 800 μl with water, after which 200 μl of the BioRad stock solution were added. The mixture was then incubated at RT between 5 and 60 min, before the $OD_{595}$was determined.

The reference curve used to calibrate the OD measurements was obtained using BSA as the standard protein in the following concentrations: 100 μg/ml, 50 μg/ml and 0 μg/ml. If the protein concentration of crude cell extract needed to be determined, cells were boiled for 15 min in 0.5 M NaOH. NaOH increases the OD_{595}; therefore, NaOH was also added to the samples for the reference curve.

3.9.4 Immunoblot Analysis of Proteins (Western Blot)

Within the framework of protein expression studies and in order to verify antibody purification, immunoblot analyses of proteins were performed according to Burnette (1981). This analysis method is known as Western blotting, and was performed using a MiniProteanII setup (BioRad, Munich, D). First, a varying quantity of protein was heated to 95°C for 5 min in Sample buffer and applied to a 12% SDS PAGE gel. The gel was run for 60 min at 200 V.

The separated proteins were transferred to a PVDF membrane (Millipore, Bedford, MA, USA) in pre-cooled Trans Blot Buffer (TBB) (see Table 3-54), for 1 h hour at 100 V with an additional ice-cooling unit to prevent methanol combustion. Afterwards, the PVDF membrane was transferred to Post Blot Buffer (PBB) (see Table 3-55) and washed twice for 5 min each, before being blocked for at least one hour at RT (or overnight at 4°C) in 50 ml PBB with 5% fat- reduced milk powder. Following blocking, the membrane was incubated for one hour in 50 ml primary antibody incubation solution (PBB, 0,05% NP40, 5% Milk powder + antibody solution in the dilution 1/10'000 for antibodies against PIG5p, PIG7p and PIG14p, respectively 1/5'000 for antibodies against PIG15p). The membrane was washed three times for 10 min in PBB + 0.05% NP40. Then the membrane was incubated for 30 min in PBB + 5% milk powder + 1/1'000 secondary antibody (Goat Anti Guinea Pig IgG Peroxidase Conjugate, Sigma, St. Louis, MI, USA for anti-bodies against PIG7p and PIG14p, and Goat Anti Rabbit IgG Peroxidase Conjugate, Sigma, St. Louis, MI, USA for antibodies against PIG5p, and PIG15p, respectively). Afterwards the membrane was washed twice for 10 min in PBB + 0.05% NP40 and twice in 10 mM Tris pH 7.4.

To develop the blot, ECL Western Blot detection reagents (Amersham Biosciences Europe GmbH, Freiburg, D) was applied according to the manufacturers manual. The blot was exposed to the film (Hyperfilm ™ ECL ™, Amersham Biosciences Europe GmbH, Freiburg, D) between 1 min. and 1 h.

Table 3-54 Trans Blot Buffer (TBB)

Contents	Final Concentration
Tris base	35 mM
Glycine	270 mM
Methanol	20 %

Table 3-55 Post Blot Buffer, pH 7.4 (PBB)

Contents	Final Concentration
NaCl	0.9 %
Tris-HCl	10.0 mM

Table 3-56 Post Blot Buffer + NP 40, pH7.4 (PBB+NP40)

Contents	Final Concentration
Tris-HCl	10 mM
NaCl	0.90 %
Nonidet P40 (Octylphenol ethylene oxide condensate), LBB-Produkter AB, Brommer, S	0.05 %

3.9.5 Immuno-Fluorescence Microscopy of Proteins in Leaf Cross-Sections

Immuno-fluorescence microscopy of proteins in leaf cross-sections was performed by Prof. Dr. K.W. Mendgen and Dipl. Ing (FH) H. Vahlenkamp. Infected *Vicia faba* "Con amore" leaves were incubated for 1 h at RT with ethanol/acetic acid (3/1) and thereafter washed three times with 100% ethanol. These steps were followed by a substitution procedure using a resin/ethanol mixture. The resin fraction (see Table 3-57) was enhanced from 25% up to 100% in steps of 25%; the duration of each incubation step was at least 1 h, with ambient air temperature of 4°C. The resin was allowed to polymerize for 4 days under long wave UV illumination at 4°C, after which the embedded leaf tissue was cut in 0.5 μm slices and placed onto microscope slides coated with Biobond (British BioCell Ltd., Cardiff, UK).

The affixed leaf cross-sections were allowed to dry and then washed for 1 min in acetone in order to remove the resin, followed by 10 min incubation in blocking reagent (5% BSA in TBS (see Table 3-58) followed by 2 h incubation in antibody serum (Anti-PIG 1:100 in Primary Antibody Incubation buffer (see Table 3-59). Thereafter the leaf cross-sections were washed 6 times for 5 min each in Incubation buffer, before they were incubated for 1 h at 20°C with the secondary antibody (Cyanin-3-conjugated goat anti-rabbit antibodies resp. anti-guinea-pig antibodies, Rockland, Gilbertsville PA, USA, 1:400 in TBS). Excess antibody was removed by washing the cross-sections 3 times for 10 min each in Incubation buffer. The thus treated tissue was then embedded in SlowFade (Molecular Probes Inc., Eugene, OR, USA). Visualizing took place with a fluorescence microscope (Axioplan 2, Zeiss, Oberkochen, D) using the following filters: excitation: 545 nm; splitter: 565 nm; emission: 610 nm.

Table 3-57 Resin (Aldrich Chemical Corp., Milwaukee, WI, USA)

Contents	Final Concentration
Butyl-methacrylate	75.5 %
Methyl-methacrylate	25.0 %
Benzoin-ethyl-ether	0.5 %
DTT	10.0 mM

Table 3-58 TBS, pH 7.4

Contents	Final Concentration
Tris-HCl	10 mM
NaCl	150 mM

Table 3-59 Primary Antibody Incubation Buffer

Contents	Volume
Incubation Buffer (TBS pH 7,4 + 1% BSA-C),	150 μl
Block Buffer (5% BSA, 5%, Goat-serum in 10 mM TBS)	50 μl
Yeast Buffer (Yeast in 75% TBS)	100 μl
Bean Leaf Buffer (thin cut bean leaves in TBS 1/1, autoclaved)	100 μl

After mixing all ingredients, the buffer mixture was spun down and the supernatant used as Primary Antibody Incubation buffer.

3.10 Antibody Production

3.10.1 Immunization of Rabbits and Guinea Pigs

The pre-immune serum of the animals to be immunized was tested for immune reactions on infected leaves (*Uromyces fabae* on *Vicia faba*) by Prof. Dr. K. Mendgen and Dipl. Ing (FH) H. Vahlenkamp using immuno-fluorescent microscopy. If no immune response was detected, the animals were used to produce antibodies. The animals were kept, cared for, and treated at the Tierforschungsanlage (TFA), University of Konstanz, Konstanz D. For the immunization of rabbits, 500 μg / 500 μl of antigen was mixed with the same volume of Freunds adjuvant (first boost complete Freunds adjuvant, followed by boosts with incomplete Freunds adjuvant). For the immunization of guinea pigs, 250 μg / 250 μl antigen was mixed with the same volume of RIBI Adjuvant System, (Corixa Corporation, Seattle, WA, USA). The animals were injected with this mixture once every two weeks over a period of 8 weeks (4 injections in total). The serum was checked using immuno-fluorescence microscopy for a sufficient immune response and if necessary, the immunization procedure was repeated.

Table 3-60 Animals used for Antibody Production

PIGp	Animal No.
PIG5p	**K767** (rabbit)
PIG7p	**MS2 (746)** (guinea pig)
PIG14p	**MS130** (guinea pig)
PIG15p	**K750** (rabbit)

Only those animals from which useable antibodies were obtained are listed above.

3.10.2 Purification of Antibodies

Purification of antibodies was performed according to Voegele et al. (2001). To remove vector-encoded protein sections, antibodies were purified by negative adsorption against an *U. fabae* protein also cloned in the vector pET28a (PIG7p for PIG5p, PIG14p, PIG15p and PIG5p for PIG7p). Then, in order to select only the required antibodies, the antibodies were purified by positive adsorption against the antigen.

For both types of adsorption, freeze-dried CNBr-activated Sepahrose 4B (Amersham Bioscience Europe GmbH, Freiburg, D) was washed with 1 mM HCl according to the manual to remove additives, whereby 1 g CNBr-activated Sepharose 4B was used per column. Then, at least 1.4 mg of each fusion protein (arbitrary protein and antigen) was coupled to its matrix for 1 h at RT or overnight at 4°C. Excess ligand was then washed away with Coupling buffer (see Table 3-61) and the remaining active groups were blocked for 2 h at RT with Glycine buffer I, (see Table 3-61). Non-covalently bound components were removed using high-salt buffers during three washing cycles of alternating pH. Each cycle consisted of a cleansing with Acetate buffer, followed by a cleansing with Coupling buffer (see Table 3-61).

Both mixtures of CNBr-activated Sepahrose 4B and Coupling buffer were filled into their own plastic column (BioRad, Poly-Prep Chromatography Columns, Munich, D) and the Coupling buffer was allowed to drain. 5 ml of the blood serum was first applied and reapplied three times onto the column that did not contain the antigen (negative adsorption). Subsequently, the flow-through was allowed to trickle through the other, antigen-containing column (positive adsorption) and re-applied three times.

For elution of the coupled antibodies, 3.6 ml Glycine buffer II was applied to the column and the flow-through was collected in a tube containing 500 μl 1.0 M Tris buffer pH 8.0. The Tris buffer served to buffer the acidic Glycine buffer at a sufficiently high pH level to prevent denaturation of the protein.

The success of the purification was verified by Western blotting (see chapter 3.9.4).

For regeneration of the columns, these were completely filled with Regeneration buffer I and allowed to drain. Subsequently, the columns were filled to the top with Regeneration buffer II and allowed to drain. Thereafter, the columns were washed twice with Coupling buffer. The columns were then stored at 4°C in Coupling buffer containing Na-Azid and reused as required.

Table 3-61 Buffers for Antibody Purification

Buffer name	Contents	pH
Coupling buffer	0.10 M $NaHCO_3$, 0.50 M NaCl	Adjusted to pH 8.3 with NaOH
Acetate buffer	0.10 M NaAcetate 0.50 M NaCl	Adjusted to pH 4.0 with glacial acetic acid
Glycine buffer I	0.20 M Glycine	Adjusted to pH 8.0 with NaOH/HCl
Glycine buffer II	0.05 M Glycine	Adjusted to pH 2.2 with HCl
Regeneration buffer I	0.10 M $NaHCO_3$	Adjusted to pH 8.3 with NaOH
Regeneration buffer II	0.10 M NaAcetate 0.50 M NaCl	Adjusted to pH 2.2 with glacial acetic acid

3.11 Data Handling and Analysis

3.11.1 Description of Hardware

Most of the data handling and analysis was carried out using an Apple Power Macintosh G3, 255 MHz, operating under system Mac OS 8.6, and an Apple Macintosh iBook G3, 500 MHz, operating under systems Mac OS 9.2.1 and Mac OS X.

3.11.2 Software Employed

Table 3-62 Software Employed

Product Name/URL	Supplier	Function
Adobe Photoshop 5.5	Adobe Systems, Inc.	Image processing
BLAST http://www.ncbi.nlm.nih.gov/BLAST/	(Altschul et al. 1997)	Comparative alignment against public database
CorelDraw 8.0.1	Pantone, Inc.	Drawing & sketching work
Edit Seq 4.00	DNAstar, Inc.	Sequencing analysis
EditView 1.0.1	Perkin Elmer Corp.	Sequencing analysis
Endnote 5.0	Niles & Associates, Inc.	Bibliographic software
Excel® Mac:2001	Microsoft Corp.	Calculation, Statistical evaluation
FileMaker©Pro 5	FileMaker, Inc.	Databases creation, Data management
GeneJockey ©1990	Biosoft	Sequencing analysis
Medline http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed	NCBI	Literature search
MegAlign 4.00	DNAstar, Inc.	Sequencing analysis
MultiAnalyst	BioRad	Image processing
Netscape Communicator 4.7	Netscape Corp.	Internet Browser
PowerPoint® Mac:2001	Microsoft Corp.	Drawing & sketching work
Primer select 4.00	DNAstar, Inc.	Primer design
SeqMan 4.00	DNAstar, Inc.	Comparative alignment against internal database
Simple Text D1-1.4	Apple Computer, Inc.	Word processing
Word® Mac:2001	Microsoft Corp.	Word processing

3.11.3 EST Database Creation

M. Hahn has done important spadework for the EST project described on the following pages. This included the development of a detailed work-sequence for isolation of *Uromyces fabae* haustoria and the establishment of a haustoria-specific cDNA library (Hahn and Mendgen 1992; Hahn and Mendgen 1997). Differential hybridization of this cDNA library led to the identification and description of 31 *in planta* induced genes (*PIG*s) (Hahn and Mendgen 1997). A substantial number of these in part highly expressed *PIGs* have either no homology to known genes, or have an interesting putative function. These include *THI1*, *THI2*, a hexose transporter gene and amino acid transporter genes. The function of these genes has been verified (Sohn 2000; Sohn et al. 2000; Voegele et al. 2001; Struck et al. 2002; Struck et al. 2004).

Based on the intriguing results obtained from the first analysis of these 31 *PIG*s, it was decided to gain further insight into gene expression in haustoria. Thus, a small-scale EST sequencing project was established with the goal of obtaining at least 1'000 readable sequences. For reasons of efficiency, PCR products of the cDNA library smaller than 500 base pairs were not included.

At the onset of the project, the above mentioned library of haustorial cDNA in the λgt10 phage (Hahn and Mendgen 1992; Hahn and Mendgen 1997) and 775 previously sequenced clones were provided by Prof. Dr. M. Hahn for further analysis. As not all unreadable sequences had been previously removed, the first step was to either remove these sequences or repeat the sequencing procedure. In addition, a further number of clones were sequenced during this project.

Prof. Dr. M. Hahn also provided a FileMaker™ database with information regarding the cDNA clones. Also using FileMaker™, an additional database was generated containing information with regard to the contigs. Both of the databases were linked together, making it possible to obtain data regarding specific contigs based on a known sequences, or vice versa.

All of the sequence data used during the project was analyzed using EditView 1.0.1 and manually trimmed of vector sequences. When possible, ambiguous base calls were corrected manually by comparison with the sequence traces. The revised sequence data was imported into DNAstar Seqman and comparatively aligned (Settings: Match size 30, Minimum Match Percentage 80). Both, non-redundant ESTs (singletons) and redundant ESTs (contigs in their original sense) were subsequently named contigs and compared with public databases (http://www.ncbi.nlm.nih.gov/BLAST/). This was performed using the gapped BLASTX algorithm, which among other things generates the so called "E value" (Altschul et al. 1997).

The Expect value (E value) is a parameter that describes the number of hits one can "expect" to find coincidentally when searching a database of a particular size. It decreases exponentially with the Score (S) that is assigned to a match between two sequences. Essentially, the E value describes the random background noise that exists for matches between sequences (http://www.ncbi.nlm.nih.gov/BLAST/blast_FAQs.shtml) (Lipman et al. 1984; Altschul and Erickson 1985). The E value is used as a convenient way to report the significance of sequence matches (Karlin and Altschul 1990). For example, an E value of 1 assigned to a hit can be interpreted as meaning that in a database of a particular size one might expect to see one coincidental match with a similar score.

Besides the BLASTX-algorithm, the NCBI server also provides the option of using the BLASTP and BLASTN algorithms. Whereas BLASTN consists of a single nucleotide query, BLASTX is a translated BLASTN query, without necessarily having a defined Open Reading Frame (ORF) and BLASTP allows a query at the protein level.

As it can be expected, BLASTX, BLASTP and BLASTN produce different results. Given that a definite ORF is assumed for a BLASTP search, fewer but more accurate hits can be expected in comparison to a BLASTX search, where all three possibilities for an ORF are tested.

Thus, if special interest exists regarding a specific gene, and the ORF is known, a BLASTP search is preferable. Within the scope of this thesis, all contigs were first “blasted” against various databases using the BLASTX algorithm. In a few cases, a subsequent BLASTP search was carried out. The BLASTN algorithm proved unsuitable for this EST project, as blasting with this algorithm led to a great number of false positive hits. Therefore, the BLASTN algorithm was not used during this thesis.

The length of the sequence to be blasted also has an effect on the E value. Doubling the length of a sequence doubles the number of high scoring pairs attaining a given score (Smith et al. 1985). Another factor is the length of the proteins in the database. The BLAST programs (Altschul et al. 1990; Altschul and Gish 1996; Altschul et al. 1997) assume that the *a priori* chance of relatedness is proportional to sequence length. In this context, the BLAST programs treat the database as a single long sequence of length, and use this approach for calculating the database E-value. For DNA sequence comparisons, the length of database records is largely arbitrary. Therefore, the BLAST approach is a very efficient method for estimating statistical significance (Altschul et al. 1990).

Other important factors affecting the E value are the size and composition of the databases against which the BLAST search is run. The smaller the database, the less results can be expected. This in turn has an effect on the E value. E values generated from different databases and even different searches of the same database can vary to a certain extent. The ESTs found during this project were primarily compared to one of the largest databases at NCBI, for which a cut off value of e-5 was defined.

Putative functions have been assigned to ESTs/contigs based on the results returned from searches using the BLASTX algorithm. Generally, non-redundant sequences with the same or very similar BLASTX annotations, but that do not group in one contig, were assumed to code for different proteins.

A clone was normally sequenced only with one primer (GT10R). In cases where this sequencing failed or the result seemed to be very promising, additional sequencing with the primer GT10F was carried out. In some cases where one clone was sequenced from both ends (with primers GT10F and GT10R) and produced no overlapping parts, two seemingly different contigs were derived from non-overlapping regions of the same gene transcript. In this case, the contig name was made up of two contig numbers (e.g. C400/C331).

All contigs were grouped in functional classes according to Kamoun et al. (1999). Any contig that did not produce a BLASTX hit or produced hits with E values greater than cut-off value of e-5 was assigned to the category 15 “Unclassified Proteins” (Kamoun et al. 1999). All contigs whose E value was < e-5 were assessed as having generated a significant hit. Sequences producing hits to known proteins with unknown function were classified under category 14 “Similarity to Proteins of Unknown Function”. The remaining contigs were sorted into 13 functional categories with additional subcategories based on their assumed function. These categories were based on a publication of an EST project on the oomycete *Phytophthora infestans* (Kamoun et al. 1999). If necessary, some categories or subcategories were added, e.g. for putative viral genes, that were not found in the *Phytophthora infestans* EST project.

Because the classification is rather close-meshed, a number of proteins could theoretically be placed in more than one category. In those cases where it was difficult to decide in which category a contig should be placed, the more specialized category was chosen. The following Table 3-63 lists the functional categories into which the contigs have been grouped.

Table 3-63 Functional Classes of Contigs according to Kamoun et al. (1999) (altered)

No.	Category
01	**Metabolism**
01.01	Amino acid metabolism
01.02	Nitrogen and sulfur metabolism
01.03	Nucleotide metabolism
01.04	Phosphate metabolism
01.05	Carbohydrate metabolism
01.06	Lipid and sterol metabolism
01.07	Biosynthesis of vitamins, cofactors and prosthetic groups
01.08	Aminosugar metabolism
02	**Energy**
02.01	Glycolysis & pyruvat metabolism
02.02	Gluconeogenes
02.03	Pentose phosphate pathway
02.04	TCA pathway
02.05	Respiration
02.06	Fermentation
03	**Cell Growth/ Division/ DNA Synthesis**
03.01	Cell growth
03.02	Meiosis
03.03	DNA synthesis/ replication
03.04	Recombination/ DNA repair
03.05	Cell cycle/ cell cycle control
03.06	Cytokinesis
03.07	Growth regulators
03.08	Other proteins
04	**Transcription**
04.01	rRNA synthesis
04.02	tRNA synthesis
04.03	RNA transport
04.04	mRNA synthesis
04.05	General transcription factors
04.06	Transcriptional control
04.07	Chromatin modifier
04.08	mRNA processing
04.09	Other proteins involved in transcription
05	**Protein synthesis**
05.01	Ribosomal proteins
05.02	Translational factors
05.03	Translational control:
05.04	tRNA synthesis
05.05	Other proteins involved in protein synthesis
06	**Protein destination**
06.01	Folding and stabilization of proteins:
06.02	Targeting/ sorting/ translocation of proteins
06.03	Modification of proteins
06.04	Assembly of protein complexes

No.	Category
06.05	Proteolysis
06.06	Other proteins involved in protein destination
7	**Transport facilitators**
07.01	Ion channels
07.02	Ion transporters
07.03	Sugar/ carbohydrate transporters
07.04	Amino acid transporters
07.05	Lipid transporters
07.06	Purine/ pyrimidine transporters
07.07	Transport ATPases
07.08	ABC transporters
07.09	Other transport facilitators
8	**Intracellular traffic**
08.01	Nuclear transport
08.02	Mitochondrial transport
08.03	Vesicular transport
08.04	Extracellular transport:
08.05	Cellular import
08.06	Other proteins involved in intracellular traffic
9	**Cellular Organization and Biogenesis**
09.01	Cell wall/ plasma membrane
09.02	Cytoplasm
09.03	Cytoskeleton
09.04	ER/ golgi
09.05	Nuclear
09.06	Chromosome structure
09.07	Mitochondrial
09.08	Peroxisomal
09.09	Vacuolar
09.10	Other proteins involved in cellular organization and biogenesis
10	**Signal transduction**
10.01	Receptors
10.02	Mediators
10.03	Kinases
10.04	Phosphatases
10.05	G proteins
10.06	Other proteins involved in signal transduction
11	**Pathogenicity factors, avirulence factors and elicitors**
11.01	Pathogenicity factors
11.02	Avirulence factors
12	**Cell Defence**
12.01	Stress response
12.02	Detoxification
13	**Viral Proteins**
14	**Similarity to proteins of unknown function**
15	**Unclassified proteins**

Included in the results are also those exceptional cases where a low score hit (E > e-5) nonetheless seemed to of interest. These are marked as exceptions to the rule. All contigs that did not produce hits within the NCBI database were compared to the special fungal databases of *Magnaporthe grisea* and *Neurospora crassa* at the Whitehead Institute, Center for Genome Research, Cambridge, MA, USA.

For this EST project, the following BLASTX settings (BLAST Version 2.0) were used: Expect value (default): 10, Word size: 3, Matrix: BLOSUM 62, Gap Costs Existence: 11, Extensions: 1.

3.11.4 Cleavage Site Prognosis

The SignalP server at http://www.cbs.dtu.dk/services/SignalP-2.0/ predicts the presence and location of signal peptide cleavage sites in amino acid sequences from various organisms: Gram-positive prokaryotes, Gram-negative prokaryotes, and eukaryotes (Nielsen et al. 1997b; Nielsen et al. 1997a; Nielsen and Krogh 1998; Nielsen et al. 1999). SignalP version 2.0 comprises two signal peptide prediction methods: SignalP-NN (based on neural networks, corresponding to SignalP, version 1.1) and SignalP-HMM (based on hidden Markov models). Thus, the SignalP server allows the prediction of a cleavage site with two independent models based on different algorithms at the same time. Both of the results are displayed in text and graphically.

For eukaryotic data, SignalP-HMM allows substantially improved discrimination between signal peptides and uncleaved signal anchors, but it has a slightly lower accuracy in predicting the precise location of the cleavage site. SignalP-HMM provides 1) prediction of the presence of a signal peptide; 2) the position of the cleavage site; 3) an approximate assignment within the signal peptide of the N-terminal positively-charged region (N-region); 4) the central hydrophobic region (H-region); and 5) the c-region (C-terminal region). These are displayed in a graph as probabilities for these three regions being at specific positions of the protein sequence (Nielsen et al. 1997b; Nielsen et al. 1997a; Nielsen and Krogh 1998; Nielsen et al. 1999).

3.11.5 Theoretical Analysis of Protein Localization and Protein Stability

The theoretical analysis of protein localization was carried out using two internet-based database applications, PSORT II (Nakai and Kanehisa 1992) and PROSITE (Gasteiger et al. 2003).

PSORT II is a computer program for the prediction of protein localization sites within cells, accessible at http://psort.ims.u-tokyo.ac.jp/ (University of Tokyo, Tokyo, Japan). It receives the information of an amino acid sequence as input, which is analyzed by applying the stored rules for various sequence features of known protein sorting signals. PSORT II reports the possibility for the tested protein to be located at each candidate site, together with additional information. The applied version of PSORT II was trained on yeast sequences from the SWISS-PROT database.

PSORT II requires an amino acid sequence and its source of origin as input information, i.e. yeasts, bacterial or plant. It analyzes the input sequence by applying the defined rules for various sequence features of known protein sorting signals, and reports the possibility for the input protein to be located at each candidate site. PSORT II predicts the presence of signal sequences by using McGeoch's method (McGeoch 1985), modified by Nakai and Kanehisa (1991; 1992). It considers the N-terminal, charged region (CR) and the central hydrophobic region (UR) of signal sequences. A discriminant score is calculated from the three values: length of UR, peak value of UR, and net charge of CR. These results are summarized as the "McG". A large positive discriminant score indicates a high possibility of possessing a signal sequence, whether it has been cleaved off or not.

PSORT II also applies von Heijne's method of signal sequence recognition (Nielsen et al. 1999). It is a weighted-matrix method that provides information about consensus patterns around the cleavage sites (the "(-3, -1)-rule"), and thus can be used to detect non-cleavable signal sequences. The output of this "GvH" method is the original weighted-matrix score (for eukaryotes) subtracted by 3.5. A large positive score indicates a high possibility that the amino acid sequence has a cleavable signal sequence. The position of a possible cleavage site is also reported. The various algorithms applied by PSORT II are detailed in annex 8.2. For further explanations see: http://psort.nibb.ac.jp/psort/ helpwww2.html.

PROSITE is a database of protein families and domains at the ExPasy server (Gasteiger et al. 2003) at the Swiss Institute of Bioinformatics (SIB, Geneva, Lausanne and Basel, CH). It consists of biologically significant sites, patterns and profiles that help to reliably identify whether a sequence contains a known protein pattern (Bucher and Bairoch 1994; Hulo et al. 2004). It is accessible at http://www.expasy.org/prosite/.

During this thesis, PROSITE was used to compare the studied *PIG*s with databases in regards to whether or not similarities in the sequences respectively sequence patterns exist. This allows a grouping of the proteins into a limited number of families.

Two different methods were used to theoretically predict PIGp stability: These were the "N-end rule" and the "Estimated Half-Life and Stability Index", both available on the ExPasy server.

The estimation of the in vivo half-life of proteins was carried out using to the "N-end rule" (Tobias et al. 1991) on the ExPASy server. The idea behind the calculation of the estimated half-life and the stability of a protein is that the identity of the N-terminal residue of a protein plays an important role in determining its stability in vivo (Bachmair and Varshavsky 1989; Gonda et al. 1989).

The “Estimated Half-Life and Stability Index” is based on research done by Guruprasad et al. (1990). A protein whose instability index is smaller than 40 is predicted to be stable, whereas a value above 40 indicates that the protein may be unstable (Guruprasad et al. 1990). Based on statistical analysis of 32 stable and 12 unstable proteins, Guruprasad et al. (1990) have shown that there are certain dipeptides that occur significantly more often in unstable proteins than in stable ones. Therefore, the authors have assigned an instability value to each of 400 different dipeptides.

4 Results

4.1 *Uromyces fabae* EST Project

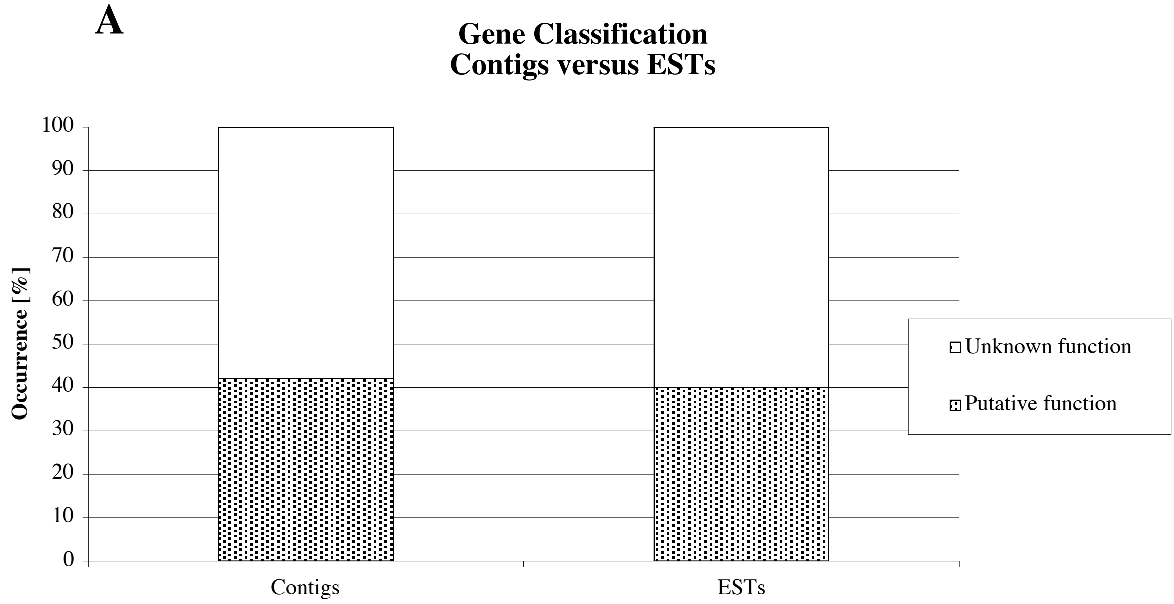
4.1.1 Statistical Analysis of ESTs and Contigs Distribution

The average length of the 1’032 haustorial ESTs sequenced was 597 base pairs. These sequences were merged into 540 “contigs”. Contigs are defined as ESTs cleansed of redundancies, each representing one specific gene. The similarity and putative gene function of sequenced ESTs was determined using gapped BLASTX (Altschul et al. 1997) at the National Center for Biotechnology Information (NCBI). Of the 1’032 ESTs sequenced, 607 did not exhibit a significant similarity (E value ‹ e-5) to entries in the GenBank database. 379 sequences matched previously identified genes and 46 matched genes with unknown or putatively known function. Out of the 540 contigs (containing the 1’032 ESTs), 309 did not exhibit a significant similarity to genes in the GenBank database.

Figure 4-1-A presents the distribution of the above-mentioned contigs respective ESTs in respect to unknown or putative function.

Figure 4-1-B displays the distribution of the above-mentioned ESTs and contigs with a putatively function into their functional classes. Genes with no similarities (ESTs 60%, contigs 58%) are only included in the first bar graph. In the second bar graph, only ESTs respectively contigs with E values ‹ e-5 are displayed (both in absolute numbers). In order to give a better impression of the activities in the haustorium, the graph compares by functional category, the occurrence of ESTs (not cleansed of redundancies) with the occurrence of contigs (cleansed of redundancies). The assignment to functional categories was performed according to Kamoun et al. (1999).

A



B

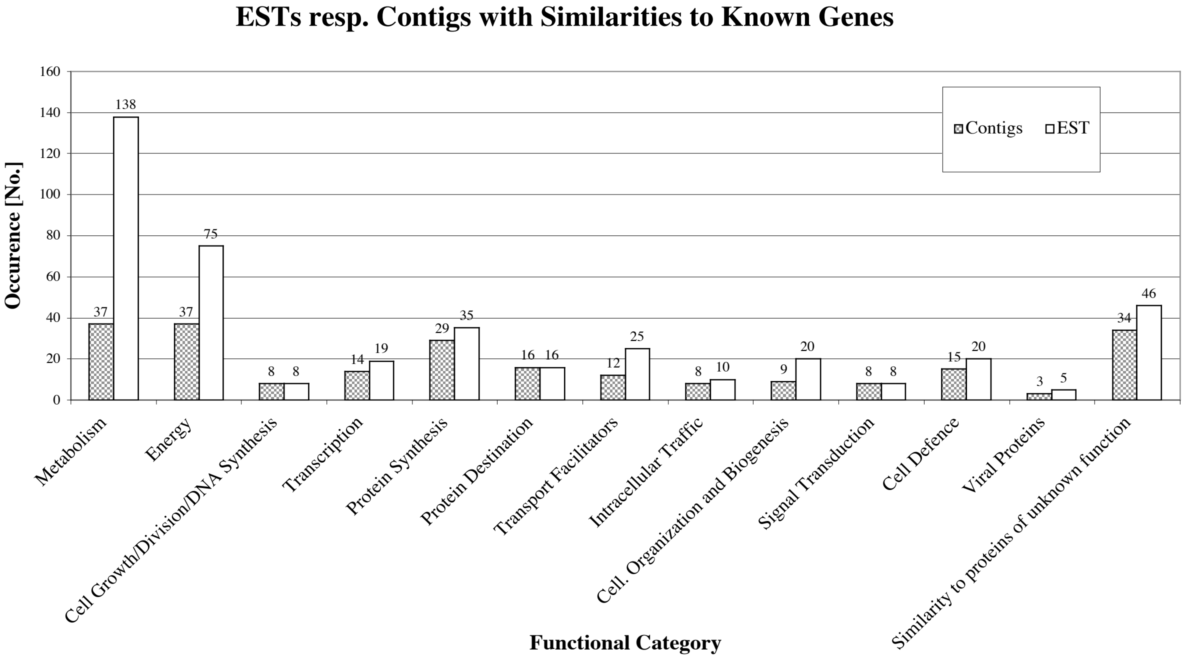


Figure 4-1 Classification of sequence similarities of *Uromyces fabae* haustorium-specific ESTs compared to the corresponding contigs of known genes using the BLASTX algorithm. A: Total number of ESTs = 1'032; total number of corresponding contigs = 540; 607 ESTs (60%) with unknown function correspond to 309 Contigs (58%) with unknown function. B: ESTs and the corresponding contigs grouped according their functional classification.

When taking a closer look at the ESTs with similarities to previously identified genes, it becomes evident that half of the sequenced EST clones with known functions represent either genes related to metabolism (138 sequences (13.4%)) or to energy transformation (75 sequences (7.3%)). The third largest category of ESTs (46 sequences (4.4%)) represents genes with similarities to proteins with an unknown function. 35 (3.4%) of the ESTs represent genes involved in protein synthesis, followed by a group of genes coding for transport facilitators (25 sequences (2.4%)). Genes for cell defense and cellular organization & biogenesis make up 20 sequences (1.9%) each, followed by genes for transcription (19 sequences (1,8%)) and protein destination (16 sequences (1,5%)). 10 sequences (0.96%) represent genes for intracellular traffic, 8 sequences (0.77%) represent genes for signal transduction, and 8 sequences (0.77%) represent genes for cell growth/division & DNA synthesis. The smallest group (5 sequences (0.5%)) contains genes of probable viral origin.

There is a conspicuous difference between the allocation of contigs and the allocation of ESTs to the different categories. Whereas in both cases most of the contigs (310 sequences (58%)) respectively ESTs (607 sequences (60%)) fit into the category „unclassified proteins", the distribution changes for the smaller categories. In the case of the contigs, genes related to metabolism (37 sequences (7%)) and energy transformation (37 sequences (7%)) are followed by genes with similarities to proteins of unknown function (34 sequences (6%)). The EST classification follows the same order, although the number of ESTs related to metabolism is significantly greater then those related to energy transformation. Despite the fact that the categories metabolism and energy transformation have the same number of contigs, the category metabolism contains nearly the double amount of ESTs. Therefore, some of these ESTs are strongly expressed (e.g. contigs C008, C023, see Annex 8.1).

Genes related to metabolism and energy transformation belong the second large group of genes (together14%), this leads to the conclusion that haustoria (the origin of these ESTs) are metabolically highly active components with a high energy turnover.

Despite the above findings, most of the genes found during this EST project have an unknown function (unclassified and unknown function, together 67%). Thus, it becomes clear that there is still a great deal to explore regarding haustorial gene expression.

28 of the most highly expressed contigs are shown in the chart below (Figure 4.2). All contigs that consist of six or more ESTs are displayed. Some of these were represented by a large number of cDNA clones (up to 36), indicating a strong expression of the corresponding genes in haustoria. The function of 4 of these genes is known, while the function of another 3 can be assumed based on high similarities to genes in the NCBI database. To the remaining 21 of the most frequently found genes, no putative function can be assigned. 10 of these are *in planta induced genes* (*PIG*s) (Hahn and Mendgen 1997). These genes are of great interest because of the possibility that they could be involved in the pathogenesis or biotrophy of *Uromyces fabae*. Therefore, they have been analyzed regarding a putative signal sequence and cleavage site (see chapter 4.2.1).

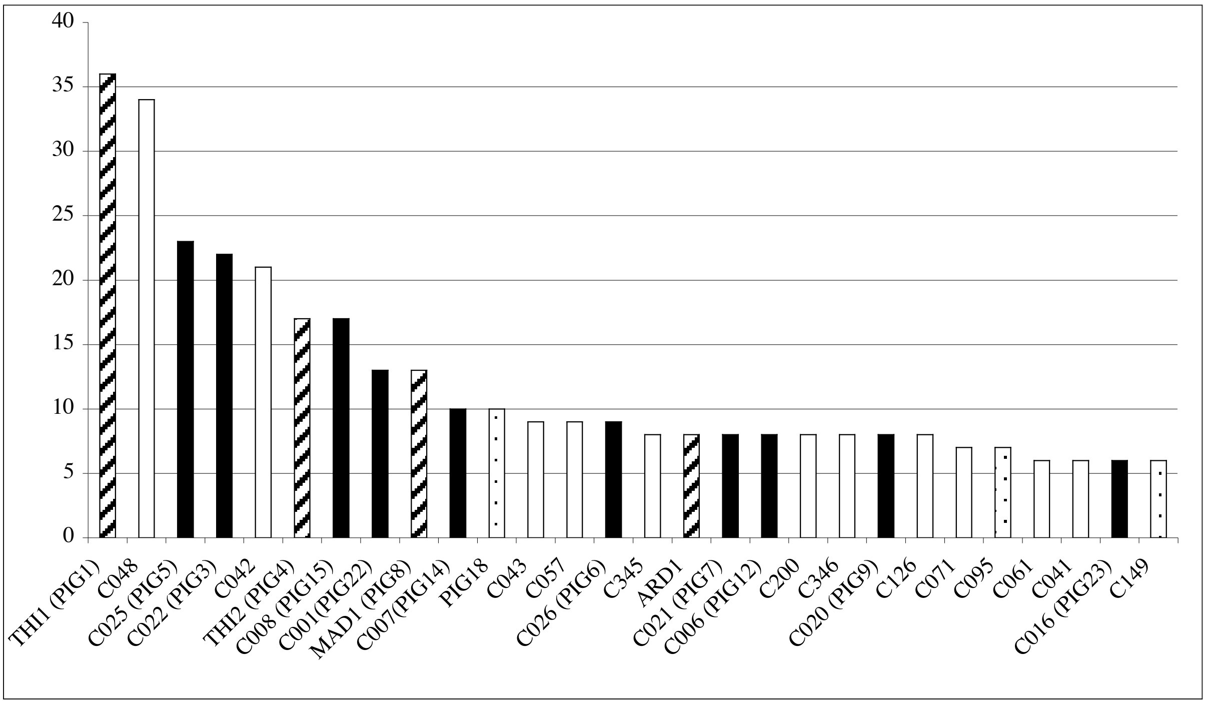


Figure 4-2 Number of ESTs found in the 28 most highly expressed contigs (out of 1032 total ESTs). Genes with a verified function are indicated by stripes (*THI* = Thiamin biosynthetic gene, *MAD* = Mannitol dehydrogenase, *ARD* = D-arabitol dehydrogenase). *PIG*s (*in planta induced genes*) of unknown function are indicated by black columns. Genes with a high similarity to genes in the NCBI database are indicated by dots. Blank columns represent genes (without *PIG*s) of unknown function.

The 4 genes with know functions are *THI1*, *THI2 MAD1* and *ARD1.* THI1p and THI2p are involved in vitamin synthesis (Sohn 2000; Sohn et al. 2000), MAD1p (Voegele et al. 2005) and ARD1p (Link et al. 2005) are involved in carbohydrate metabolism. Of the 3 genes with high similarities to genes in the NCBI database, *PIG18* has a strong similarity to an endochitinase (BLASTX P(N) = 2e-44), and the contigs C095 and C149 to malate dehydrogenase (BLASTX P(N) = e-10 and e-105 respectively).

The next table shows the *PIG*s found during the EST project which were previously described by Hahn et. al (1997).

Table 4-1 *PIG*s found during the EST project

PIG	Contig	(Probable) Function / NCBI BLASTX "hit"	Frequency of PIG found during the EST project
PIG1	C003	*THI1* (Sohn et al. 2000)	36
PIG2	C013	*AAT2* (Hahn et al. 1997)	4
PIG3	C022	Unknonwn function	22
PIG4	C051	*THI2* (Sohn et al. 2000)	17
PIG5	C025	Unknonwn function	23
PIG6	C026	Unknonwn function	9
PIG7	C021	*RTP1* (this treatise)	8
PIG8	C023	*MAD1* (Voegele et al. 2005)	13
PIG9	C020	(*Rattus norvegicus*) Heavy neurofilament subunit (E value: 2e-07)	8
PIG10	C004	Unknonwn function	4
PIG11	C346	Unknonwn function	8
PIG12	C006	Unknonwn function	8
PIG13	C043	Unknonwn function	9
PIG14	C007	Unknonwn function	10
PIG15	C008	(*Oryza sativa*) Put. Trehalose-phosphate synthase (E value: 1e-7)	17
PIG16	C009	P-450 monooxygenase (Hahn and Mendgen 1997)	1
PIG17	C010	(*Magnaporthe grisea*) 70-15 chromosome (E value:1e-15)	1
PIG18	C011	(*Puccinia triticina*) Chitinase (E value: 9e-157)	10
PIG19	C012	Unknonwn function	3
PIG20	C014	Unknonwn function	5
PIG21	C017	Unknonwn function	1
PIG22	C001	Unknonwn function	13
PIG23	C016	Unknonwn function	6
PIG25	C019	Unknonwn function	2
PIG27	C015	*AAT1* (Struck et al. 2002)	2
PIG28	C024	*PIG28* (Hahn and Mendgen 1997), (*Malassezia pachydermatis*) Cyclophilin (E value: 7e-50)	4
PIG34	C217	*AAT3* (Struck et al. 2004)	1

Not all of the *PIG*s as described by Hahn and Mendgen (1997) have been found during this treatise.

245 out of 1032 ESTs found during this project are *in planta* induced genes (*PIG*s). These findings show that more than 23% of the ESTs of this EST project are presumably related to biotrophy or pathogenesis of *U. fabae*.

4.1.2 Putative Functions of Selected Genes

Table 4-2, Table 4-3, and Table 4-4 below provide an overview of those contigs (i.e. genes), which are of interest with respect to presumed (or already proven) haustorial function. During the EST project, a number of contigs/genes of putative viral origin were also unexpectedly discovered. These are listed separately in Table 4-5.

The tables below also include contigs of interest whose E values are higher than the cut-off value of 1e-5. A short description of the best hit is provided, including name and accession number of the coded gene. In doing so, the notation of the source organism was adopted. Therefore it is possible to find listed both the teleomorph- and the anamorph names of the same fungus, e.g. *Emericella nidulans* and *Aspergillus nidulans*. In those cases where the function of an *Uromyces fabae* gene has already been verified, this fact has been explicitly mentioned. The assignment to functional categories was performed according to (Kamoun et al. 1999).

For the sake of legibility, only the most interesting contigs discovered during the EST project are described in this chapter. A short description of all contigs with similarities to the NCBI database can be found in Annex 8.1.

Table 4-2 lists five contigs found to have a significant similarity to proteases, as well as one contig with a probable similarity to pepsin.

Table 4-2 Evidence for Proteases Produced in Rust Haustoria

Contig No.	Protein (best hit)	Organism	BLASTX E value
C123	Extracellular elastinolytic metalloproteinase precursor	*Aspergillus fumigatus*	6e-23
C155	Proteinase A	*Pichia angusta*	4e-46
C175	*Pepsin	*Sus scrofa*	6e-04
C210	Probable zinc metallo-protease	*Neurospora crassa*	2e-42
C537	Leucine aminopeptidase	*Coprinopsis cinerea*	1e-48
C545	Cellular serine Proteinase	*Aspergillus fumigatus*	1e-104
C569	Metalloproteinase	*Aspergillus fumigatus*	6e-23

*E value is greater than the cut-off value of e-5

In Table 4-3, contigs with similarities to genes that may be specific for the haustorial function, amino acid uptake, nutrition and vitamin provision are described.

Table 4-3 Evidence for the Specific Haustorial Functions Amino Acid Uptake, Nutrition–, and Vitamin Provision in *U. fabae*

Contig No.	No of corresponding ESTs	Protein	Reference
C013	4	AAT2p, Amino Acid Transporter 2 (PIG2p)	(Hahn et al. 1997)
C015	2	AAT1p, Amino Acid Transporter 1 (PIG27p)	(Struck et al. 2002)
C217	1	AAT3p, Amino Acid Transporter 1 (PIG34p)	(Struck et al. 2004)
C023	13	MAD1p, Mannitol dehydrogenase 1 (PIG8p)	(Voegele et al. 2005)
C003	36	THI1p Thiamine Biosynthetic Enzyme 1 (PIG1p)	(Sohn et al. 2000)
C051	4	THI2p. Thiazole Biosynthetic Enzyme (PIG4p)	(Sohn et al. 2000)
C224	3	HXT1p, Hexose Transporter 1	(Voegele et al. 2001)
C098	8	ARD1p, D-arabitol dehydrogenase 1	(Link et al. 2005)

Because the NCBI database does not contain all public data about genome projects, the sequences of the *U. fabae* EST project were blasted against the *Magnaporthe grisea* and *Neurospora crassa* databases at the Whitehead Institute, Center for Genome Research, Cambridge, Ma, USA. Table 4-4 lists 12 contigs with no similarities to entries in the NCBI database, but with similarities to the above-mentioned fungal databases.

Table 4-4 EST Contigs Displaying Similarity to Genes not in the NCBI Database

Contig	*Magnaporthe grisea* (best hit) Accession No./ Description✣	BLASTX E value⁕	*Neurospora crassa* (best hit) Accession No./ Description✣	BLASTX E value⁕
C047	MG09736.1 predicted protein	3e-06	-	-
C093	MG04522.1 predicted protein	1e-05	NCU00903.1 predicted protein	4e-07
C097	-	-	NCU09460.1 predicted protein	7e-09
C141	MG09372.1 predicted protein	7e-07	NCU00695.1 predicted protein	3e-07
C208	MG06241.1 hypothetical protein	4e-06	NCU06410.1 hypothetical protein	1e-06
C291	MG02948.1 hypothetical protein	1e-07	NCU01914.1 hypothetical protein	9e-08
C323	-	-	NCU01021.1 hypothetical protein	8e-06
C366	MG03182.1 hypothetical protein	2e-06	NCU05452.1 hypothetical protein (related to spliceosome associated protein)	8e-06
C467	MG00998.1 hypothetical protein	2e-17	NCU07059.1 hypothetical protein	2e-16
C476	-	-	NCU03970.1 hypothetical protein	2e-07
C506	MG03673.1 hypothetical protein	4e-24	NCU00659.1 predicted protein	3e-28

⁕ BLASTX E value for a random hit. The cut-off value is >1e-05. ✣ Whitehead Institute, Center for Genome Research, Cambridge, MA, USA

Some of the ESTs studied here, which generated no matches to the NCBI database, generated matches to *N. crassa* and/ or *M. grisea* database at the Whitehead Institute, Center for Genome Research, Cambridge, MA, USA. Most of these ESTs display “hits” just below the cut off value of e-5 (see Table 4-4). In addition, all “hits” are against predicted or hypothetical proteins only. This leads to the conclusion that these genes are most likely very specific genes of the *Uromyces* family, but may have analogues in the above-mentioned species.

As mentioned before, a surprising number of contigs showing similarities to viral DNA was found. Even though viral sequences in rust fungi have already been described, (Dickinson and Pryor 1988b; Dickinson and Pryor 1988a; Dickinson et al. 1989; Zhang et al. 1994) not much is known about viral infections in *Uromyces fabae*. Those contigs having a homology to viral genes are displayed in the following table.

Table 4-5 Evidence for the Expression of Viral Genes

Contig No.	Protein (best hit)	Organism	BLASTX E value
C228	2a protein	Broad bean mottle virus	1e-12
C341	186K protein	Cucumber green mottle mosaic virus	7e-25
C404	186K protein	Cucumber green mottle mosaic virus	5e-09
C528	141K protein*	Pea early browning virus	8e-04
C544	beta-b protein*	Barley stripe mosaic virus	7e-01

*E values are higher than the cut off value of e-5

The number of viral sequences found during this EST project leads to the conclusion that in addition to cDNA of *U. fabae,* cDNA of at least one virus is included in the cDNA library used for this EST project. The origin of this viral cDNA will be discussed in chapter 5.1.9.

4.2 Analysis of Highly Expressed *PIG*s with Unknown Function

4.2.1 Cleavage Site Prognosis

A signal sequence with a probable cleavage site at the beginning of a protein provides an indication about the future localization of the mature protein. Therefore, all highly expressed *PIG*s were tested for the presence of a N-terminal, hydrophobic signal sequence. As a prerequisite for this test, the starting point of the cDNA sequence coding for the *PIG* must be known (which could not easily be confirmed). The results indicated that six of the highly expressed *PIG*s (*PIG*s found six times or more during the EST project) with unknown function (a BLASTX search showed no, or very minor homology to known genes, see Table 4-1, Annex 8.1) have a probable cleavage site.

Table 4-6 lists the protein sequences of the *PIG*s that provided positive cleavage site prognosis results. As recommended for SignalP-neural networks, only the first 70 amino acids were used for the signal sequence prediction. The predicted signal sequence is underlined.

Table 4-6 Protein sequence of highly expressed *PIG*s which showed positive cleavage site prognosis results using SignalP- neural networks (signal sequences underlined)

Protein Sequence of *PIG5*
<u>MQLHHLVTVLGLAFSQAQA</u>AIPMVANHDQVIRQIVTLITDPTHVMEAVPGLKELVPEAKLSPTLSHAQ
LNQGYREIFTSLRDVDNQHAPLIIDSLAGYHHTQDPARAEIYRNSLYSRIRDLSEHHPAQIRPHLDLTHH
RLLTTLNLQLQNLEESHLPLIGSDKAVQELKEFVARFKNRNEKEIKQLYHEFLSTIPPTPKSEMVYLSPE
QVNKALMETNSASKTPALPVDGNLWISTITHHHHHQTAPVAQTSPSLAAHAPQPVAASA

Protein Sequence of *PIG7* (*RTP1*)*
<u>MSNLRLLFTIISLAAIARA</u>XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXKRELDQDANPGHRR
HKSXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXX

Protein Sequence of *PIG9*
<u>MKSEQLFILVVLCLPCLYLG</u>RPLITAETAHIPVIAGPVQDALVSNEAIHSGSNLGLEVKDATAGGPSSIK
DVVGHANKGASPASEIAPQADKVKPEANEVAPDASDPVPKSAKSKWSDRFKEGMKKTETYSKDLFER
WFAKSKSAIKRWLAFIKDYFSRLRPKKATEKGTKVAAEEVELGLKPISQGKHDGPNVASPVTIVKTHP
QSEATIERQTIAKNPLADKEVKSSSEDQQSFNDATVEVPSQVHPPVEPQTIAKNPVGDEVKTISEDQKG
VSDAVPVKTLNPVGEEVTTISEDHQSLNNAVPVKVPSQVHPPVEPQTIAKNPVGEEVKTISEDQQGLNV
AIPAKTPTHVQHPLESPGIAKSPQQSSSDTPPSKPS

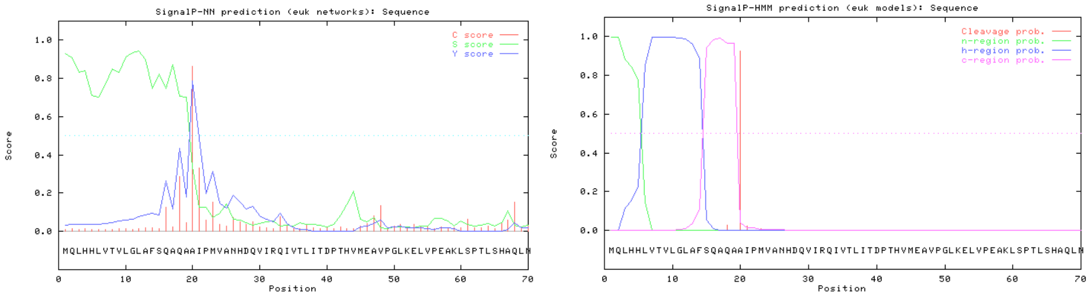
Protein Sequence of *PIG14*
<u>MPLTFAFLATILLSGSASA</u>NATGPSHVSRLAPRDEGANLRGTFIMSDNIDYSKGPLPVYYPNGSVAYLY
DQWYKQAGISTSTLTPVFPGTHHPTPVITLHSVDDGCAAKSHYSEFEVPGTSHLTYKIDPRGIKSDRWY
FEFVNHQGVRFRYRYHRNILSKGGKVYQYADKRPVRLVARLQEQLRWESWLNPGKSGAPTFTLSYDE
STGLGDKLVTLMALVMSRVENCGL

Protein Sequence of *PIG15*
<u>MHYSLFLVLSLFLQHARFAQS</u>VRPYNRIPILTVNEYKSAWHSADNRLILLDNDGVLLPQHGRNDHDVK
KAQDLLEALARDPKNTVWVITARGFDHVQEQYRVLRDRKVRLNLAGQLGTESRKWSEDTVRIQEGA
RDDVAALHRSLVEATGATFMPRNCCILYPKNKGQNDKLVLKEMQRLAGDFTFEPNSGDHATLTHPTI
NKGGFARDLFEKHERKTFVMSFGDADIDEKMHEAVNNSGFQTAVSTYVNTNGQSAARTRLDSHHDV
HRFFREMVGPEWF

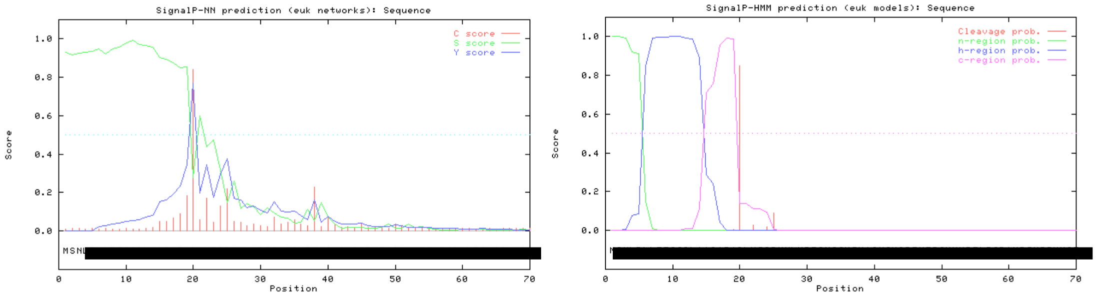
Protein Sequence of *PIG23*
<u>MRSYLLQFFLIGIAPAVRS</u>FSQAPVLMPRGMDHPKQLLKGKVSIEEISTTFTEATQSVVNTVSSPRSKTD
SATIAQQVTNVHIYAQQLTISVENLNDHKKMVTHQDTMIGAFVSYISMINAISDSKTRTTQCRNQLVSI
NIAFRSISTTYLASGIDLREEYNKHPHSPQYDPAKFAALDLDPLFDQVNPPAADLSTPEFDQSGEETDHA

*As agreed with Prof. Dr. K. Mendgen, only the signal sequence and the bipartite NLS of PIG7p (RTP1p) are shown

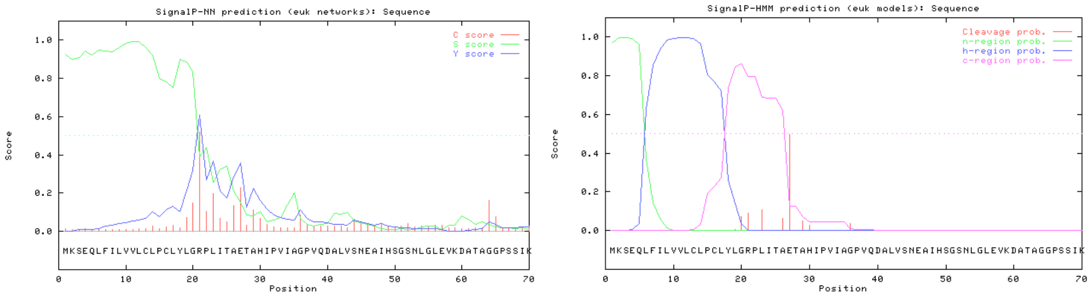
In Figure 4-3, the results for cleavage site prognosis using SignalP2.0 are presented in the form of graphs. The results for neural networks (NN) and hidden Markov models (HMM) trained on eukaryotes are shown next to each other.



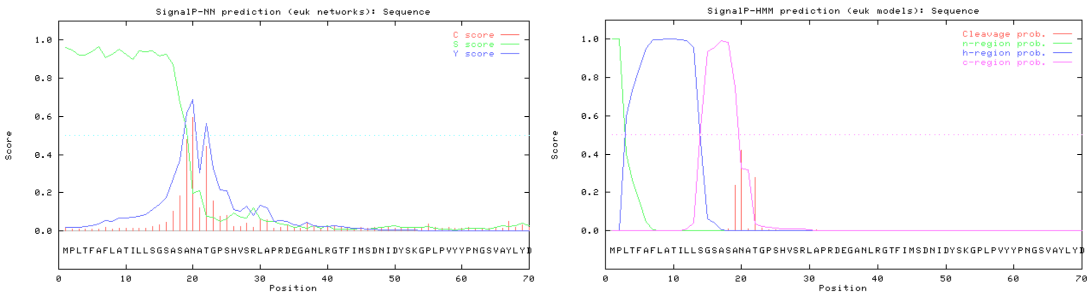
PIG5



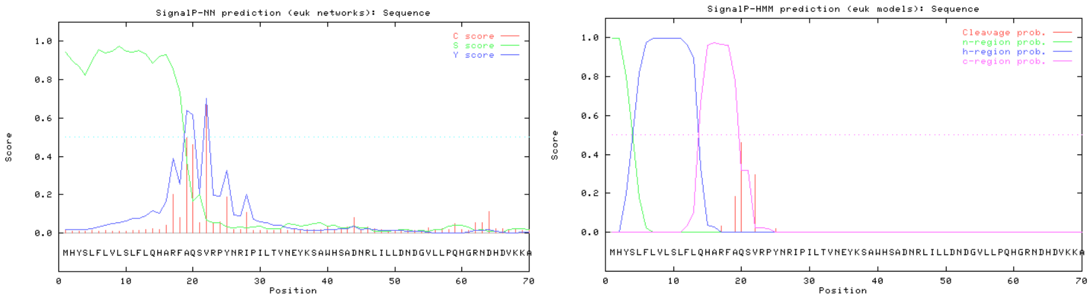
PIG7



PIG9



PIG14



PIG15

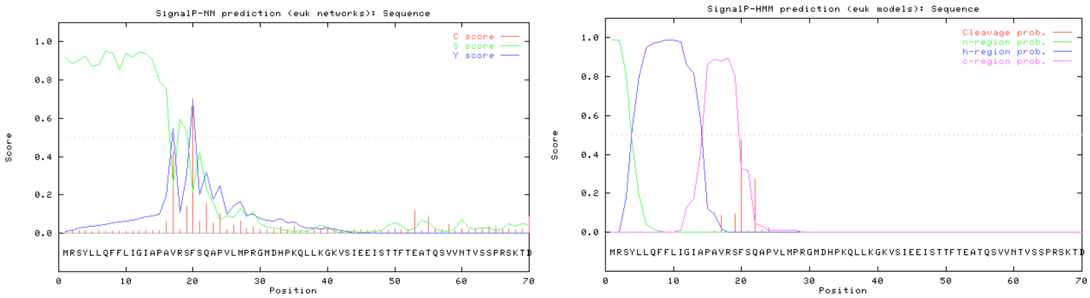


Figure 4-3 Cleavage site prognosis for the predicted proteins encoded by *PIG5, PIG7, PIG9, PIG14, PIG15* and *PIG23* using SignalP2.0. Only the first 70 amino acids of the ORF were used. On the left side, the graphical results for Neural Networks (NN) can been seen, while on the right, the results for Hidden Markov Models (HMM) trained on eukaryotes are displayed

Because of differences in cleavage site predictions between Neural Networks and Hidden Markov Models in SignalP2.0, a second prediction program (PSORT II, University of Tokyo, Tokyo, Japan) (Nakai and Kanehisa 1992) was used as well. Without exception, PSORT II provided the same result as Neural Networks, with all of the predicted cleavage sites being located between the 19th and 21st amino acid. Thus, for subsequent phases of this project, the Neural Network result was assumed to correctly identify the cleavage site.

Table 4-7 Cleavage site prognosis for different *PIG*s using Neural Networks (NN), Hidden Markov Models (HMM) trained on eukaryotes and PSORT

	Possible cleavage site (after amino acid position-no), as identified by:		
Gene	**Neural Networks**	**Hidden Markov Models**	**PSORT II (using G.V. Heinje)**
PIG5	19	19	19
PIG7	19	19	19
PIG9	20	26	20
PIG14	19	19	19
PIG15	21	19	21
PIG23	19	19	19

4.2.2 Analysis of Further Sequence-Specific Characteristics

Using PROSITE at http://www.expasy.org/prosite/ and PSORT II, at http://psort.nibb.ac.jp/form2.html, the six *PIG*s were tested for sequence-specific characteristics. The scope and functionality of PROSITE and PSORT II are explained in chapter 3.11.

Below, the most important results obtained from the comparison between the studied *PIG*s and the abovementioned databases are briefly described. In this context, special emphasis was placed on the theoretical prediction of the localization of the PIG-proteins and their stability.

- *Mature Protein Localization*

  Using the MTOP algorithm (Hartmann et al. 1989; Nakai and Horton 1999), the N-terminal side was predicted to be inside the cell for PIG5p, PIG7p (RTP1p) and PIG23p. PIG14p and PIG15p were predicted to have the C-terminal side inside the cell. This result is inconsistent with the predicted signal sequence.

  Using the NNCN algorithm (Reinhardt and Hubbard 1998; Nakai and Horton 1999), all the PIGps were predicted to be nuclear proteins, with the exception of PIG14p and PIG15p, which were predicted to be cytoplasmic proteins.

  According to the k-NN algorithm (Horton and Nakai 1997; Nakai and Horton 1999), PIG5p, PIG9p, PIG14p and PIG15p are located extracellularly or attached to the cell wall. PIG7p and PIG23p seem to be mitochondrial proteins, which is inconsistent with the nuclear localization sequence found for PIG7p (see below) and with the NNCN prediction for both proteins.

  Because of the inconsistent results provided by the various prediction programs, the data of the *in silico* analysis can only be interpreted in combination with the data from the lab experiments. Both together are discussed in chapter 5.2.3

- *Presence of NLS*

  Using the NUCDISC algorithm (Nakai and Horton 1999), PIG7p (RTP1p) and PIG9p are predicted to have a nuclear localization sequence (NLS). Especially for PIG7p, this result is extremely interesting in relation to the *in situ* localization results, which will be discussed later.

- *Protein Stability*

  The estimation of the in vivo half-life of proteins was carried out using to the "N-end rule" algorithm (Tobias et al. 1991). All six PIG-proteins were predicted to have a half-life of > 20 h in yeast and > 10 h in *E. coli,* which indicates that they are stable. In contrast, according to the Stability Index (SI) algorithm (Guruprasad et al. 1990), only PIG5p (SI: 36.98) and PIG15p (SI: 34.83) are predicted to be stable, whereas the other four proteins were classified as unstable. The cut off value for the stability index is generally defined as 40, with proteins with values above 40 being classified as unstable.

  During lab work, only PIG9p (SI: 56.71) and PIG23p (SI: 57.37) demonstrated instability and degraded during purification. Therefore, and because *PIG23* does not seem to be a real *in planta* induced gene (see chapter 4.8), *PIG9* and *PIG23* were not further studied.

Isolation and purification of PIG5p and PIG15p, PIG7p (RTP1p) and PIG14p could be successfully carried at RT, despite the theoretically calculated instability of RTP1p (PIG7p) (SI: 41.50) and PIG14p (SI: 43.76). In this context, it must be noted that the over expression, isolation and purification of PIG14p was repeated several times to compensate for a low protein yield.

A table with the complete results generated by the prediction programs, can be found in the annex 8.3.

4.3 Comparative Infection Studies

To study whether multiple fungus infections of *Vicia faba* have a (reciprocal) amplifying or inhibiting effect, parallel inoculations were carried out with *Uromyces fabae* and *Ascochyta fabae* or *Botrytis fabae* respectively.

The results of these experiments were unsatisfactory, because even under controlled laboratory conditions, the *A. fabae* and *B. fabae* infection-rates were very poor. Repeat experiments resulted in strongly differing infection times. The duration from inoculation to visible infection was between one to three months. Better results were achieved when the infected bean plants were cultivated outside for several weeks with minimal care. The infections were more pronounced than under laboratory conditions, but the infection time was still between one to three months. Therefore, this experimental method was judged to be inadequate, and the comparative infection studies were not continued.

4.4 Infiltration Experiments

To study a possible effect of antibodies on the different PIG-proteins *in vivo*, antibodies in a dilution of 1/10 or 1/100 were infiltrated into bean leaves immediately prior and up to 3 days after infection with *U. fabae* spores.

Although various antibody infiltration methods were tested, no conclusive results regarding *U. fabae* infection inhibition were obtained. When comparing the effects of infiltration with antibody solution, buffer and plain water, with leaves left completely untreated, a one day delay in the appearance on Uredospores on the lower side of the infiltrated leaves could be observed, independent of the applied infiltration solution. This effect can most likely be attributed to the (partial) separation of the epidermis from the underlying leaf structures as a result of the internal pressure caused by infiltration.

Based on these results, the applied method for infiltration experiments proved to be inadequate, and, the infection studies were not continued.

4.5 Gene Frequency Analysis

In order to analyze the frequency with which the various *PIG*s occurred, a Southern Blot Analysis as described in chapter 3.8.5.1was carried out. Restriction enzymes were employed as indicated in Figure 4-4. Since the restrictions sites for these enzymes were known only for the cDNA of *U. fabae*, the estimation of the number of gene copies relied on the existence of distinct hybridization signals. The estimated numbers of gene copies are listed in Table 4-8.

Regarding *PIG7* (*RPT1*), the results of the preceding EST project indicated that there may be a mutated copy of this gene present in *U. fabae*. The cDNAs in question differ in only one amino acid. The location of this difference coincides with a restriction site for the enzyme *Mun*I (*Mfe*I). Therefore, this enzyme was used as an additional restriction enzyme for the Southern Blot of *PIG7*.

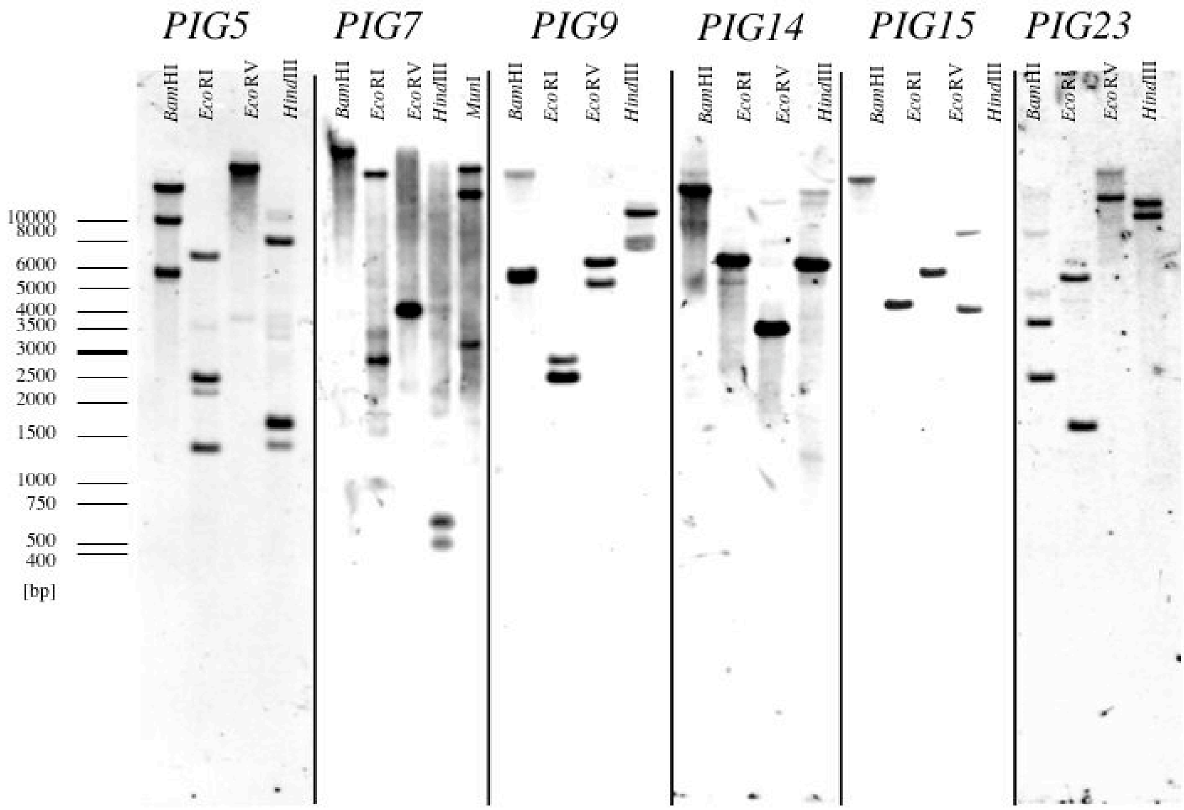


Figure 4-4 Quantification of gene copies via Southern Blotting: 1 μg of *U. fabae* DNA per lane was digested with the enzymes indicated. DIG-labeled cDNA probes of the indicated *PIG*s were used for hybridization.

Table 4-8 Number of hybridization signals after digestion with the respective restriction enzyme in *U. fabae* total DNA and number of gene copies determined by Southern Blotting

Enzymes	*PIG5*		*PIG7* (*RTP1*)		*PIG9*		*PIG14*		*PIG15*		*PIG23*	
*Bam*HI	3	(0)	1	(0)	2	(0)	1	(0)	1	(0)	2	(1)
*Eco*RI	4	(1)	2	(0)	2	(0)	1	(0)	1	(0)	2	(0)
*Eco*RV	1-2	(0)	1	(0)	2	(0)	1	(0)	1	(0)	1-2	(0)
*Hind*III	4	(1)	2	(1)	2	(0)	1	(0)	2	(0)	2	(0)
*Mun*I		-	3	(1/0)		-		-		-		-
Gene Copies	**2**		**2**		1-**2**		**1**		**1**		**1**	

The actual numbers of known restriction sites in the cDNA are shown in parenthesis.

Based on the above results, it becomes evident that in *U. fabae,* two copies of *PIG5* and *PIG7* and one copy of *PIG14, PIG15* and *PIG23* exist. As expected based on the different cDNAs found for *PIG7* during the EST project, the digestion with the restriction enzyme *Mun*I resulted in three hybridization signals. Therefore, it can be assumed that two different copies of *PIG7* with a small difference in one amino acid exist. Unfortunately, no clear conclusion can be drawn for *PIG9* based on the results of the Southern blots. In this case, either one or two copies exist in *U. fabae*.

4.6 Homology Analysis of *PIG7* (*RTP1*) to other Fungi

In order to determine whether *PIG7* has any significant homology to genes of other fungi, two Southern blot analyses as described in chapter 3.8.5.1 were carried out. Various fungal DNAs were tested for gene copies of *PIG7* using a DIG-labeled probe generated from cDNA cloned into the vector pET28a$^+$.

In order to improve the reliability of the experimental results, the Southern blotting was carried out twice with fungal DNA that had been digested with two different restriction enzymes. The restriction enzymes *Nco*I and *Eco*RI were chosen based on the best restriction results gained during prior tests.

In Figure 4-5, the results of the Southern blotting of *Nco*I-digested fungal DNA probed against *PIG7* are shown.

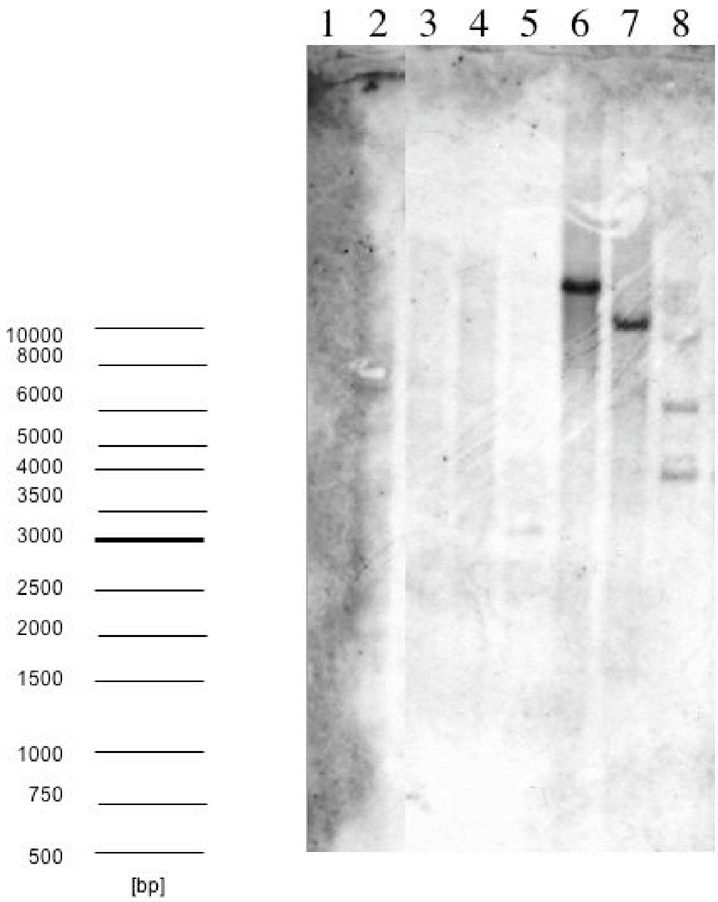


Figure 4-5 Detection of hybridization signals via Southern blotting: 10 μg of fungal DNA per lane were digested with *Nco*I. The lanes are labeled as follows: 1: *Vicia faba*; 2: *Ascochyta fabae*; 3: *Botrytis fabae*; 4: *Ustilago maydis*; 5: *Aspergillus nidulans*; 6: *Uromyces fabae*; 7: *Uromyces appendiculatus*; 8: *Uromyces striatus*. A DIG-labeled cDNA probe of the *PIG7* was used for hybridization (65°C hybridization temperature, washing with 2x SSC). (S. Bunk did blotting during a laboratory internship).

As can be seen, the *Nco*I-digested DNA of all three *Uromyces* species (lanes 6, 7, 8,) showed hybridization signals against *PIG7* with the DIG-labeled probe.

As was to be expected, no hybridization occurred with DNA from bean plants (*Vicia faba*) in lane 1. Also, none of the other tested fungal DNAs showed hybridization signals under the defined conditions (65°C hybridization temperature, washed with 2x SSC).

In Figure 4-6, the results of the Southern blotting of *Eco*RI-digested fungal DNA probed against *PIG7* are shown.

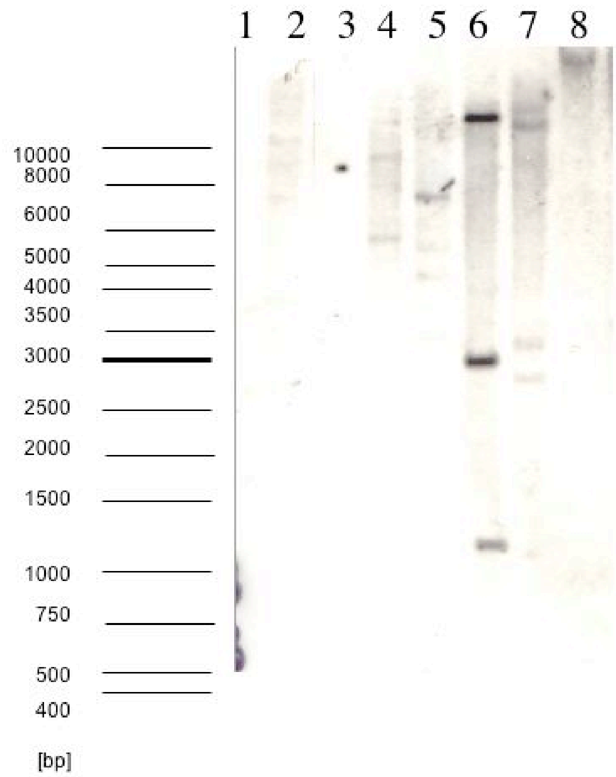


Figure 4-6 Detection of hybridization signals via Southern blotting: 10 μg of fungal DNA per lane was digested with *Eco*RI. The lanes are labeled as follows: 1: *Vicia faba*; 2: *Ascochyta fabae*; 3: *Botrytis fabae*; 4: *Ustilago maydis*; 5: *Aspergillus nidulans*; 6: *Uromyces fabae*; 7: *Uromyces appendiculatus*; 8: *Uromyces striatus*. DIG-labeled cDNA probes of the *PIG7* was used for hybridization (65°C hybridization temperature, washing with 2x SSC). (S. Bunk performed blotting during a laboratory internship).

As was the case for the *Nco*I-digested DNA, the *Eco*RI-digested DNA of all three *Uromyces* species (lanes 6, 7, 8) showed hybridization signals with the DIG-labeled probe against *PIG7*. In lane 4 (*U. maydis*) a lane diffuse hybridization signal is visible at about 5.5 kb and in lane 5 (*A. nidulans*) a diffuse hybridization signal is visible at about 7 kb.

Once again, no hybridization occurred with DNA from bean plants (*Vicia faba*) in lane 1. Also, none of the other tested fungal DNAs showed hybridization signals under the defined conditions (65°C hybridization temperature, washing with 2x SSC).

Both Southern blot analyses thus generated similar results, with the exception of lane 4 (*U. maydis*) and lane 5 (*A. nidulans*). These results confirm that analogues to *PIG7* exist in other *Uromyces* species besides *U. fabae*.

4.7 *In vitro* Germination of *U. fabae* Spores

In preparation for Northern and Western blots, RNA and proteins were isolated from *in vitro* structures of *U. fabae* during different stages of development. For optimal results, this isolation must take place when the desired development stage reaches its peak, but before the subsequent structure is formed. To determine the specific stage of spore development, probes with spores germinated on polyethylene foil were examined. This examination was carried out as described in chapter 3.1.4, and according to Deising et al. (1991). The experiments of Deising et al. (1991) were repeated in order to verify that the previously determined differentiation stages were consistent with the actual differentiation of the *in vitro* structures of *U. fabae*, as differentiation stages may vary between *U. fabae* spore batches.

Non-germinated spores were not included in the evaluation, because non-germinated spores could not be broken up with the applied method for RNA- and protein isolation. The results are displayed in Figure 4-7.

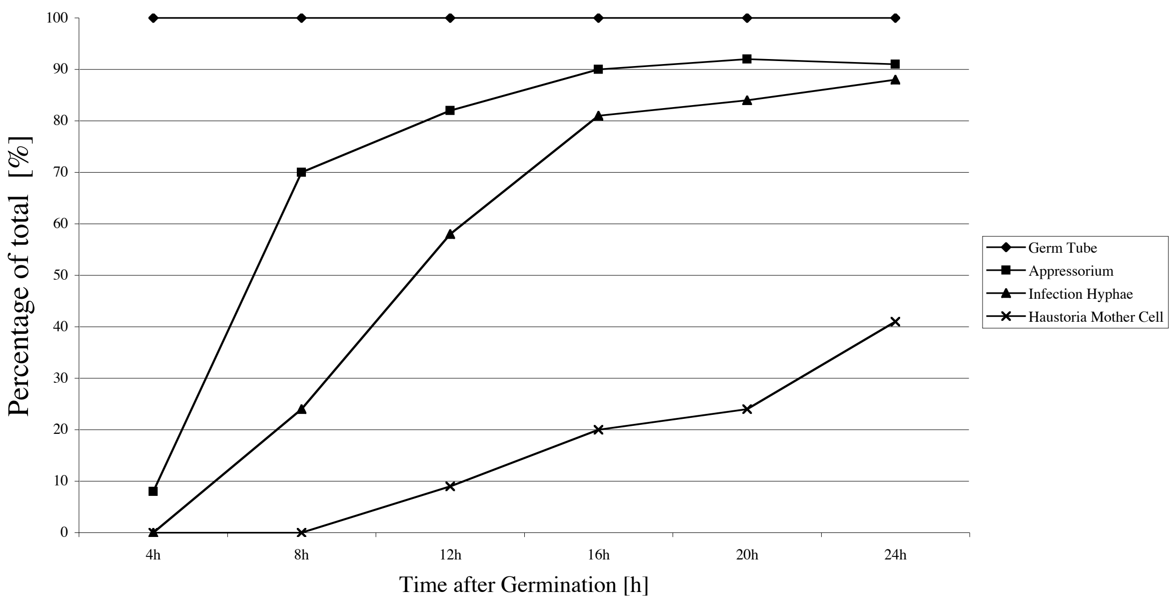


Figure 4-7 Time dependent percentage of *U. fabae* spores on polyethylene foil having reached or passed a certain differentiation stage, carried out according to Deising et al. (1991).

The above data served to verify the optimum times for harvesting specific infection structures and isolating the corresponding RNA and proteins.

4.8 Analysis of mRNA Transcription Profile

To verify that the studied *PIG*s are definitely *in planta* induced, the development stage-dependent mRNA transcription profile was analyzed using Northern blotting. The results are shown in Figure 4-8.

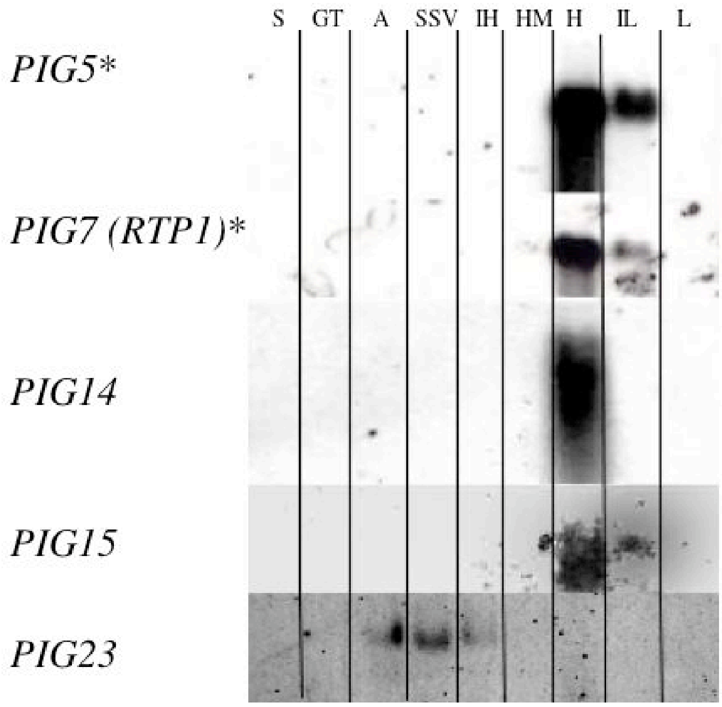


Figure 4-8 Transcription profile analysis via Northern blotting: 5 μg total RNA were applied per lane. DIG-labeled cDNA probes of the *PIG*s indicated on the left were used for hybridization. The lanes are labeled as follows: S: spore, GT: germ tube, A: appressorium, SSV: substomatal vesicle, IH: infection hypha, HM: haustorial mother cell, H: haustorium, IL: rust (*U. fabae*) infected bean leaf (*V. faba*), L: non-infected bean leaf. *Blots marked with an asterisk were kindly provided by Prof. Dr. M. Hahn

From the blot, it is clearly visible that *PIG5*, *PIG7*, *PIG14*, and *PIG15* are initially expressed in the haustorium stage (*in planta*!), as a clear hybridization signal is present for the first time in the H lane. RNA from infected *Vicia faba* leaves also shows a clear hybridization signal for *PIG5*, *PIG7* and *PIG15*, whereas the signal for *PIG14* is also present, but barely visible.

PIG23 does not seem to be an *in planta* induced gene, because hybridization signals are clearly visible in the A, SSV and IH lane (and are therefore not *in planta* induced), whereas the hybridization signal in the H lane is barley visible. Thus, *PIG23* was not studied any further during this thesis.

Finally, as expected, the RNA from uninfected leaves in the L lane (control lane) produced no signal for all *PIG*s, verifying that it did not hybridize with the DIG-labeled probes. Thus, all *PIG*s shown in Figure 4-8, with the exception of *PIG23*, are truly *in planta* induced.

4.9 Detection and Localization of *PIG* Products

4.9.1 *In vitro* and *in vivo* Detection of *PIG* Products

To detect the presence of the remaining PIG-proteins (PIGp 5, 7, 14, 15) in the different development stages, proteins from *in vitro* and *in vivo* infection structures were isolated and analyzed by Western blots (see chapter 3.9.4). In addition, to positively verify the existence and function of the predicted secretion signal in each of the above PIGps (see chapter 4.2.1), it was tested whether yeast would recognize these signals and secrete PIGp into the growth medium. For this test, each of the four *PIG*s was expressed in the heterologous system *Saccharomyces cerevisiae* using the yeast vector pDR195. Proteins from the resulting yeast pellets and supernatants were isolated and - in some cases – deglycosylated, and submitted to Western blotting, together with the abovementioned proteins from *in vivo* infection structures. All Western blots were carried out using polyclonal antibodies against the respective PIGp.

On the following pages, the Western blot results for the PIG-proteins in haustoria, in *in vitro* structures, as well as in the heterologous expression systems of *E. coli* and *S. cerevisiae,* are described.

In Figure 4-9 below, the Western blot for PIG7p detection in haustoria and heterologous expression are shown:

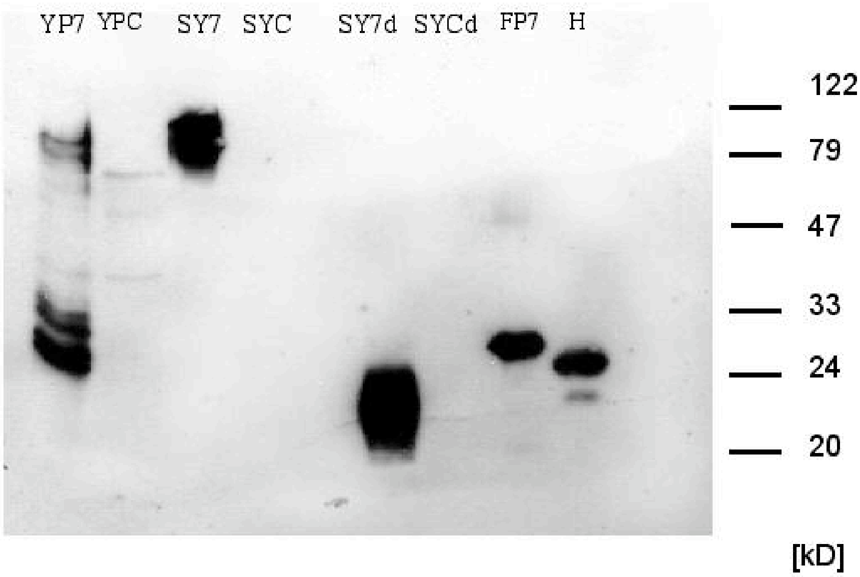


Figure 4-9 Protein detection by Western blot using polyclonal antibodies against PIG7p (RTP1p). The lanes are labeled as follows: YP7: pellet of *S. cerevisiae*:pDR195:PIG7 (20μg), YPC: pellet of S. cerevisiae:pDR195 (20μg), SY7: supernatant of *S. cerevisiae*:pDR195:PIG7 (5μg), SYC: supernatant of *S. cerevisiae*:pDR195 (5μg), SY7d: deglycosylated supernatant of *S. cerevisiae*:pDR195:PIG7 (5μg), SYCd: supernatant of *S. cerevisiae*:pDR195, FP7: fusion protein PIG7p (antigen) (0.1μg), H: haustorial protein extract

Four distinctive antibody cross-reaction signals (at ca. 25, 32, 80 and 90kDa) are visible in lane YP7, where protein extract from the *PIG7*-transformed yeast pellet was applied. The protein extracts from the non-deglycosylated supernatant (lane SY7) displays one strong signal (at about 100 kDa) that is most likely a combination of two signals similar to the upper two signals in lane YP7. Lane SY7d with the deglycosylated supernatant of the transformed yeast contains one strong signal at ca. 22 kDa. Thus, it can be concluded that the yeast *S. cerevisiae* recognizes the secretion signal of PIG7p.

In addition, we see that in contrast to the protein produced in yeast (lane SY7), the protein extract from haustoria (lane H) shows two smaller antibody cross reactivity signals at ca. 23.5 kDa for and at ca. 25.0 kDa. The smaller protein may be explained as a PIG7p pre-protein or a non-glycosylated form of the assumed mature protein. The signal for a larger protein in lane SY7 is due to the fact that the resulting heterologous protein in yeast is hyper-glycosylated compared to the original protein produced by the rust fungus in the haustorium.

The Western blotting results for PIG7p *in vitro* detection in are not shown, as no clear antibody cross-reactivity was detected in the *U. fabae* spores, and no antibody cross-reactivity at all were detected in *U. fabae in vitro* infection structures (2h, 4h, 20h, 24h) using antibodies against PIG7p.

Next, in Figure 4-10 below, the Western blot for PIG14p detection in haustoria and in heterologous expression systems is shown carried out with proteins extracted from transformed yeast, *E. coli* and haustoria:

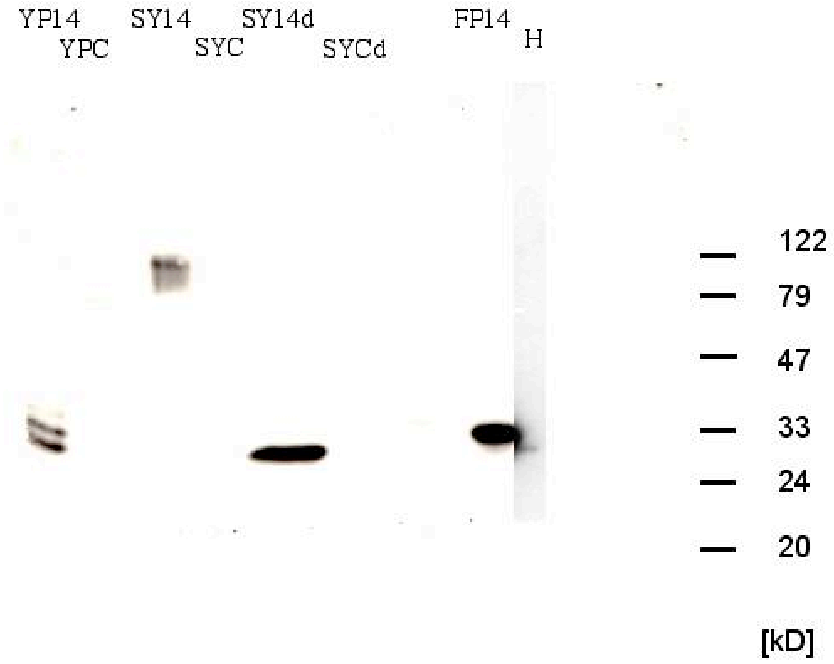


Figure 4-10 Protein detection by Western blot using polyclonal antibodies against PIG14p. The lanes are labeled as follows: YP14: pellet of *S. cerevisiae*:pDR195:PIG14, YPC: pellet of *S. cerevisiae*:pDR195, SY14: supernatant of *S. cerevisiae*:pDR195:PIG14, SYC: supernatant of *S. cerevisiae*:pDR195, SY14d: deglycosylated supernatant of *S. cerevisiae*:pDR195:PIG14, SYCd: supernatant of *S. cerevisiae*:pDR195, FP14: fusion protein PIG14 (antigen) (expressed in *E. coli*), H: haustorial protein extract

Two distinctive antibody cross-reactivity signals (at ca. 29 and 32 kDa) are visible in lane YP14, where protein extract from the *PIG14*-transformed yeast pellet was applied. The protein extract from the supernatant (lane SY14) shows one signal at ca. 120 kDa. In lane SY14d with the deglycosylated supernatant of the transformed yeast, one signal can be seen at 29 kDa. The conclusion that follows from these results is that the yeast *S. cerevisiae* recognizes the secretion signal of PIG14p.

In contrast to the protein produced in yeast (lane SY14), the protein extract from haustoria (lane H) shows one weak antibody cross-reaction signal at 29 kDa. As with PIG7p, this is due to the fact that the heterologous protein in yeast is more strongly deglycosylated than the original protein produced by the rust fungus in the haustorium.

The Western blot for PIG14p *in vitro* detection in are not shown, as no clear antibody cross-reaction signals were detected in the *U. fabae* spores, and no signals at all were detected in *U. fabae in vitro* infection structures (2h, 4h, 20h, 24h) using antibodies against PIG14p.

To positively verify the function of the predicted secretion signal in PIG15p, it was tested whether yeast would recognize this signal and secrete PIG15p into the growth medium. The results of the Western blot for PIG15p, carried out with protein extracts from transformed yeast, *E. coli*, as well as from haustoria, can be seen in Figure 4-11 below.

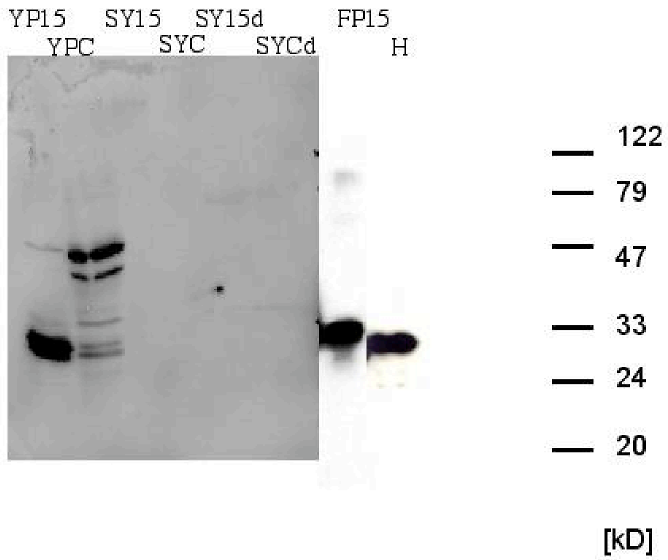


Figure 4-11 Protein detection by Western blot using polyclonal antibodies against PIG15p. The lanes are labeled as follows: YP15: pellet of *S. cerevisiae*:pDR195:PIG15, YPC: pellet of *S. cerevisiae*:pDR195, SY15: supernatant of *S. cerevisiae*:pDR195:PIG15, SYC: supernatant of *S. cerevisiae*:pDR195, SY15d: deglycosylated supernatant of *S. cerevisiae*:pDR195:PIG15, SYCd: supernatant of *S. cerevisiae*:pDR195, FP15: fusion protein PIG15p (antigen), H: haustorial protein extract

Two distinctive antibody cross-reaction signals (at about 29 and 32 kDa) are visible in the lane YP15, where protein extract from the *PIG15*-transformed yeast pellet was applied. Unexpectedly, two antibody cross-reaction signals (at about 45 kDa and 47 kDa) also occur in the lane YPC (control), were protein extract from the vector-transformed yeast was applied.

The protein extract from the supernatant (lane SY15) displays no signal, as does the deglycosylated supernatant of the transformed yeast (lane SY15d). The protein extract from haustoria (lane H) shows one clear antibody cross-reaction signal at about 31 kDa,

Whereas *S. cerevisiae* produces the *U. fabae* protein PIG15p, the secretion of PIG15p could not be verified with this Western blot. Several possible conclusions can be drawn from the above results. The results will be discussed in chapter 5.2.7.

Figure 4-12 below displays a second Western blot for PIG15p, showing the occurrence of PIG15p protein in *in vitro* structures and haustorium of *U. fabae:*

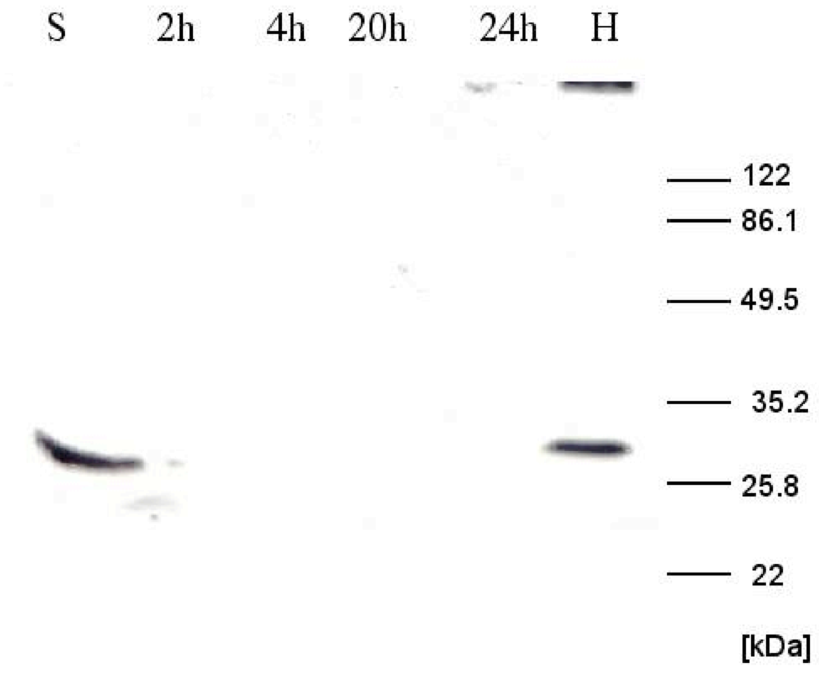


Figure 4-12 Protein detection in *in vitro* structures by Western blot of using polyclonal antibodies against PIG15p. 10 μg protein were applied per lane. The lanes are labeled as follows: S: spore, 2h: infection structure after 2h (germ tube), 4h: infection structure after 4h (germ tube/ appressorium), 20h: infection structure after 20h (infection hypha), 24 h: infection structure after 24h (haustorial mother cell), H: haustoria.

The protein-extract from haustoria (lane H) and the protein-extract from non-germinated spores display an antibody cross-reaction signal at ca. 31 kDa. No signals are visible in the lanes with protein extract from germinated spores (*in vitro* structures).

Unfortunately, for PIG5p, it was not possible to produce a valid Western blot for heterologous protein expression and protein expression in haustoria. The antibody cross-reaction resulted in an indistinguishable smear over large sections of the respective lanes, making the identification of a clear signal impossible (results not shown). In addition, in the Western blots for PIG5p *in vitro* detection, no clear antibody cross-reaction signals were detected in the *U. fabae* spores, and no signals at all were detected in *U. fabae in vitro* infection structures (2h, 4h, 20h, 24h) using antibodies against PIG5p (results not shown).

The following table (Table 4-9) provides an overview and comparison between the theoretically calculated and the experimentally derived protein masses (by Western blot) for of PIG7p, PIG14p and PIG15p*:*

Table 4-9 Protein Mass and Occurrence of Studied *in planta* Induced Genes

	PIG7p (RTP1p)	**PIG14p**	**PIG15p**
	[kDa]	[kDa]	[kDa]
Protein, calculated value	24.2	25.6	32.1
Protein without signal, calculated value (presumed mature protein)	22.1	23.7	29.6
His-tagged fusions-protein in *E. coli*	29	32	33
Transformed yeast (pDR195) pellet	25 / 32 80 / 90	29/32	30 / 32
Transformed yeast (pDR195) supernatant	100	120	Not detected
Transformed yeast (pDR195) supernatant deglycosylated	22	29	Not detected
Spore	-	-	31
2h infection structure	Not detected	Not detected	Not detected
4h infection structure	Not detected	Not detected	Not detected
20h infection structure	Not detected	Not detected	Not detected
24h infection structure	Not detected	Not detected	Not detected
Haustoria	23.5 / 25.0	29	31

4.9.2 Localization of *PIG* Products

In order to localize the gene products of the four remaining *PIG*s, slices of infected *V. faba* leaves were stained and examined using immuno-fluorescence microscopy. The primary polyclonal antibodies used in the protein-labeling process were the same as used for the Western blots described in chapter 4.9.1 above. The subsequent fluorescent staining with labeled secondary antibodies served to localize the gene products of the different PIGp *in situ*.

The following series of photographs (Figure 4-13) shows the results of the immuno-fluorescence microscopy:

Localization of Proteins with Potential Export Signals

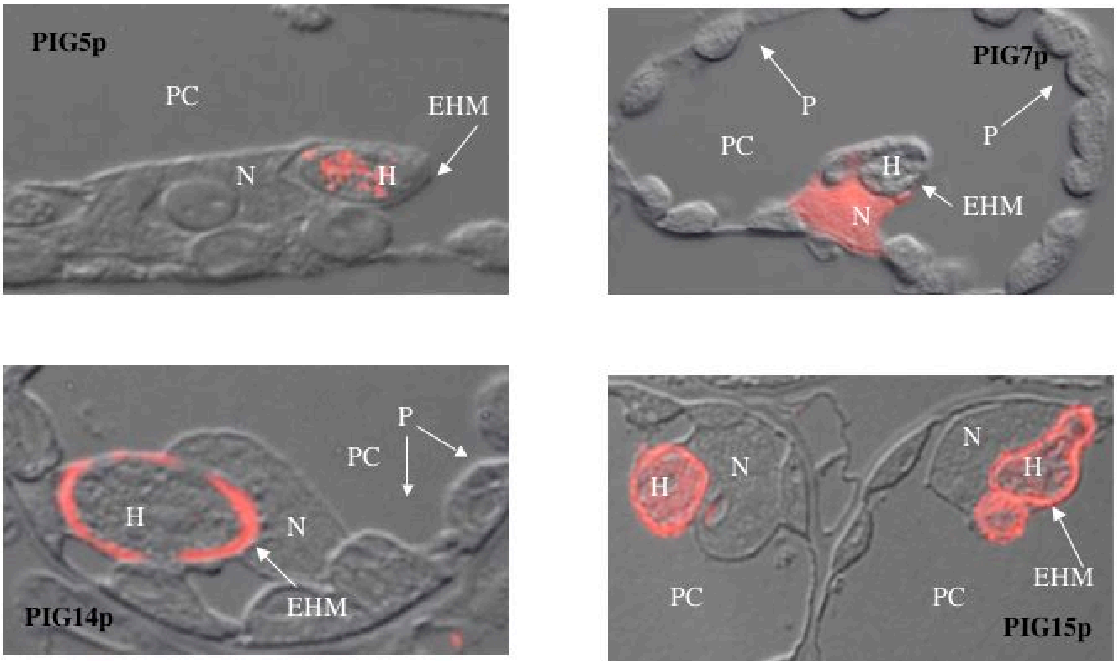


Figure 4-13 *In situ* localization of various PIGps in rust-infected bean leaves. The light-microscopy images of the plant structures are overlaid with fluorescence microscopy images of the indicated PIGp. (EHM): extrahaustorial membrane; (H): haustorium; (N): plant nucleus; (P): plastids; (PC): plant cell. Staining and photography by Prof. Dr. K. Mendgen and Dipl. Ing. (FH) H. Vahlenkamp.

As seen in the first photograph, PIG5p is located in the inside the haustoria. No staining of the extrahaustorial matrix or the host cells is visible.

As can be seen in the second and third photographs, the immuno-fluorescence staining clearly shows that PIG14p and PIG15p are both located in the extrahaustorial matrix at the fungal plant borderline. In both cases, the host cells are not stained.

As made visible in the fourth photograph, PIG7p (RTP1p) is located in the nucleus of the host *V. faba*. PIG7p can also be found in the extrahaustorial matrix und to a limited extend in the cytoplasm of the host, whereby none of the uninfected neighbor cells show a signal (results not shown).

The immuno-fluorescence staining provides strong evidence that PIG7p (RTP1p) is transferred from the rust haustoria into the nucleus of the host plant *V. faba.* This is the first time it has been shown that a plant pathogenic fungal protein is transferred to the nucleus of its host. The implications of this finding will be discussed in more detail in chapter 5.2.8.

5 Discussion

5.1 *U. fabae* EST Project Analysis

5.1.1 General Observations on the EST Project

The interest in fungal genes, which are expressed specifically in the biotrophic stage, has been aroused by the idea that such genes play a central role in pathogenesis. Spore germination, appressoria induction, the formation of penetration hyphae and the development of the haustorial mother cell, are all metabolically independent from the host plant, and draw on the spore energy reserves (Deising and Siegrist 1995). In contrast, the subsequent haustorium formation only takes place within the living plant cells, and possibly requires other, yet unknown, signals from the host (Brendel et al. 1992; Heath 1997). After invention of a new method to isolate the haustoria of *Uromyces fabae* (Hahn and Mendgen 1992), it was possible to isolate sufficient RNA to assemble a cDNA library of *U. fabae* haustoria (Hahn and Mendgen 1997). This library was the prerequisite for this EST project.

Genomic resources available to researchers studying phytopathogenic fungi are limited (Soanes et al. 2002). While the *Magnaporthe grisea* sequence is available at the Whitehead server, and cDNA databases of *Aspergillus nidulans, Aspergillus flavus* and *Neurospora crassa* are accessible at the university of Oklahoma (http://www.genome.ou.edu/fungal.html), there are still many fungal genomes to investigate. For example, genome data has been or is currently being acquired from a number of crop disease-causing fungi worldwide, including *Fusarium graminearum, Botrytis cinerea, Ustilago maydis, Blumeria graminis, Cochliobolus heterostrophus*, and *Septoria tritici*. Much of the data is held in inaccessible private sector databases, or has been deposited in public databases in an un-annotated form. Whereas not all of the above mentioned EST projects deal with biotrophic organisms, they all are fungal or fungal/host plant EST projects.

When looking at published fungal EST projects, it can be noted that most of them show a high percentage of unknown genes. For example, an EST project with expressed genes from wheat spikes infected with *Fusarium graminearum* yielded 49% unknown and 2% fungal genes, while the remaining genes were of plant origin (Kruger et al. 2002). A different cDNA library from a forced diploid culture of the smut fungus *U. maydis* growing as filaments, generated 7455 ESTs that were assembled into 3074 contiguous sequences. Of these *U. maydis* contigs, only 59% matched homologues in the NCBI and dbEST databases (Nugent et al. 2004). Finally, a cDNA library of *Eucalyptus globules-Pislithus tinctorius ectomycorrhiz*a contained 43% novel ectomycorrhizal genes (Voiblet et al. 2001). In contrast, ESTs from a mycelial library of the oomycete *Phytophthora infestans* (which is more similar to algae than to higher fungi) resulted in only 10% unknown genes (Kamoun et al. 1999).

As experienced during other published fungal EST sequencing projects (Györgyey et al. 2000; Thomas et al. 2001; Kruger et al. 2002), the largest portion of the contigs (310 contigs out of 541) studied during this EST project display no similarities to genes of known function.

5.1.2 cDNA Clone Redundancy

Of the 1032 cDNA clones resp. ESTs sequenced during this EST project, 409 are unique and the remaining 623 sequences (60%) are redundant, reappearing between 2 and 36 times (see below).

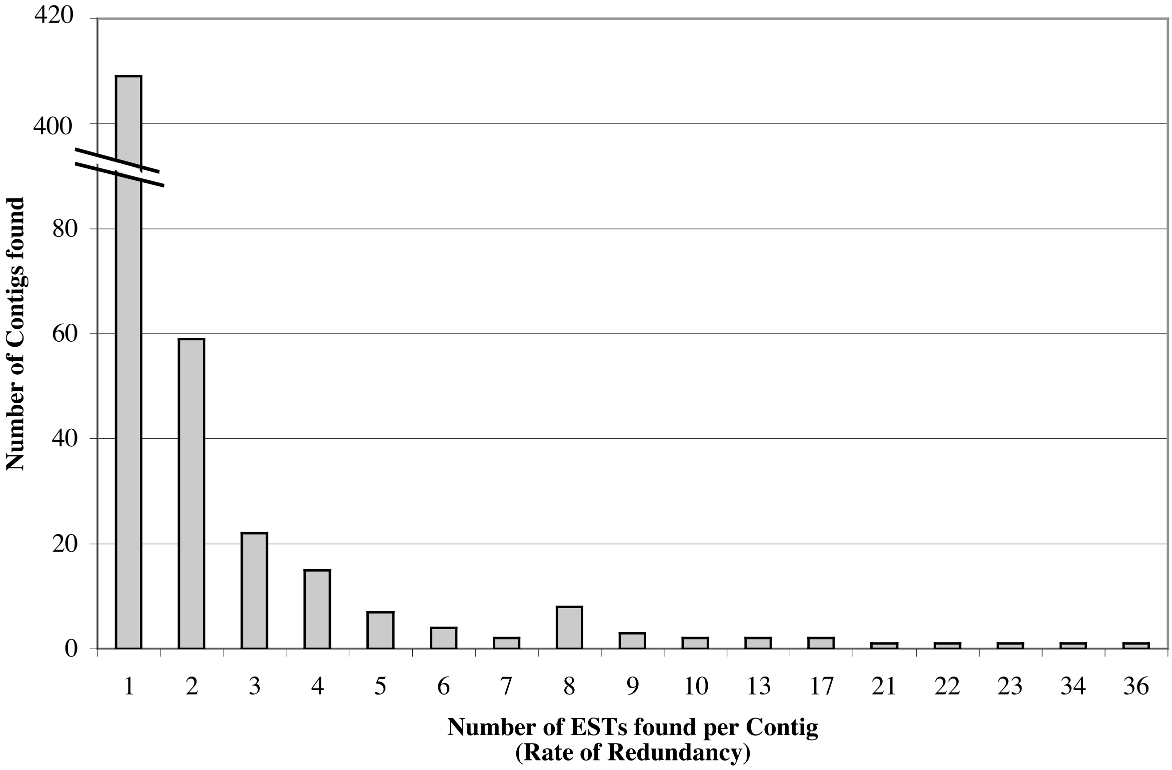


Figure 5-1 Rate of Redundancy for cDNA clones in the *Uromyces fabae* EST library based on contigs generated with SeqMan™. The number of multiple sequences found per EST is shown. The number of unique sequences is 409 (as shown in the first column).

It is notable that only a few genes of this EST project display a high redundancy. This suggests that considerably more than the sequenced 1’032 cDNAs are expressed in the biotrophic stage. This in turn is consistent with the knowledge that filamentous fungi have more than 10’000 genes (Galagan et al. 2003), and it is likely that several thousand genes are expressed during one stage (in this case the biotrophic stage). Thus, it becomes clear that there is still a great deal to uncover regarding gene expression during pathogenesis.

The rate of redundancy of *U. fabae* cDNA clones observed during this EST project is slightly higher than redundancy of genes obtained during other EST projects of similar scale (Nelson et al. 1997; Sterky et al. 1998; Thomas et al. 2001; Nugent et al. 2004). As mentioned above, the total redundancy of the 1’032 ESTs is 60%, which is higher than the rate of 37% achieved with *P. infestans* (1’000 clones) (Kamoun et al. 1999), the 49% achieved with *N. crassa* (1’409 clones) (Nelson et al. 1997), the 35% achieved with tissue of poplar, *Populus tremula*

L. x *tremuloides* Michx. and *Populus trichocarpa* 'Trichobel (5'692 clones) (Sterky et al. 1998), the 26% achieved with flower buds of Chinese cabbage, *Brassica campestris* L. ssp. *pekinensis* (1'216 clones) (Lim et al. 1996), and the 45% achieved with *Blumeria graminis* (4'908 clones) (Thomas et al. 2001).

About 20% of all cDNAs in the haustorial library sequenced during this EST project were found to be *in planta* induced (*PIG*s), which had been previously described by Hahn and Mendgen (1997). In addition, half of the most frequently found genes (six and more clones) belong to this group of *PIG*s. The most frequently found gene *THI1* (*PIG1*) (3,4% of the sequenced cDNAs code for this protein) codes for Thiamin (vitamin B1) synthesis (Hahn and Mendgen 1997; Sohn 2000; Sohn et al. 2000). The second most frequently found cDNA (3,3%) shows no similarities to genes of known function. BLAST searches for this gene and for all to date unidentified PIGs were performed without success against the NCBI database, as well as against several (partially incomplete) fungal databases for organisms such as *Neurospora crassa, Aspergillus fumigatus, Cryptococcus neoformans, Magnaporthe* grisea, *Blumeria graminis, Botryotinia fuckeliana, Fusarium sporotrichioides, Gibberella zeae, Mycosphaerella graminicola, Magnaporthe grisea, Phytophthora infestans, Phytophthora sojae* and *Verticillium dahliae* (for examples of available databases see http://www.cadre.man.ac.uk/Aspergillus_fumigatus/; http://cbr-rbc.nrc-cnrc.gc.ca/services/cogeme/; http://www.tigr.org/tdb/e2k1/cna1/ and Soanes et al (2002). This second most frequently expressed cDNA also corresponds to none of the *PIG*s previously described by Hahn and Mendgen (1997).

In a non-normalized cDNA-library, strongly expressed genes are generally found more frequently than others. Therefore, the frequency pattern of cDNAs found in the non-normalized haustorial library used for this EST project potentially reflects relative expression levels of the corresponding genes. The number of EST representing a certain *PIG* found during this EST project (see Table 4-1) in most cases correspond very well to the fraction of haustorial cDNA as determined by Hahn and Mendgen (1997), whereas several cDNAs (e.g. those encoding *PIG11, PIG13*) seem to be underrepresented.

One possible source of cDNA under-representation could be the exclusion of λ phage inserts smaller than 500 bp from the cDNA library, when it was originally created. As a result of this exclusion, two types of sequences could not be counted: those cDNA which are naturally shorter than 500 bp, and those mRNA or cDNA which have degraded or otherwise been reduced to below 500 bp during the transcription- or isolation process.

Non-full-length cDNAs result either from mRNA degradation, premature termination of reverse transcriptase during first strand cDNA synthesis, or DNA polymerase during second strand cDNA synthesis, or possibly DNA degradation before linker addition and ligation into the bacteriophage vector. A combination of these effects is also possible. Based on the fact that the protocol applied during this EST project for the purification of the haustoria used to create the cDNA library intrinsically leads to a substantial degradation of the mRNA (M. Hahn, personal communication), we can assume this to be the primary reason for the observed cDNA under-representation. In addition, other sources have also reported a weak correlation between the frequency of occurrence of a gene in an EST library and the actual transcript level (Thomas et al. 2001).

For large EST projects, it can make sense to not sequence redundant clones in order to reduce the required effort. Since the frequency of a gene in an EST library does not always reflect the actual transcript level, this approach may be deemed acceptable, depending on the goals of the project in question. A possible method is to filter out redundant clones before they are sequenced. One filtering method is the pre-selection of clones by oligonucleotide fingerprinting. This is an efficient high-throughput strategy to reduce redundancy in large-scale sequencing projects (Radelof et al. 1998). Another method to avoid the sequencing of redundant clones is the normalization and subtraction of cDNAs (Carninci et al. 2000; Clark et al. 2001).

Using the above methods, an opportunity to gain insight in the transcript levels is lost. Nonetheless, they represent useful approaches for reducing costs and - depending on the throughput - time. For the small scale EST project described in this treatise, redundancy-reducing methods were not applied, since the effort required for filtering out already sequenced clones would have been too great compared to the potential costs savings.

5.1.3 Homology to Other Fungi

Two important factors influence the amount and quality of homology found between the studied ESTs and genetic databases publicly accessible on the Internet. These are firstly the relationship in evolutionary terms between the two organisms whose genes are being compared, and secondly the number and type of organisms whose genes have been made available in the respective databases.

More than a fifth (20,6%) of the ESTs sequenced during this project exhibited their highest similarity to fungal sequences when blasted against the NCBI database. These ESTs actually make up 49.9% of the ESTs with homologues to other genes. This result was to be expected, because a number of fungal genomes such as *S. cerevisiae* (Goffeau et al. 1996) and *S. pombe* (Wood et al. 2002) have already been completely sequenced and are available for blasting at the NCBI database. The total number of ESTs whose best hit is a Basidiomycete gene is 138 (13.3%). This is not surprising, as *U. fabae* is a member of the Basidiomycota. A lower percentage (7.3%) of ESTs have their best homology to other fungal genes. This is simply due to the fact that the number of publicized non-basidiomycetal fungal genes on the NCBI server far exceeds those available for Basidiomycetes.

The NCBI blasting results described above are verified by the fact that 33.2% of the sequenced ESTs also have homology to genes in the public release of the *Magnaporthe grisea* genome at the Whitehead Institute server. *M. grisea* is a hemibiotroph and therefore has an initial developmental phase resembling biotrophic fungi.

5.1.4 Evidence for Proteases Produced in *U. fabae*

During this EST project, several genes with homology to proteases were found. These include three different putative metalloproteases, a putative cellular serine proteinase and several putative proteasome subunits. Because proteasomes primarily break down endogenous proteins such as transcription factors, cyclins, proteins encoded by viruses, other intracellular parasites and incorrectly folded proteins and are therefore unlikely to be involved in pathogenesis, these proteases are not discussed further.

The fact that intercellular hyphae and haustoria secrete lytic enzymes including proteases has been proven for many different fungi. For example, the necrotrophic fungus *Sclerotina sclerotiorum* secretes several proteases when grown in the presence of sunflower cell walls (Poussereau et al. 2001a; Poussereau et al. 2001b). The saprophytes *Aspergillus nidulans* and *Neurospora crassa* secrete enzymes against a broad spectrum of protein and carbohydrate substrates. This is indicative for their less specialized nutrition uptake. Broad-spectrum proteases are also secreted by *Verticillium albo-atrum* and *Verticillium dahliae* (St Leger et al. 1997).

Contig C545, coding for a putative serine protease, shows its highest homology to a subtilisin-related serine proteinase (ALP2) of *Aspergillus fumigatus*. Targeted disruption of the ALP2-encoding gene in *A. fumigatus* resulted in a slightly decreased rate of vegetative growth and in a more than 80% reduction of sporulation in the alp2-negative mutants (Reichard et al. 2000). This correlated with an approximately 50% reduction of the median diameter of conidiophore vesicles (Reichard et al. 2000). The high similarity (e-104) of the contig C545 to ALP2 suggests a correlation of this putative serine protease with the fitness and pathogenesis of *U. fabae*.

C210 shows similarity to a putative zinc metalloprotease from *Neurospora crassa* (E value 2e 42). Metalloproteases have been described to be involved in pathogenesis of various microbiological pathogens (Joshi and St Leger 1999; Miyoshi and Shinoda 2000; Brouta et al. 2002). The contigs C123 and C569 have similarities to an extracellular metalloprotease of *A. fumigatus*. *A. fumigatus* secretes a serine alkaline protease (ALP) and a metalloprotease (MEP) when the fungus is cultivated in the presence of collagen as sole nitrogen and carbon source. An alp mep mutant, deficient in proteolytic activity at neutral pH *in vitro*, showed no difference in pathogenicity to the wild-type strain, suggesting that ALP and MEP are not essential for the invasion of the lung tissues by *A. fumigatus* (Jaton-Ogay et al. 1994). Previous studies of the extracellular proteases of the rust fungus *Uromyces fabae* gave hints regarding a metalloprotease involved in the pathogenesis of the fungus (Rauscher et al. 1995). In consequence, no conclusion between the correlation of the gene product of the contigs C210, C123 and C569 and the fitness and pathogenesis of *U. fabae* can be drawn, except that metalloproteases may be involved in the infection process.

It must be noted that not all proteases necessarily have an influence of fungal pathogenicity. For example, targeted inactivation of prt1 in *F. oxysporum* f. sp. *lycopersici* had no detectable effect on mycelial growth, sporulation, and pathogenicity in tomato plants (Di Pietro et al. 2001). In addition, 3 extracellular serine proteases from *Cochliobolus carbonum* were purified and characterized (Murphy and Walton 1996). Using transformation-mediated gene disruption, *C. carbonum* mutants were created, but the disease symptoms on maize were identical to the growth and disease symptoms of the wild type (Gorlach et al. 1998).

In contrast to those examples of proteases of phytopathogenic fungi where the essential function could not yet be proven, there are some examples of proteases of *Candida* ssp. with known essential function (Lee et al. 2001; Felk et al. 2002). For instance, *Candida albicans* mutants lacking different secreted aspartic proteinase (SAP) have been analyzed (Felk et al. 2002). Mutants lacking SAP1 to SAP3 had invasive properties indistinguishable from those of wild-type organisms. In contrast, a triple mutant lacking SAP4 to SAP6 showed strongly reduced invasiveness but still produced hyphal cells. When comparing the damage caused by wild-type cells and several different mutants, all mutants who lacked functional SAP6 showed significantly reduced liver and pancreas tissue damage in the host organism. These data demonstrate that strains, which produce hyphal cells but lack hyphae-associated proteases, particularly, those encoded by SAP6, are less invasive (Felk et al. 2002). Even though *Candida albicans* is a human-pathogenic yeast and not closely related to *U. fabae*, the fact that mutants of *Candida albicans* show a reduced pathogenesis leads to the conclusion that the concerted action of different proteases may influence the pathogenicity and fitness of phytopathogenic fungi as well.

Based on the above-mentioned studies of various pathogenic fungi, it can be concluded that proteases in general play a decisive role in pathogenesis, although not all proteases may be required for host infection.

5.1.5 Evidence for Genes Involved in Amino Acid Metabolism and Transport

Once the fungus has infected its host and the biotrophic phase has been established, it must absorb nutrition in order to live and reproduce (Agrios 1997; Prell and Day 2001). Biotrophic plant pathogens acquire their nutrients from living host cells, with the haustorium playing a decisive role (Mendgen and Hahn 2002). Among the nutrients required are amino acids. Several amino acid transporters have been characterized in *U. fabae* and their function has been thoroughly discussed (Hahn and Mendgen 1997; Martin 2001; Müller 2001; Struck et al. 2002; Struck et al. 2004). The *U. fabae* amino acid transporter AAT1p displayed high transport activities for lysine and histidine (Struck et al. 2002) and the amino acid transporter AAT3p shows substrate preferences for leucine, methionine and cysteine (Struck et al. 2004). In addition to the amino acids absorbed from the host plant, the fungus must produce certain amino acids on its own, because the amino acid transporters are usually substrate specific (Struck et al. 2002; Struck et al. 2004). In the EST project described in this treatise, several genes coding for putative amino acid synthesis or degradation have been found (Table 5-1, Table 5-2).

Table 5-1 Genes Putatively Involved in Amino Acid Synthesis

Contig	Enzyme putatively coded for	EC No.	Resulting Amino Acid	E Value
C398	Prephenate dehydrogenase	1.3.1.13 (synonyme)	Tyrosine	2e-47
C040	Delta 1-pyrroline-5-carboxylate reductase	1.5.1.2 (synonyme)	Proline	6e-64
C037	Imidazoleglycerol-phosphate dehydratase (IGPD)	4.2.1.19	Histidine	2e-50
C127	Saccharopine dehydrogenase	1.5.1.10	Lysine/ L-glutamate	1e-55
C170	Glutamine synthetase	6.3.1.2	Glutamine	4e-73
C308	Threonine synthase	4.2.99.2/ 4.2.3.1	Threonine	2e-10
C206	Homoserine dehydrogenase	Q96VZ7 (primary accession no. at ExPSAy, possible synonyme 1.1.1.3)	Threonine, Isoleucine, Methionine	3e-29
C462	Adenosylhomocysteinase	3.3.1.1 (synonyme)	Adenosine, Homocysteine Methionine	9e-59

Interestingly, the contigs C037, C127, C206 and C462 code for proteins with homology to enzymes involved in the synthesis for amino acid for which amino acid transporters in *U. fabae* have already been described (Struck et al. 2002; Struck et al. 2004). A possible conclusion is that these enzymes are involved in degradation or alteration of the amino acids in question.

The expected function of the putative homoserine dehydrogenase (C206) and the putative adenosylhomocysteinase (C462) is to take part in the biosynthesis of methionine, threonine and isoleucine respectively adenosine and homocysteine. Because methionine is transported by AAT3p from the host to the fungus (Struck et al. 2004) it can be assumed that the putative enzymes are at least involved in the biosynthesis of threonine and isoleucine or adenosine and homocysteine respectively.

One function of the putative saccharopine dehydrogenase in *U. fabae* beside the biosynthesis of lysine can be the degradation of excess lysine and the forming of L-glutamate. The function of the putative imidazoleglycerol-phosphate dehydratase (IGPD) in *U. fabae* would be the catalyzation of D-erythro-1-(imidazol-4-yl)glycerol 3-phosphate to 3-(imidazol-4-yl)-2-oxopropyl phosphate + H_2O. This is the sixth step in the histidine synthesis (starting from phosphoribosyl pyrophosphate (PRPP)). It is questionable whether the fungus, which is able to extract histidine from the plant, produces additional histidine. Indeed, IGPD, also participates in the synthesis of L-glutamate and 2-oxo-gluterate (Michal 1993).

These findings described above, underline the fact that despite the uptake of nutrients from the host cells, many amino acids required by plant pathogens are synthesized by the pathogen itself. This is consistent to other research results. For example, genes involved in histidine and methionine biogenesis have been identified in *Magnaporthe grisea,* using REMI (Restriction Enzyme Mediated Integration) (Sweigard et al. 1998; Balhadère et al. 1999). These results may suggest that the host does not supply sufficient amounts of all metabolites required by the fungus.

In addition to finding putative genes for amino acid synthesis, evidence for the presence of genes presumed to be involved in amino acid degradation was also found during the EST project (see Table 5-2).

Table 5-2 Genes Putatively Involved in Amino Acid Degradation

Contig	Enzyme putatively coded for	EC No.	Degraded Amino Acid	E Value
C554	Delta-1-pyrroline-5-carboxylate dehydrogenase precursor p5c	1.5.1.12	Proline	4e-60
C069	GABA transaminase	2.6.1.19	Glutamine	2e-32

In summary, it can be stated that the existence of amino acid transporters and putative genes involved in amino acid synthesis and degradation in *Uromyces fabae* haustoria provides strong evidence for the assumption that haustoria play a central role in the anabolic and catabolic metabolism of the rust fungus.

5.1.6 Evidence for Genes Involved in Vitamin Synthesis

In addition to amino acids, vitamins also seem to be essential for the life cycle of the rust fungus. Two of the *PIG*s most frequently found during the EST project, *PIG1* (= *THI1*) and *PIG4* (= *THI2*), are related to vitamin B1 synthesis and have been previously found to be highly transcribed in haustoria (Hahn and Mendgen 1997; Sohn 2000; Sohn et al. 2000). Vitamin B is necessary for metabolism of sugar and starch to provide energy (Bender 2003). The function of *THI1* and *THI2* in *U. fabae* have been discussed in detail (Sohn 2000; Sohn et al. 2000). Sohn et al. (2000) came to the conclusion that: (1) *U. fabae* does not extract thiamine from its host and therefore synthesizes thiamine in it's own right; (2) The synthesis of thiamine is presumably stage-specific; (3) not only do haustoria of *U. fabae* extract nutrition from the host plant, they are also involved in biosynthesis (Sohn 2000; Sohn et al. 2000). The number of amino acid biosynthesis-encoding genes (see chapter 4.1) found during the described EST project is supportive of Sohn et al.'s third conclusion.

5.1.7 Evidence for Genes Involved in Signaling in Rust Pathogenesis

Signaling plays an important role in all cellular processes. One example of a gene involved in signaling in the *U. fabae* haustoria is contig C287. C287 displays high homology to calmodulin from *Drossphila melanogaster* (BLASTX E value = 5e-55). Calmodulin acts as a calcium sensor and is involved in the regulation of the activity of several proteins. Calmodulin has also been isolated from *Uromyces appendiculatus*, a close relative of *U. fabae* (Laccetti et al. 1987).

Inhibition of a calmodulin-dependent kinase in *Colletotrichum gloeosporioides* and *Magnaporthe grisea* inhibits germination and appressorium formation, suggesting that calmodulin is necessary for this process (Kim et al. 1998; Lee and Lee 1998; Liu and Kolattukudy 1999). As germination in *Colletotrichum gloeosporioides* and *Magnaporthe grisea* starts after hard surface contact, it can be deduced that hard surface contact actually induces calcium-calmodulin signaling (Kim et al. 1998; Lee and Lee 1998; Liu and Kolattukudy 1999). Homologues to calmodulin and calmodulin-dependent kinase were also identified in *Blumeria graminis* (Thomas et al. 2001).

The function of calmodulin in *Uromyces fabae* remains unclear at present, but since this protein has been shown to be necessary in other phytopathogenic fungi, some relevant function can be expected.

5.1.8 Evidence for Genes Involved in Stress Response in Haustoria

Pathogen infection induces numerous changes in the host's cellular metabolism and structure, which are part of the plant's defense strategy. Besides more common defense reactions such as reinforcement of cell walls, there are many examples of more dramatic reactions, such as modifications to the host cell cycle and the shut-off of selected host genes to avoid metabolic competition between host and pathogen (Prell and Day 2001). Thus, not only does the host plant have to react to the stress induced by the pathogen, but the fungus must also react to stress induced by its host. Several fungal genes putatively coding for stress response were found during this EST project. Two of them, the putative cytochrome P450 monooxygenase (*PIG16*) and the putative cyclosporin-binding protein (*PIG28*) have previously been described (Hahn and Mendgen 1997).

Functions of P450 enzymes are extremely diverse. In fungi, they include synthesis of membrane sterols and mycotoxins, detoxification of phytoalexins, and metabolism of lipid carbon sources (Werck-Reichhart and Feyereisen 2000). Next to these functions it has been shown that a P450 monooxygenase is essential for abscisic acid biosynthesis, which is thought to be involved in host-pathogen interaction in the phytopathogenic fungus *Botrytis cinerea* and its host (Siewers et al. 2004).

The immunosuppressive drug cyclosporin A (CsA) targets cyclophilins (Viaud et al. 2002). Cyclophilins are peptidyl prolyl cis-trans isomerases that are highly conserved in eukaryotes. CsA forms a complex with cyclophilin A and inhibits the calmodulin-dependent phosphoprotein phosphatase calcineurin. By using targeted gene replacement to investigate the role of CYP1, a cyclophilin-encoding gene in *Magnaporthe grisea*, it was found that cyclophilin acts as virulence determinant during plant infection (Viaud et al. 2002).

In addition, four different putative heat shock proteins were found during this project (contigs C209, C417, C282, C511). Heat shock proteins (HSPs) are a group of proteins that are present in all cells of all life forms. HSPs are strongly induced when a cell is subjected to environmental stress such as heat, cold and oxygen deprivation, but are also present in cells under normal conditions. HSPs help nascent or distorted proteins to fold into shape, which is essential for their function. They also shuttle proteins from one compartment to another inside the cell, and help to remove spent proteins from the cell. It remains unclear whether the putative HSPs found during this EST Project are expressed due to ambient environmental stress, stress exerted from the host plant on the fungus, or if they are continually present in the haustorium in their function as chaperones for proteins.

It is not surprising that the above-described putative stress response-related genes of *U. fabae* are found in the haustorium. Once the fungus has overcome the plants so-called basic resistance (Prell and Day 2001), it has to live in a potentially hostile environment. By effectively suppressing the host defense (using mechanisms which are today largely unknown), the fungus can substantially reduce the threat to its survival and maintain a state of compatible interaction.

5.1.9 Evidence for Viral Genes

Although the broad bean plants (*Vicia faba*) used for isolation of *U. fabae* haustoria during this EST project showed no visible virus infection (personal communication M. Hahn), some evidence for the expression of viral genes was found (see Table 4-5). Since all of the putative viral genes found during this project are homologous to genes of plant viruses, a viral infection of the bean plants seems to be in evidence. On the other hand, there are three circumstances that speak for a viral infection of *U. fabae* by a mycovirus instead:

- The similarity of the newly discovered putative viral genes to plant viruses remains quite low (the lowest E values is 7e-25).
- During this EST project, only one putative plant cDNA was found (result not shown). This suggests that the fungal cDNA library is relatively clean and plant cDNA contaminations have been kept to a minimum (under 0.1%).
- As mentioned above, the bean plants used during this EST project showed no visible virus infection.

Thus, the viral genes isolated most likely stem from fungal cDNA.

A wide range of other research on phytopathogenic fungi supports the assumption that there is indeed a mycoviral infection of *U. fabae*. The most well known viruses in plant pathogenic fungi may be the *Cryphonectria parasitica* hypoviruses CHV 1-3. Other examples are viral infections of the rust *Puccinia recondita* (Zhang et al. 1994) and *Ustilago maydis* (Peery et al. 1987). A great number of additional viral sequences in fungi and protozoa can be found in the Internet under http://www.dpvweb.net/seqs/fungusviruses.htm.

For rust fungi, virus-like-particles (VLP) have been previously described (Dickinson and Pryor 1988b; Dickinson and Pryor 1988a; Dickinson et al. 1989; Zhang et al. 1994). These VLP can be represented by dsRNA (Dickinson and Pryor 1988b). Most rusts contain large amounts of dsRNA, whereby the number and size is highly species-specific (Dickinson et al. 1989). As a matter of fact, in the *U. fabae* I2 isolate used during this thesis, a preparation of undigested DNA separated on an agarose gel, often showed a band at about 5 kb. This probably represents a dsRNA as mentioned above (personal communication M. Hahn). Some of this dsRNA might form virus-encoded proteins.

Regarding the location of VLPs, Zhang et al. (1994) report that virus-like-particles are located in the extrahaustorial matrix of *Puccinia recondita*, which was confirmed by immuno-gold staining.

Concerning the interaction between fungus and virus, research has shown that the presence of dsRNAs in rust is not associated with fungal disease or alteration of virulence (Zhang et al. 1994). On the other hand, studies of the corn smut fungus *Ustilago maydis* have shown that the *U. maydis* P4 strain secretes a fungal toxin, KP4. KP4 is encoded by a fungal dsRNA virus (UMV4), that persistently infects *U. maydis* cells. UMV4 does not spread to uninfected cells by release into the extracellular milieu during its normal life cycle, and is thus dependent upon host survival for replication. In symbiosis with the host fungus, UMV4 encodes KP4 to kill other competitive strains of *U. maydis*, thereby promoting both host- and virus survival (Peery et al. 1987; Gu et al. 1995).

The uncovered evidence of mycoviral infection of *U. fabae* opens the door for further questions: What kind of mycovirus is present, and how, if at all, does the infection affect the metabolism or pathogenicity of *U. fabae*? Does *U. fabae* possibly benefit from this assumed infection? Are there signs of enhanced fitness compared to other *U. fabae* strains? These questions and more may be the subject of future phytopathological studies.

5.2 Characterization of Highly Expressed *U. fabae PIG*s with a Signal Sequence

5.2.1 Validity of Data

The possibility of analyzing gene- and protein sequences with publicly available bio-informatics tools constitutes a substantial improvement in molecular biological research. The analysis of the studied *PIG*s with bio-informatics tools more quickly permits a deeper insight in the potential function of those gene-products than previously possible with conventional methods. Nevertheless, without experimental data, these theoretically calculated results are difficult to interpret. Thus, a cautious approach to interpretation is advisable in order to avoid an overvaluation of the achieved results.

5.2.2 PIG-Protein Stability

One of the characteristics of the studied PIG-proteins, that was theoretically analyzed and observed under laboratory conditions, was their stability.

The theoretical stability analysis was carried out primarily in order to verify whether the instabilities experienced during laboratory work were due to experimental conditions or resulting from specific sequence characteristics. This analysis was based on prior research which has shown that proteins which contain specific sequence characteristics can be expected to be more unstable than other (Guruprasad et al. 1990).

In addition, a general relationship between protein stability and protein function can be expected (Shoichet et al. 1995). It is known that proteins which recognize ligands have active sites with hydrophobic characteristics on their outer side (Pauling et al. 1943; Warshel 1978; Warshel et al. 1988). On the other hand, in the interest of stability, the proteins minimize their free energy by folding into ordered structures with a hydrophobic interior and a hydrophilic exterior (Kauzmann 1959). Thus, protein residues which contribute to catalysis or ligand binding are not optimal for protein stability (Shoichet et al. 1995; Beadle and Shoichet 2002). With reference to the PIGps studied during this thesis, it can be inferred that a reduced stability may indicate the presence of a binding site (Beadle and Shoichet 2002).

Two different methods were used to theoretically predict PIGp stability during this thesis. These were the “N-end rule” and the “Estimated Half-Life and Stability Index”, both available on the ExPasy server (http://au.expasy.org/).

According to the "N-end rule" algorithm (Tobias et al. 1991), all six PIG-proteins were predicted to have a half-life of > 20 h in yeast and > 10 h in *E. coli, in vivo*, which would indicate that they are stable. Based on the "Estimated Half-Life and Stability Index" (Guruprasad et al. 1990), only PIG5p and PIG15p were predicted to be stable with values of 36.98 respectively 34.83. PIG7p (RTP1p) and PIG14p slightly exceeded the cut-off value of 40 with results of 41.5 and 43.76 respectively, whereas PIG9p and PIG23p both generated a very poor theoretical stability index at 56.71 and 57.37 respectively. The results for both protein stability prediction methods compared to the experimental results are displayed in the following table.

Table 5-3 Comparison of PIG-Protein Stability with Predictions

	Experimental Results	"Estimated Half-Life and Stability Index", (Guruprasad et al. 1990)	"N-end rule" algorithm, (Tobias et al. 1991),
PIG5p	stable	stable	stable
PIG7p	stable	unstable *	stable
PIG9p	unstable	unstable	stable
PIG14p	stable (low yield)	unstable*	stable
PIG15p	stable	stable	stable
PIG23p	unstable	unstable	stable

*just above the cut-off value of 40 and therefore just barely classified as unstable

As correctly predicted by both theoretical methods, PIG5p and PIG15p were stable for some hours to days under laboratory conditions. PIG7p (RTP1p) and to a lesser extent PIG14p also proved to be stable under laboratory conditions. While the "Estimated Half-Life and Stability Index "rating was only slightly above the cut-off value for stable proteins (Guruprasad et al. 1990), the "N-end rule" (Tobias et al. 1991) algorithm again correctly predicted the stability of these two proteins. Finally, PIG9p and PIG23p were not stable under the applied laboratory conditions, as correctly predicted by the "Estimated Half-Life and Stability Index", but not the "N-end rule".

From the comparison between the theoretically predicted stability of the proteins and the observations in the laboratory, it can be concluded that the "N-end rule" algorithm (Tobias et al. 1991), although applicable for bacteria (where the PIG-proteins were expressed), is not valid for the PIG-proteins studied during this work. This method incorrectly predicted all proteins, without exception and differentiation, to be stable for more then 20 h in yeast and more then 10 hours in *E. coli* (see Table A-2). The stability prediction using the "Estimated Half-Life and Stability Index" (Guruprasad et al. 1990), seems to be considerably more precise, assuming that the cut-off value of 40 is not applied too rigidly.

The differing stability of the studied proteins allows speculation on whether proteins with a short lifespan are actually designed to have a short working life, or if, because of their function, e.g. catalysis or ligand binding, they have a structure sub-optimal for long-term stability, as described by Shoichet et al. (1995). To find answers to these and other intriguing questions, further research into the function of these PIG-proteins is required.

5.2.3 Secretion and Localization of PIG-Proteins

5.2.3.1 Analysis of Theoretical and Experimental Secretion Behavior and Localization

A further important characteristic that was studied during this thesis was the secretion behavior and subsequent *in situ* localization of the PIG-proteins. As explained in chapter 2.5.1, a (theoretically predicted) signal sequence was the prerequisite for the selection of the *PIG*s to be studied, as a signal sequence indicates that the associated protein is secreted into the host plant and may therefore contribute to fungal-plant interactions.

While all cDNA clones of *PIG*s analyzed during this thesis were predicted to encode secreted proteins using various algorithms (see Table 4-7), the theoretical prediction of their localization generated very different and partially contradictory to the secretion predictions. As described below, practical verification or rebuttal of the theoretically derived results for 4 of the 6 *PIG*s could be achieved through experimental work.

As described in chapter 4.9.1, it was not possible to verify the predicted secretion of PIG5p via Western blot for technical reasons. The initial antibody resulted only in unspecific antibody cross-reaction signals. A second antibody against a different PIG5p antigen was then applied. After purification, this second antibody invariably produced weak signals with the control protein (results not shown). Therefore no Western blot results for PIG5p are shown.

As a matter of fact, despite the circumstance that the NCNN prediction algorithm within the PSORT II database application predicted a nuclear localization (76.7% reliability) for PIG5p, and the k-NN algorithm within the same application predicted extracellular localization (44.4% probability), the immunofluoroscence pictures show that this protein is not located in the extrahaustorial matrix, but rather inside the haustorium itself. This indicates that PIG5p may not be secreted at all. Further studies confirmed that PIG5p is localized in the ER (personnel communications E. Kemen).

The Western blot results of PIG7p (RTP1p) (Figure 4-9) clearly show that the yeast *S. cerevisiae* recognizes the theoretically predicted secretion signal of *PIG7*, verifying that PIG7p is truly a secreted protein. The two antibody cross reactivity signals at ca. 23.5 kDa for and at ca. 25.0 kDa in the protein extract from haustoria (lane H) may be explained as two different glycosylated forms of the mature PIG7p.

Curiously, the NUCDISC and NNCN prediction algorithm within the PSORT II database application for PIG7p localization seems to contradict the secretion predictions and experimental results. Despite the fact that the protein is secreted, it was also predicted by NUCDISC to have an NLS (70.6% reliability), and by NCNN to be nuclear (30.4% probability). At first glance, it does not seem likely that a protein which is secreted from a cell, which should be located in the nucleus of that cell.

Fortunately, both the secretion and localization predictions regarding PIG7p (RTP1p) are explicable and reconcilable when looking at further experimental data. By using immuno-microscopy, it was again shown that PIG7p (RTP1p) is secreted from the fungus, and also that it can be located in the <u>nucleus of the host plant (</u>Figure 4-13). Thus, the NLS of PIG7p relates not to its parent organism, but to its host instead.

The above results demonstrate that a fungal protein can be structured in a way that it is secreted and then imported to the nucleus of a totally different species.

Because of the poor stability of PIG9p under laboratory conditions, no experiments regarding protein secretion or localization could be performed using this protein. Nonetheless it can be noted that the theoretical analysis for PIG9p showed very similar results to PIG7p regarding secretion, presence of an NLS, and nuclear localization. Thus, there is a possibility that PIG9p has similar properties to PIG7p regarding host nucleus localization. For further studies it would be of interest to generate a stable PIG9p antigen, by excluding unstable regions of the protein, and to proceed with experiments analog to those done with PIG7p.

For PIG14p, the theoretically predicted secretion signal was verified by the Western blot results (Figure 4-10). The transformed yeast recognized the secretion signal, and a strongly glycosylated form of PIG14p was found in the supernatant of the *S. cerevisiae* culture. As can be seen in the Western Blot for PIG7p (Figure 4-9), one could expect to see two antibody cross-reaction signals for PIG14p in the Haustoria lane. Unfortunately, the single antibody cross-reaction signal for PIG14p in the Haustoria lane was weak and a second signal was not visible. This may be due to a to small amount of haustorial protein in the Western blot. An alternate explanation for the existence of only one antibody cross-reaction signal in the H lane could be that the signal sequence was not cleaved from the protein, although this is not a common occurrence (Muesch et al. 1990). The most likely explanation is that only one form of PIG14p exists.

While the NNCN prediction algorithm indicated that PIG14p might be cytoplasmic (76.7% reliability), the k-NN algorithm indicated a 77.8% probability that PIG14p is extracellular. The k-NN prediction matches well with the immunofluoroscence analysis, which shows PIG14p to be located in the extrahaustorial matrix (see Figure 4-13).

Interestingly, for PIG15p, we find antibody cross-reaction signals at about 31 kDa in both the protein extract from spores and from haustoria. Two distinctive antibody cross-reaction signals (at 30 and 32 kDa) were visible in the lane where protein extract from the *PIG15*-transformed yeast pellet was applied. In addition, two antibody cross-reaction signals occurred at about 45 kDa and about 47 kDa in the lane were protein extract from the vector-transformed yeast (control) was applied.

One explanation for the antibody cross-reaction signals in the control lane (YPC) could be that these signals result from unspecific reaction with parts of the yeast. Another possible explanation is that a structural homologue to PIG15p exists in yeast. Because no significant BLASTX matches between *PIG15* and *S. cerevisiae* (which is fully sequenced) were found, it is most likely that no structural homologues exist and the antibody cross-reaction signals are unspecific. The lack of strong reactions in lane YP15 at the same level as in the control lane (YPC) can be due to a preferred reaction of the antibody with its antigen. For further discussion of the development-stage depended expression of PIG15p, see chapter 5.2.7.

In contrast to PIG14p, the theoretically predicted secretion signal for PIG15p could not be verified by Western blot (Figure 4-11). Both the protein extract from the supernatant as well as the deglycosylated supernatant of the transformed yeast did not exhibit an antibody cross-reaction signal. On the other hand, the immunofluoroscence pictures show that PIG15p is located in the extrahaustorial matrix, so it is possible that PIG15p is secreted. The failure to detect the PIG15p secretion by Western blot may be due to several reasons. One possibility is that the yeast *S. cerevisiae* does not recognize the secretion signal of PIG15p. The second possibility is that the gene product of *PIG15* is too large to be secreted by *S. cerevisiae*.

As had been the case with PIG9p, the poor stability of PIG23p under laboratory conditions prevented the performance of experiments regarding protein secretion or localization using this protein. As a matter of fact, as will be discussed later in chapter 5.2.7, *PIG23* turns out not to be an *in planta* induced gene after all, but rather is induced during *in vitro* stages of infections structures (see Figure 4-8).

It may be noted that despite the fact that the secretion and the localization in the extrahaustorial matrix could be shown for PIG7p (RTP1p) and PIG14p, and the localization in the extrahaustorial matrix could be shown for PIG15p (see Figure 4-13), it remains unclear how the secretion takes place. Within the framework of this thesis, no evidence for a fungal secretion system similar to the ones found in bacteria (Hueck 1998; Bonas and Van den Ackerveken 1999; Fischer et al. 2002; Ghosh 2004; Gomis-Ruth et al. 2004), has been found in the public databases or *in vivo*.

5.2.3.2 *Evaluation of the Applied Algorithms*

When comparing the theoretically predicted localization and secretion behavior of the studied PIG-proteins generated by the various algorithms (NUCDISC, NNCN and K-NN) described in chapters 3.11.4, 3.11.5, and annex 8.2 with the actual experimental results, it becomes evident that not all of the these algorithms provide reliable results for the fungal organism studied during this thesis.

Analyzing the NUCDISC results, we see that this algorithm predicted PIG7p (RTP1p) and PIG9p to have a NLS, and PIG5p, PIG14p, PIG15p and PIG23p to not have such an NLS. The non-existence of an NLS for PIG5p, PIG14p and PIG15p could be experimentally verified. In contrast, the nuclear localization of PIG7p (RTP1p) could be experimentally verified, albeit not in the nucleus of the fungus, but rather in the nucleus of the host plant. Although it could not be proven during this thesis that the discovered NLS is responsible for the alien nuclear localization, the evidence supports this assumption (see Table 4-6). The nuclear- resp. non-nuclear localization of PIG9p and PIG23p could not be experimentally verified, due to the (correctly predicted) instability of these proteins. Thus, the NUCDISC algorithm proved to be reliable for those four PIGps, which could also be experimentally studied.

In contrast, the only NNCN result, which could be experimentally shown to be correct pertained to PIG7p (RTP1p), which was correctly predicted to be a nuclear protein. NCNN also predicted PIG9p and PIG23p to be nuclear proteins (although NUCDISC did not predict PIG23p to have a NLS), but this could not be tested during this work due to the instability of these proteins. The NNCN algorithm predicted PIG5p to be nuclear, whereas it was experimentally shown during this treatise to be located inside the haustorium, but not in the nucleus. Further studies by E. Kemen revealed PIG5p to be most likely located in the endoplasmic reticulum (ER) (personal communication E. Kemen). The NNCN algorithm also predicted PIG14p and PIG15p to be cytoplasmic proteins, whereas according to the experimental results obtained during this work, they are actually located in the extrahaustorial membrane or secreted. Since three out of four NCNN predictions were shown to be incorrect, the predictions of the NNCN algorithm proved to be not reliable for use with *U. fabae* within the framework of this thesis.

The k-NN algorithm provided somewhat more ambiguous results. For example, PIG5p was predicted to be an extracellular- or cell wall protein (44.4% probability) or a cytoplasmic protein (33.3% probability). According to the experimental results, PIG5p is most possibly located in the ER (personal communications E. Kemen). If PIG5p is an extracellular- or cell wall protein as predicted by the k-NN algorithm, the localization of PIG5p in the ER is most likely an intermediate step, because (pre-)proteins destined for secretion, the cell organelles, the plasma membrane, or the cell membrane, are first transported to the ER (Kelly 1985; Alberts 1994). Thus, despite not having given the correct location the highest probability, the k-NN prediction provided a valuable clue for the localization of this protein.

Using the same algorithm, PIG7p (RTP1p) was predicted to be a mitochondrial protein (34.8% probability), a nuclear protein (30.4% probability), or an extracellular-, or cell wall protein (26.1% probability). According to the experimental results achieved during this work, PIG7p (RTP1p) is an extracellular protein (secreted from the fungus) as well as a nuclear protein (located in the host nucleus). Thus, two of the three most probable k-NN predictions were correct. It can even be hypothesized that the incorrect mitochondrial prediction resulted from confusion caused by the fact that the algorithm did not foresee PIG7p having two different locations.

The k-NN algorithm predicts PIG14p to be an extracellular or cell wall protein with a probability of 77.8%. This is in concordance with the experimental results. PIG15p is predicted to be an extracellular- or cell wall protein (44.4% probability), or a plasma membrane protein (22.2% probability). Although the secretion of PIG15p could not be experimentally shown during this thesis, the localization in the extrahaustorial matrix could be clearly seen in the immuno-fluorescence microscopy pictures (see Figure 4-13). Therefore the results of the k-NN prediction are also concordant with the experimental results for PIG15p. Based on the above results, we see that the k-NN algorithm was helpful in predicting the localization of all PIG-proteins studied during this thesis, making it a suitable algorithm for use in further molecular studies of *U. fabae*.

In summary it can be said that the NCNN algorithm should be used with caution in connection with *U. fabae*, whereas the k-NN algorithm as well as the NUDISC algorithm provide more or less accurate and valuable information regarding the localization of the studied proteins.

5.2.4 Molecular Mass of PIG-Proteins

A further characteristic of the PIGps analyzed in conjunction with their secretion behavior was their molecular mass. When comparing the theoretically calculated molecular weights of the PIG7p (RTP1p), PIG14p and PIG15p pre-proteins (24.2 kDa, 25.6 kDa, 32.1 kDa respectively) and of their mature proteins (22.1 kDa, 23.7 kDa, 29.6 kDa respectively) with the experimental results (23.5 kDa/25.0 kDa, 29 kDa, 31 kDa respectively), we see that the experimentally derived results from haustoria are 6.3%/13.1%, 22.4% and 4.7% larger respectively. No mass comparison could be carried out for PIG5p, PIG9p and PIG23p, as Western blots were not possible with these PIGps. As can be seen in table 4.8, the experimentally derived weights of the PIG-proteins in haustoria were on average 11% larger than the calculated mass. This phenomenon is most likely due to a weak glycolysation of the proteins.

5.2.5 Homology of the Studied *PIG*s to other Genes

One method to more reliably characterize the studied *PIG*s is to compare them for homology to *PIG*s or genes in general from other organisms. Developmentally-regulated genes in plant fungal interactions other than *V. faba* / *U. fabae* have been described to a great extent in literature (Bhairi et al. 1989; Liu et al. 1993; de Wit et al. 1994; Wubben et al. 1994; Heath 1995; Kahmann et al. 1999; Basse et al. 2000; Kahmann and Basse 2001; Breuninger and Requena 2004; Schulze Gronover et al. 2004). One example is the arbuscular mycorrhizal symbiosis. This symbiosis in the system *Glomus mosseae* / *Petroselinum crispum* is induced upon a series of recognition events involving the reorganization of both plant and fungal cellular programs, culminating in the formation of appressoria on the epidermal root cells (Breuninger and Requena 2004). It is also known that in the ectomycorrhizal basdiomycete *Pisolithus tinctorius*, genes of the hydrophobin family, a group of small cysteine-rich fungal proteins involved in morphogenesis, are expressed during plant-fungus interactions (Tagu et al. 1996). Other examples for development-regulated genes in plant fungal interactions are the *usp*-genes expressed during germination of *Puccinia graminis* (Liu et al. 1993), and the *INF*-genes expressed during infection structure development of *Uromyces appendiculatus* (Bhairi et al. 1989; Xuei et al. 1992). In addition, several genes of *Botrytis cinerea* up-regulated during infection of host plants have been identified (Schulze Gronover et al. 2004). *In planta* induced genes have also been described for *Phytophthora infestans* (Pieterse et al. 1991; Pieterse et al. 1993) and *Cladosporium fulvum* (Wubben et al. 1994).

Despite the above research, and although the complete genomes of many fungi have already been made public by the Whitehead Institute (http://www-genome.wi.mit.edu/ annotation/fungi/fgi/ status.html), none of the *in planta* induced genes studied in this work displayed a convincing similarity to genes published in the public databases. Results of BLASTX searches at NCBI only produced an extremely weak similarity to a heavy neurofilament subunit of *Rattus norvegicus* (2e –07) for *PIG9* and to a putative trehalose-phosphate synthase of *Oryza sativa* (5e –08) for *PIG15*. Both scores were just below the cut-off value of 1e –05.

These findings, respectively the lacks thereof, do not rule out the possibility that homologues to these *in planta* induced genes may exist in other fungi not fully published in the public databases. For example, M. Ernst showed that Southern blotting with DIG-labeled probes of *PIG5* generates a visible hybridization signal with genomic DNA of *Ustilago maydis*, *Coprinus cinereus*, *Aspergillus nidulans* and *Saccharomyces cerevisiae* (Ernst 1997). Unfortunately, no convincing similarities to the fully sequenced and published genomes of the abovementioned organisms were found during this thesis. Thus, the existence of homologues to *PIG5* in distantly related fungi must be questioned.

For *PIG7* (*RTP1*), it has been shown that DIG-labeled DNA probes generate a medium-strength hybridization signal with the genomic DNA of *Uromyces vignae, Puccinia graminis* and *Aspergillus nidulans* (Ernst 1997). But none of these signals was as strong as the hybridization signal with *U. fabae* DNA, suggesting only a weak homology (Ernst 1997). In addition, it is known that a homologue for *PIG7* (*RTP1*) exists in *Uromyces striatus* (personal communication by M. Hahn).

The difference in Southern blotting results achieved during this thesis and work done by M. Ernst (Ernst 1997) may result from of the use of two different probes against *PIG7*. The probe used during this work was generated from *PIG7* cDNA, cloned into the vector pET28a to exclude cDNA from genes other than *PIG7*. The thus-generated probe led to strong hybridization signals with the genomic DNA of *Uromyces fabae*, *Uromyces striatus* and *Uromyces appendiculatus*. This result was to be expected as these *Uromyces* species are closely related. Depending on the restriction enzymes used, weak and diffuse hybridization signals could also be seen in conjunction with the DNA from *Aspergillus nidulans*, also suggesting only a weak similarity at best.

Based on the experimental results discussed above, it could be concluded that (remotely related) homologues to *PIG7* exist in fungal organisms other than the members of the *Uromyces* family. But as also mentioned, no homologues to *PIG7* could be found in the public databases searched during this work. Therefore, it seems unlikely that these hybridization signals results indicate the existence of real homologues to *PIG7* in other fungi outside the *Uromyces* family.

M. Ernst also showed that hybridization with DIG-labeled probes of *PIG9* generates a strong hybridization signal with *Phytophthora infestans* DNA and medium-strength hybridization signals with the genomic DNA of *Puccinia graminis, Coprinus cinereus*, *Magnaporthe grisea* and *Vicia faba* (Ernst 1997). But as mentioned earlier, only weak homologues to *PIG9* could be found in the public databases searched during this thesis.

M. Ernst demonstrated that in contrast to the hybridization signals with probes against *PIG7* and *PIG9*, DIG-labeled probes of *PIG14* provide neither strong nor medium-strength hybridization signals with the tested DNA probes other than *U. fabae* DNA (Ernst 1997).

One possible explanation for the contradiction between the positive homology findings in the Southern blotting done during this thesis and in the work by M. Ernst, and the negative findings resulting from comparisons to genes published in the public databases, is the undefined variance in the genomic sizes of the tested fungi. In all Southern blots, the same amount of DNA was digested for each fungus, although the genomic size of all fungi was not known. Therefore, the comparative quantitative analysis of the resulting signals is not reliable, as a smaller genome leads to an over-evaluation of individual genes. Under these circumstances, visible hybridization signals can be misinterpreted as an indication that there may be homologues to the applied probe.

5.2.6 Gene Frequency of the Studied *PIG*s

To further characterize the studied *U. fabae PIG*s, each was analyzed in respect to the number of gene copies present in the fungus using Southern blotting. This analysis is of interest because of the general notion that multiple copies of the same gene in an organism may indicate a heightened importance of the gene in question. These gene families may arise by gene duplication. Gene duplication can result from unequal crossing over, retroposition or chromosomal (or genome) duplication (Zhang 2003).

Summarizing the results, it was found that two copies of *PIG5* and *PIG7* (with *PIG*7 having most likely an "single nucleotide polymorphism" (SNP)), but only one copy of *PIG14, PIG15* and *PIG23* exist. As discussed below, no definite conclusion can be drawn for *PIG9* based on the results of the Southern blots. In this case, either one copy or two copies may exist.

Despite the fact that rust fungi have predominately dikaryotic stages, thus having two nuclei per cell (Agrios 1997), we cannot generally expect to find more than one copy of each *PIG* in *U. fabae*. This is because although the DNA isolated from uredospore germ tubes during this thesis originated from both nuclei, it can be assumed that the entirety of the genes from both nuclei are more or less identical. In consequence, the *PIG*s for which one copy was found are not conspicuous. Those *PIG*s where more than one copy may have been uncovered (*PIG5, PIG7, PIG9* and *PIG23)* warrant further analysis.

When reviewing the EST project (results not shown) in conjunction with the Southern Blotting results, we see that while the cDNA of *PIG5* does not seem to vary, the cDNA of both copies of *PIG7* differs in one amino acid at the restriction site for *Mun*I (*Mfe*I). For *PIG5*, we can thus conclude that two copies of this gene exist per nucleus. For *PIG7*, the deviation in one amino acid has three possible differing explanations:

- During experimental cDNA synthesis, a translation error occurred, or
- There are at least two copies of *PIG7* per nucleus, or
- One of the two nuclei of *U. fabae* may contain a different (point-mutated) copy of *PIG7*.

To shed further light on this mystery, *Mun*I was used as a fifth restriction enzyme for the Southern blotting of *PIG7*. The third hybridization signal found in the *Mun*I lane supports the hypothesis that another copy of *PIG7* may exist. Solving the riddle of the origin of these two *PIG7* (*RTP1*) copies would require extensive supplementary research, including cross-breeding of the monokaryotic basidiospores and subsequent isolation of DNA originating from only one nucleus (karyon).

Based on the Southern blotting and EST project results for *PIG9,* it can be assumed that two copies of this gene most likely exist. Nonetheless, the possibility that all employed restriction enzymes cut at a restriction site in the introns of one and the same gene copy cannot be excluded. Therefore, based on the information available, the number of copies must be stated as one or two.

For *PIG14* and *PIG15* the Southern blotting and EST project results are clear: each gene has only one copy.

While the EST project results show no variation in the cDNA, the Southern Blotting results for *PIG23* are more vague. But since we know that a restriction site for *Bam*HI exists in the cDNA and we only find two distinctive hybridization signals in the lane were *Bam*HI digested DNA was applied, it is most likely that only one copy of *PIG23* exists in *U. fabae*.

The above described Southern blot results correspond to the findings of Hahn and Mendgen (1997) that *PIG*s are single or low copy rust genes. Based on the hypothesis that the number of gene copies is linked to its importance, and at a maximum two copies of a studied *PIG*s has been found, it cannot be stated that any one of the *PIG*s is more critical to the *U. fabae* life cycle than the others.

5.2.7 Development- and Time Dependent Regulation of *PIG* Expression

Developmentally-dependent expressions of *U. fabae* genes have been previously described (Deising et al. 1991; Hahn and Mendgen 1992; Heiler et al. 1993; Deising et al. 1995; Rauscher et al. 1995; Struck et al. 1996; Hahn and Mendgen 1997; Hahn et al. 1997; Rauscher et al. 1999; Sohn et al. 2000; Voegele et al. 2001; Wirsel et al. 2001; Mendgen and Hahn 2002; Link et al. 2005). During the early stages of germination, cutinases and esterases are secreted to assist the attachment of uredospores to the host cuticle (Deising et al. 1992). When appressoria are formed, extracellular proteases (Rauscher et al. 1995) as well as cellulases appear (Heiler et al. 1993). When the rust fungus forms haustorial mother cells, it also secretes lyases (Deising et al. 1995).

Another class of genes is expressed when the fungus reaches the biotrophic stage. Prominent examples are some amino acid transporters (Struck et al. 2002; Struck et al. 2004), a hexose transporter (Voegele et al. 2001), genes coding for thiamine synthesis (Sohn et al. 2000), and the *in planta* induced genes, with the majority of these being of unknown function (Hahn and Mendgen 1997).

To analyze the developmental stage- and thus time-dependent expression of the studied *PIG*s, RNA- and total-protein extracts from the different infection structures (*in vitro*) and haustoria (*in vivo*) were isolated and analyzed using Northern-, and Western blotting respectively.

As can been seen from the results of the Northern Blots (see chapter 4.8), *PIG5, PIG7, PIG14*, and *PIG15* exhibit a time-dependent expression behavior. All these genes are only expressed *in planta*, (when the biotrophic stage is reached), as the name "*in planta* induced gene" indicates. A DIG-labeled probe of *PIG23* cDNA was able to hybridize with RNA isolated from different *in vitro* stages of *U. fabae*. This fact leads to the conclusion that *PIG23* is not truly *in planta* induced, and was therefore not studied any further.

With one exception, the results of the immuno-blot analysis (see chapter 4.9.1) confirm the results of the Northern blotting. Whereas in the Northern Blots, the first RNA expression of the *PIG*s is visible in haustoria, the immuno-blot of PIG15p shows antibody cross-reaction signals with protein isolated from haustoria as well as non-germinated spores.

One explanation for this unexpected phenomenon could be that the additional cross-reaction signal observed in the immuno-blot analysis occurs due to unspecific reaction of the antibody with parts of the non-germinated spores. This assumption is rebutted by the observation that the position of the antibody cross-reaction signal in protein extract from non-germinated spores is the same compared to the antibody cross-reaction signal in protein extract from the haustoria (see Table 4-9).

Another explanation could be that the spore-building tissue deposits the PIG15p in the spores, similar to the findings for THI1p (Sohn 2000). J. Sohn (2000) has shown that the presence of THI1p can be verified in non-germinated spores. The amount of THI1p decreases during further *in vitro* development stages and increases in the haustoria, whereas the RNA of *THI1* is first detectable in *in vitro* structures after 18 h of germination. The spores of *U. fabae* contain THI1p but no mRNA to generate new THI1p (Sohn 2000). Despite this similarity, without further knowledge of the functions of PIG15p, no conclusions can be drawn regarding the possible PIG-protein storage in spores of *U. fabae*.

Whereas the function of the genes studied during this thesis could not be determined, their stage-dependent development, their export signals and the localization of the gene products suggest an important role for these PIGps during the biotrophic stage. The findings described in this chapter indicate the existence of previously unknown interactions between the rust fungus and its host during the establishment of the biotrophic phase. Proteins, which are *in planta* induced and additionally secreted into the plant, or at least found at the fungal/plant borderline, are likely to be involved in the establishment of the biotrophic phase.

5.2.8 Import of Fungal Proteins into the Plant Nucleus

The import of alien proteins into mammalian nuclei has been described for the SV40 T antigen (Kasamatsu and Nakanishi 1998; Fouchier and Malim 1999). In addition, bacterial protein transfer into plant nuclei by *Agrobacterium tumefaciens* has been frequently described (Mysore et al. 1998; Ziemienowicz et al. 1999; Escobar and Dandekar 2003; Gelvin 2003; Veena et al. 2003) *A. tumefaciens* causes crown gall disease in a variety of plants. During the infection process, *Agrobacterium* transfers a nucleoprotein complex, the so-called VirD2 T-complex, and at least two Vir proteins, into the plant cell. These proteins are then transported into the cell nucleus (Herrera-Estrella et al. 1990; Citovsky et al. 1992). VirD2 is imported by a mechanism conserved between animal, yeast and plant cells (Herrera-Estrella et al. 1990; Citovsky et al. 1992; Tzfira et al. 2001).

In contrast, knowledge about fungal proteins imported into plant nuclei is marginal. Fungal protein import into plant cells has been described for the pine blister rust fungus *Cronartium ribicola*, but no studies regarding localization in the nucleus are mentioned (Ekramoddoullah et al. 1999). The flax rust *Melampsora lini* expresses avirulence genes in haustoria and their products are recognized inside plant cells, but as above, no localization in the plant nucleus is reported (Dodds et al. 2004).

Nuclear localization signals (NLS) have been described for different species (Macara 2001; Merkle 2001). Two of the PIGps (PIG7p and PIG9p) with an export signal sequence (see chapter 4.2.2) studied during this thesis also exhibit a NLS. In addition, immuno-fluorescence staining of PIG7p (RTP1p) using polyclonal antibodies has shown localization in the plant nucleus (see Figure 4-13). Because PIG9p proved to be unstable under laboratory conditions, further analysis via immuno-fluorescence staining was not possible.

This discovery allows speculation on how the import of fungal proteins into the host plant nucleus can be accomplished. Since the research done during this thesis has shown that PIG7p (RTP1p) is exported from the haustorium, it must pass the haustorial wall, the extrahaustorial matrix and the extrahaustorial membrane before entering the host cell. Once inside the plant cell, PIG7p has to navigate its way to the plant nucleus. In general, the process of protein trafficking is influenced by the ability of the protein to interact with the membrane systems of the cell. By virtue either of its sequence or covalent groups added to it, a protein is recognized by receptors in a particular membrane (Alberts 1994). Apparently, the plant cell recognizes the NLS of PIG7p (RTP1p) and is able to execute the instructions contained within the NLS.

The research described above reveals PIG7p (RTP1p), to which no homology was found in public databases, to be located in the extrahaustorial matrix as well as in the nucleus of infected *V. faba* cells. Therefore PIG7p (RTP1p) is the first known protein that is transferred from the rust haustoria into the nucleus of a host plant, suggesting that it plays an important role in the biotrophic interaction between *U. fabae* and *V. faba*.

6 List of Abbreviations

°C	Degree Celsius
μm	Micrometer
A. fabae	*Ascochyta fabae*
Amp^r	Ampicilline resistance
B. fabae	*Botrytis fabae*
BLAST®	Basic Local Alignment Search Tool
bp	Base pairs
BSA	Bovine Serum Albumin
cDNA	Complementary DNA (see below)
CH	Switzerland
ConA	Concanavalin A
CSPD	Disodium 3-(4-metho xyspiro {1,2-dioxetane-3,2-(5-chloro)tricyclo [3.3.1.13,7]decan}-4-yl)phenyl phosphate
cv.	cultivar
D	Germany
DNA	Deoxyribonucleic Acid
dsRNA	Double stranded RNA (see below)
EDTA	EthyleneDiamine Tetraacetic Acid
EST	Expressed Sequence Tag
E-value	Expect value
g	Gravitational constant
g	Gram
h	*Hora* (hour)
HSP	Heat Shock Protein
IPTG	Isopropyl β-D-thiogalactopyranoside
Kan^r	Kanamycin resistance
M	Mol
Max.	Maximum
min	Minute
ml	Milliliter
mM	Millimol
mRNA	messenger RNA (Ribonucleic acid)
NCBI	National Center for Biotechnology Information
NL	The Netherlands
NLS	Nuclear Localization Signal
NPC	Nuclear Pore Channel
ON	Over Night
p.a.	Per analysis
PCR	Polymerase Chain Reaction
PEG	Polyethylenglycol
pH	Negative decade logarithm of the H^+ ion concentration
PIG	*In Planta* Induced Gene
PIGp	Gene product (protein) of the *PIG*
REMI	Restriction Enzyme Mediated Integration
R-gene	Resistance gene
RNA	RiboNucleic Acid

rpm	Revolutions per minute
RT	Room temperature
sec.	Second
SNP	Single Nucleotide Polymorphism
T	Temperature
Tris	2-amino-2-hydroxymethyl-1,3-propanediol
U. fabae	*Uromyces fabae*
URL	Uniform Resource Locator
V. faba	*Vicia faba* (Broad Bean)
VLP	Virus Like Particle

7 Literature

Adam, E. J. and S. A. Adam (1994). "Identification of cytosolic factors required for nuclear location sequence-mediated binding to the nuclear envelope." J Cell Biol **125**(3): 547-555.

Adam, S. A., R. S. Marr and L. Gerace (1990). "Nuclear protein import in permeabilized mammalian cells requires soluble cytoplasmic factors." J Cell Biol **111**(3): 807-816.

Agriculture & Resource Management Council of Australia & New Zealand (2001). "Australian & New Zealand Environment & Conservation Council and Forestry Ministers. Weed of National Signidicance Blackberry (*Rubus fruticosus* L. agg) Strategic Plan." National Weeds Strategy Executive Committee, Launceston: 1-26.

Agrios, G. N. (1997). Plant pathology. San Diego, Academic Press.

Alberts, B. (1994). Molecular biology of the cell. New York, Garland Pub.

Alexopoulos, C. J., C. A. Mims and M. Blackwell (1996). Introductory Mycology. New York, John Wiley & Sons.

Altschul, S. F. and B. W. Erickson (1985). "Significance of nucleotide sequence alignments: a method for random sequence permutation that preserves dinucleotide and codon usage." Mol Biol Evol **2**: 526-538.

Altschul, S. F. and W. Gish (1996). "Local alignment statistics." Methods Enzymol **266**: 460-480.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-410.

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucl Acids Res **25**(17): 3389-3402.

Bachmair, A. and A. Varshavsky (1989). "The degradation signal in a short-lived protein." Cell **56**(6): 1019-1032.

Balhadère, P. V., A. J. Foster and N. J. Talbot (1999). "Identification of Pathogenicity Mutants of the Rice Blast Fungus *Magnaporthe grisea* by Insertional Mutagenesis." Mol Plant-Microbe Interactions **12**(2): 129-142.

Ballas, N. and V. Citovsky (1997). "Nuclear localization signal binding protein from *Arabidopsis* mediates nuclear import of *Agrobacterium* VirD2 protein." PNAS **94**(20): 10723-10728.

Basse, C. W., S. Stumpferl and R. Kahmann (2000). "Characterization of a *Ustilago maydis* gene specifically induced during the biotrophic phase: evidence for negative as well as positive regulation." Mol Cell Biol **20**(1): 329-339.

Beadle, B. M. and B. K. Shoichet (2002). "Structural bases of stability-function tradeoffs in enzymes." J Mol Biol **321**(2): 285-296.

Bechinger, C., K. F. Giebel, M. Schnell, P. Leiderer, H. B. Deising and M. Bastmeyer (1999). "Optical measurements of invasive forces exerted by appressoria of a plant pathogenic fungus." Science **285**(5435): 1896-1899.

Bender, D. A. (2003). Vitamin B1: Thiamin. Nutritional Biochemistry of the Vitamins. New York, Cambridge University Press**:** 148-169.

Bhairi, S. M., R. C. Staples, P. Freve and O. C. Yoder (1989). "Characterization of an infection structure-specific gene from the rust fungus *Uromyces appendiculatus*." Gene **81**(2): 237-243.

Birnboim, H. C. and J. Doly (1979). "A rapid alkaline extraction procedure for screening recombinant plasmid DNA." Nucl Acids Res **7**: 1513-1523.

Blattner, F. R., B. G. Williams, A. E. Blechl, K. Denniston-Thompson, H. E. Faber, L. Furlong, D. J. Grunwald, D. O. Kiefer, D. D. Moore, J. W. Schumm, E. L. Sheldon and O. Smithies (1977). "Charon phages: safer derivatives of bacteriophage lambda for DNA cloning." Science **196**(4286): 161-169.
Bogerd, H. P., R. E. Benson, R. Truant, A. Herold, M. Phingbodhipakkiya and B. R. Cullen (1999). "Definition of a consensus transportin-specific nucleocytoplasmic transport signal." J Biol Chem **274**(14): 9771-9777.
Bonas, U. and G. Van den Ackerveken (1999). "Gene-for-gene interactions: bacterial avirulence proteins specify plant disease resistance." Curr Opin Microbiol **2**(1): 94-98.
Bonner, W. M. (1978). Proteinmigration and accumulation in nuclei. The Cell Nucleus. H. Busch. New York, Academic. **6, Part C:** 97-148.
Boos, W. (1976). "Cell envelope proteins involved in the transport of maltose and sn-glycerol-3-phosphate in *Escherichia coli*." J Cell Physiol **89**(4): 529-541.
Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." Anal Biochem **72**: 248-254.
Brendel, V., P. Bucher, I. R. Nourbakhsh, B. E. Blaisdell and S. Karlin (1992). "Methods and algorithms for statistical analysis of protein sequences." PNAS **89**(6): 2002-2006.
Breuninger, M. and N. Requena (2004). "Recognition events in AM symbiosis: analysis of fungal gene expression at the early appressorium stage." Fungal Genet Biol **41**(8): 794-804.
Brouta, F., F. Descamps, M. Monod, S. Vermout, B. Losson and B. Mignon (2002). "Secreted metalloprotease gene family of *Microsporum canis*." Infect Immun **70**(10): 5676-5683.
Bruns, T. D., R. Vilgalys, S. M. Barns, D. Gonzalez, D. S. Hibbett, D. J. Lane, L. Simon, S. Stickel, T. M. Szaro, W. G. Weisburg and M. L. Sogin (1993). "Evolutionary relationships within the fungi: analysis of nuclear small subunit rRNA sequences." Molecular Phylogenetics and Evolution **1**: 231-241.
Bruns, T. D., T. J. White and J. W. Taylor (1991). "Fungal molecular systematics." Annu Rev Ecol Syst **22**: 525-564.
Bruzzese, E., F. Mahr and I. Faithfull (2000). "Best practice management guide 5." Best practice management guide for invironmental weeds CRC for Weed Management Systems(5): 1-6.
Bucher, P. and A. Bairoch (1994). "A generalized profile syntax for biomolecular sequence motifs and its function in automatic sequence interpretation." Proc Int Conf Intell Syst Mol Biol **2**: 53-61.
Burgess, T. L. and R. B. Kelly (1987). "Constitutive and regulated secretion of proteins." Annu Rev Cell Biol **3**: 243-293.
Burnette, W. H. (1981). "Western blotting: Electrophoretic transfer of proteins from SDS-polyacrylamide gels to unmodofied nitrocellulose and radiographic detection with antibody and radioiodinated proteinA." Anal Biochem **112**: 195-203.
Büttner, D. and U. Bonas (2002). "Port of entry--the type III secretion translocon." Trends Microbiol **10**(4): 186-192.
Büttner, D. and U. Bonas (2003). "Common infection strategies of plant and animal pathogenic bacteria." Curr Opin Plant Biol **6**(4): 312-319.
Carlile, M. J., S. C. Watkinson and G. W. Gooday (2001). The Fungi. San Diego, Academic Press.

Carninci, P., Y. Shibata, N. Hayatsu, Y. Sugahara, K. Shibata, M. Itoh, H. Konno, Y. Okazaki, M. Muramatsu and Y. Hayashizaki (2000). "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes." Genome Res **10**(10): 1617-1630.

Cato, M. C. (ca. 150 b. C.). De Agricultura.

Chi, N. C., E. J. Adam and S. A. Adam (1995). "Sequence and characterization of cytoplasmic nuclear protein import factor p97." J Cell Biol **130**(2): 265-274.

Chomczynski, P. (1992). "One-hour downward alkaline capillary transfer for blotting of DNA and RNA." Anal Biochem **201**: 134-139.

Citovsky, V., D. Warnick and P. Zambryski (1994). "Nuclear import of *Agrobacterium* VirD2 and VirE2 proteins in maize and tobacco." PNAS **91**(8): 3210-3214.

Citovsky, V., J. Zupan, D. Warnick and P. Zambryski (1992). "Nuclear localization of *Agrobacterium* VirE2 protein in plant cells." Science **256**(5065): 1802-1805.

Clark, M. D., S. Hennig, R. Herwig, S. W. Clifton, M. A. Marra, H. Lehrach, S. L. Johnson and W.-G. E. Group (2001). "An oligonucleotide fingerprint normalized and expressed sequence tag characterized zebrafish cDNA library." Genome Res **11**(9): 1594-1602.

Clement, J. A., R. Porter, T. M. Butt and A. Beckett (1997). "Characteristics of adhesion pads formed during imbibition and germination of urediniospores of *Uromyces viciae-fabae*." Mycol Res **101**: 1445-1458.

Clever, J., D. A. Dean and H. Kasamatsu (1993). "Identification of a DNA binding domain in simian virus 40 capsid proteins Vp2 and Vp3." J Biol Chem **268**(28): 20877-20883.

Cokol, M., R. Nair and B. Rost (2000). "Finding nuclear localization signals." EMBO Rep **1**(5): 411-415.

Davis, L. I. (1995). "The nuclear pore complex." Annu Rev Biochem **64**: 865-896.

Dayan, F. E., A. M. Rimando, M. R. Tellez, B. E. Scheffler, T. Roy, H. K. Abbas and S. O. Duke (2002). "Bioactivation of the fungal phytotoxin 2,5-anhydro-D-glucitol by glycolytic enzymes is an essential component of its mechanism of action." Z Naturforsch [C] **57**(7-8): 645-653.

de Wit, P. J., M. H. Joosten, G. Honee, J. P. Wubben, G. F. van den Ackerveken and H. W. van den Broek (1994). "Molecular communication between host plant and the fungal tomato pathogen *Cladosporium fulvum*." Antonie Van Leeuwenhoek **65**(3): 257-262.

Dean, D. A. (1997). "Import of plasmid DNA into the nucleus is sequence specific." Exp Cell Res **230**(2): 293-302.

Deising, H., A. K. Frittrang, S. Kunz and K. Mendgen (1991). "Differentiation-related proteins of the broad bean rust fungus *Uromyces viciae-fabae*, as revealed by high solution two-dimensional polyacrylamid gel electrophoresis." Arch Microbiol **155**: 194-198.

Deising, H., A. K. Frittrang, S. Kunz and K. Mendgen (1995). "Regulation of pectin methylesterase and polygalacturonate lyase activity during differntiation of infection structures in *Uromyces viciae-fabae*." Microbiology **141**: 561-571.

Deising, H., R. L. Nicholson, M. Haug, R. J. Howard and K. Mendgen (1992). "Adhesion pad formation and the Involvement of cutinase and esterases in the attachment of uredospores to the host cuticle." Plant Cell **4**(9): 1101-1111.

Deising, H. and J. Siegrist (1995). "Chitin deacetylase activity of the rust fungus *Uromyces viciae-fabae* is controlled by fungal morphogenesis." FEMS Microbiology Letters **127**: 207-212.

Deising, H. B., S. Werner and M. Wernitz (2000). "The role of fungal appressoria in plant infection." Microbes Infect **2**(13): 1631-1641.

DeRisi, J. L., V. R. Iyer and P. O. Brown (1997). "Exploring the metabolic and genetic control of gene expression on a genomic scale." Science **278**(5338): 680-686.
Deslandes, L., J. Olivier, N. Peeters, D. X. Feng, M. Khounlotham, C. Boucher, I. Somssich, S. Genin and Y. Marco (2003). "Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus." PNAS **100**(13): 8024-8029.
Di Pietro, A., M. D. Huertas-Gonzalez, J. F. Gutierrez-Corona, G. Martinez-Cadena, E. Meglecz and M. I. Roncero (2001). "Molecular characterization of a subtilase from the vascular wilt fungus *Fusarium oxysporum*." MPMI **14**(5): 653-662.
Dickinson, M. J. and A. Pryor (1988a). "Encapsidated and unencapsidated double-stranded RNAs in flax rust, *Melampsora lini*." Can J Bot **67**: 1137-1142.
Dickinson, M. J. and A. Pryor (1988b). "Isometric virus-like particles encapsidate the double-stranded RNA found in *Puccinia striformis, Puccinia recondita,* and *Puccinia sorghi*." Can J Bot **67**: 3420-3425.
Dickinson, M. J., C. R. Wellings and A. Pryor (1989). "Variation in the double-stranded RNA phenotype between and within different rust species." Can J Bot **68**: 559-604.
Dickman, M. B., G. K. Podila and P. E. Kolattukudy (1989). "Insertion of cutinase gene into a wound pathogen enables it to infect intact host." Nature **342**: 446-448.
Dodds, P. N., G. J. Lawrence, A. M. Catanzariti, M. A. Ayliffe and J. G. Ellis (2004). "The *Melampsora lini* AvrL567 avirulence genes are expressed in haustoria and their products are recognized inside plant cells." Plant Cell **16**(3): 755-768.
Edwards, M. C. and D. J. F. Bowling (1986). "The growth of rust germ tubes toward stomata in relation to pH gradients." Physiol Plant Pathol **29**: 185-196.
Ekramoddoullah, A. K. M., Y. Tan, X. Yu, D. W. Taylor and S. Misra (1999). "Identification of a protein secreted by the blister rust fungus *Cronartium ribicola* in infected white pines and its cDNA cloning and characterization." Can J Bot **77**: 800-808.
Elble, R. (1992). "A simple and efficient procedure for transformation of yeasts." Biotechniques **13**(1): 18-20.
Emr, S. D., R. Schekmann, M. C. Flessel and J. Thorner (1983). "An MF-alpha-1-SUC-2 aplpha factor invertase EC-3.2.1.26 gene fusion for study pf protein localization and gene expression in yeast *Saccharomyces cerevisiae*." PNAS **80**: 7080-7084.
Ernst, M. (1997). Pflanzeninduzierte Gene des Ackerbohnenrostes *Uromyces fabae*: Gibt es Homologe in anderen Pilzen? University of Konstanz, Konstanz, Germany
Escobar, M. A. and A. M. Dandekar (2003). "*Agrobacterium tumefaciens* as an agent of disease." Trends Plant Sci **8**(8): 380-386.
Felk, A., M. Kretschmar, A. Albrecht, M. Schaller, S. Beinhauer, T. Nichterlein, D. Sanglard, H. C. Korting, W. Schafer and B. Hube (2002). "Candida albicans hyphal formation and the expression of the Efg1-regulated proteinases Sap4 to Sap6 are required for the invasion of parenchymal organs." Infect Immun **70**(7): 3689-3700.
Felle, H. H. and S. Hanstein (2002). "The apoplastic pH of the substomatal cavity of *Vicia faba* leaves and its regulation responding to different stress factors." J Exp Bot **53**(366): 73-82.
Figueiredo, M. B. (2000). "The Plasticity of Rust Life Cycles." O Biologico **62**(1).
Fischer, W., R. Haas and S. Odenbreit (2002). "Type IV secretion systems in pathogenic bacteria." Int J Med Microbiol **292**(3-4): 159-168.
Flor, H. H. (1942). "Inheritance of Pathogenicity in *Melampsora lini*." Phytopathology(32): 653-669.
Flor, H. H. (1946). "Genetics of pathogenicity in *Melampsora lini*." Jour Agr Res(73): 335-366.

Flor, H. H. (1955). "Host-parasite interactions in flax rust - its genetic and other implications." Phytopathology(45): 680-685.
Flor, H. H. (1956). "The complementary genic system in flax and flax rust." Adv Genetics(8): 29-54.
Flor, H. H. (1971). "Current status of the gene-for-gene concept." Annu Rev Phytopathol **9**: 275-296.
Fouchier, R. A. and M. H. Malim (1999). "Nuclear import of human immunodeficiency virus type-1 preintegration complexes." Adv Virus Res **52**: 275-299.
Freytag, S., L. Bruscaglioni, R. E. Gold and K. Mendgen (1988). "Basidiospores of rust fungi (*Uromyces* species) differentiate infections structures *in vitro*." Experimental Mycology **12**: 275-283.
Fried, H. and U. Kutay (2003). "Nucleocytoplasmic transport: taking an inventory." Cell Mol Life Sci **60**(8): 1659-1688.
Galagan, J. E., S. E. Calvo, K. A. Borkovich, E. U. Selker, N. D. Read, D. Jaffe, W. FitzHugh, L. J. Ma, S. Smirnov, S. Purcell, B. Rehman, T. Elkins, R. Engels, S. Wang, C. B. Nielsen, J. Butler, M. Endrizzi, D. Qui, P. Ianakiev, D. Bell-Pedersen, M. A. Nelson, M. Werner-Washburne, C. P. Selitrennikoff, J. A. Kinsey, E. L. Braun, A. Zelter, U. Schulte, G. O. Kothe, G. Jedd, W. Mewes, C. Staben, E. Marcotte, D. Greenberg, A. Roy, K. Foley, J. Naylor, N. Stange-Thomann, R. Barrett, S. Gnerre, M. Kamal, M. Kamvysselis, E. Mauceli, C. Bielke, S. Rudd, D. Frishman, S. Krystofova, C. Rasmussen, R. L. Metzenberg, D. D. Perkins, S. Kroken, C. Cogoni, G. Macino, D. Catcheside, W. Li, R. J. Pratt, S. A. Osmani, C. P. DeSouza, L. Glass, M. J. Orbach, J. A. Berglund, R. Voelker, O. Yarden, M. Plamann, S. Seiler, J. Dunlap, A. Radford, R. Aramayo, D. O. Natvig, L. A. Alex, G. Mannhaupt, D. J. Ebbole, M. Freitag, I. Paulsen, M. S. Sachs, E. S. Lander, C. Nusbaum and B. Birren (2003). "The genome sequence of the filamentous fungus *Neurospora crassa*." Nature **422**(6934): 859-868.
Gasteiger, E., A. Gattiker, C. Hoogland, I. Ivanyi, R. D. Appel and A. Bairoch (2003). "ExPASy: The proteomics server for in-depth protein knowledge and analysis." Nucl Acids Res **31**(13): 3784-3788.
Gavel, Y. and G. von Heijne (1990). "Cleavage-site motifs in mitochondrial targeting peptides." Protein Eng **4**(1): 33-37.
Gelvin, S. B. (2003). "*Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool." Microbiol Mol Biol Rev **67**(1): 16-37.
Ghosh, P. (2004). "Process of protein transport by the type III secretion system." Microbiol Mol Biol Rev **68**(4): 771-795.
Goffeau, A., B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin and S. G. Oliver (1996). "Life with 6000 genes." Science **274**(5287): 546, 563-547.
Gomez-Gomez, E., M. C. Ruiz-Roldan, A. Di Pietro, M. I. Roncero and C. Hera (2002). "Role in pathogenesis of two endo-beta-1,4-xylanase genes from the vascular wilt fungus *Fusarium oxysporum*." Fungal Genet Biol **35**(3): 213-222.
Gomis-Ruth, F. X., M. Sola, F. de la Cruz and M. Coll (2004). "Coupling factors in macromolecular type-IV secretion machineries." Curr Pharm Des **10**(13): 1551-1565.
Gonda, D. K., A. Bachmair, I. Wunning, J. W. Tobias, W. S. Lane and A. Varshavsky (1989). "Universality and structure of the N-end rule." J Biol Chem **264**(28): 16700-16712.
Gorlach, J. M., E. Van Der Knaap and J. D. Walton (1998). "Cloning and targeted disruption of MLG1, a gene encoding two of three extracellular mixed-linked glucanases of *Cochliobolus carbonum*." Appl Environ Microbiol **64**(2): 385-391.

Görlich, D., S. Kostka, R. Kraft, C. Dingwall, R. A. Laskey, E. Hartmann and S. Prehn (1995). "Two different subunits of importin cooperate to recognize nuclear localization signals and bind them to the nuclear envelope." Curr Biol **5**(4): 383-392.
Görlich, D. and U. Kutay (1999). "Transport between the cell nucleus and the cytoplasm." Annu Rev Cell Dev Biol **15**: 607-660.
Görlich, D., S. Prehn, R. A. Laskey and E. Hartmann (1994). "Isolation of a protein that is essential for the first step of nuclear protein import." Cell **79**(5): 767-778.
Gu, F., A. Khimani, S. G. Rane, W. H. Flurkey, R. F. Bozarth and T. J. Smith (1995). "Structure and function of a virally encoded fungal toxin from *Ustilago maydis*: a fungal and mammalian $Ca2^{+}$ channel inhibitor." Structure **3**(8): 805-814.
Guruprasad, K., B. V. Reddy and M. W. Pandit (1990). "Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence." Protein Eng **4**(2): 155-161.
Györgyey, J., D. Vaubert, J. I. Jimenez-Zurdo, C. Charon, L. Troussard, A. Kondorosi and E. Kondorosi (2000). "Analysis of *Medicago truncatula* nodule expressed sequence tags." MPMI **13**(1): 62-71.
Hadwiger, L. A. and D. E. Culley (1993). "Nonhost resistance genes and race-specific resistance." Trends Microbiol **1**(4): 136-141.
Haerter, A. C. and R. T. Voegele (2004). "A novel beta-glucosidase in *Uromyces fabae*: feast or fight?" Curr Genet **45**(2): 96-103.
Hahn, M. and K. Mendgen (1992). "Isolation by ConA binding of haustoria from different rust fungi and comparison of their surface qualities." Protoplasma **170**: 95-103.
Hahn, M. and K. Mendgen (1997). "Characterization of in planta-induced rust genes isolated from a haustorium-specific cDNA library." MPMI **10**(4): 427-437.
Hahn, M., U. Neef, C. Struck, M. Gottfert and K. Mendgen (1997). "A putative amino acid transporter is specifically expressed in haustoria of the rust fungus *Uromyces fabae*." MPMI **10**(4): 438-445.
Haralampidis, K., M. Trojanowska and A. E. Osbourn (2002). "Biosynthesis of triterpenoid saponins in plants." Adv Biochem Eng Biotechnol **75**: 31-49.
Hartmann, E., T. A. Rapoport and H. F. Lodish (1989). "Predicting the orientation of eukaryotic membrane-spanning proteins." PNAS **86**(15): 5786-5790.
Heath, M. C. (1989). "A comparison of fungal growth and plant responses in cowpea and bean cultivars inoculated with urediospores or basidiospores of the cowpea rust fungus." Mol Plant Pathol **34**: 415-426.
Heath, M. C. (1991). "Evolution of resistance to fungal parasitism in natural ecosystems." New Phytol **119**: 331-343.
Heath, M. C. (1995). "Signal exchange between higher plants and rust fungi." Can J Bot **73**: 616-623.
Heath, M. C. (1997). "Signalling between pathogenic rustfungi and resistant or susceptible host plants." Annals of Botany **80**: 713-720.
Heath, M. C. (2000). "Nonhost resistance and nonspecific plant defenses." Curr Opin Plant Biol **3**(4): 315-319.
Heath, M. C. and D. Skalamera (1997). Cellular interactions between plants and biotrophic fungal parasites. Advances in botanical research. T. I. C. and A. H. J. eds. San Diego, London, Academic Press. **24:** 195-225.
Heiler, S., K. W. Mendgen and H. Deising (1993). "Cellulotic enzymes of the obligate biotrophic rust fungus *Uromyces viciae-fabae* are reugulated differntiation-specifically." Mycol Res **97**: 77-85.

Herrera-Estrella, A., M. Van Montagu and K. Wang (1990). "A bacterial peptide acting as a plant nuclear targeting signal: the amino-terminal portion of *Agrobacterium* VirD2 protein directs a beta-galactosidase fusion protein into tobacco nuclei." PNAS **87**(24): 9534-9537.
Hicks, G. R. and N. V. Raikhel (1995a). "Nuclear localization signal binding proteins in higher plant nuclei." PNAS **92**(3): 734-738.
Hicks, G. R. and N. V. Raikhel (1995b). "Protein import into the nucleus: an integrated view." Annu Rev Cell Dev Biol **11**: 155-188.
Hoch, H. C. and R. C. Staples (1987). "Structural and chemical changes among the rust fungi during appressoria developement." Annu Rev Phytopath **25**: 231-247.
Hoch, H. C., R. C. Staples, B. Whitehead, J. Comeau and E. D. Wolf (1987). "Signaling for growth orientation and cell differentiation by surface topography in *Uromyces*." Science **235**: 1659-1662.
Hodel, M. R., A. H. Corbett and A. E. Hodel (2001). "Dissection of a nuclear localization signal." J Biol Chem **276**(2): 1317-1325.
Holt, B. F., 3rd, D. Mackey and J. L. Dangl (2000). "Recognition of pathogens by plants." Curr Biol **10**(1): R5-7.
Hooley, R. (1994). "Gibberellins: perception, transduction and responses." Plant Mol Biol **26**(5): 1529-1555.
Hori, S. (1898). "Some observations on "Bakanae" disease of the rice plant." Mem Agric Res Sta (Tokyo) **12**(1): 110-119.
Horton, P. and K. Nakai (1997). "Better prediction of protein cellular localization sites with the k nearest neighbors classifier." Proc Int Conf Intell Syst Mol Biol **5**: 147-152.
Howard, R. J., M. A. Ferrari, D. H. Roach and N. P. Money (1991). "Penetration of hard substrates by a fungus employing enormous turgor pressures." PNAS **88**(24): 11281-11284.
Howard, R. J. and B. Valent (1996). "Breaking and entering: host penetration by the fungal rice blast pathogen *Magnaporthe grisea*." Annu Rev Microbiol **50**: 491-512.
Hueck, C. J. (1998). "Type III protein secretion systems in bacterial pathogens of animals and plants." Microbiol Mol Biol Rev **62**(2): 379-433.
Hulo, N., C. J. Sigrist, V. Le Saux, P. S. Langendijk-Genevaux, L. Bordoli, A. Gattiker, E. De Castro, P. Bucher and A. Bairoch (2004). "Recent improvements to the PROSITE database." Nucl Acids Res **32 (Database issue)**: D134-137.
Izard, J. W. and D. A. Kendall (1994). "Signal peptides: exquisitely designed transport promoters." Mol Microbiol **13**(5): 765-773.
Jakel, S. and D. Gorlich (1998). "Importin beta, transportin, RanBP5 and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells." Embo J **17**(15): 4491-4502.
Jaton-Ogay, K., S. Paris, M. Huerre, M. Quadroni, R. Falchetto, G. Togni, J. P. Latge and M. Monod (1994). "Cloning and disruption of the gene encoding an extracellular metalloprotease of *Aspergillus fumigatus*." Mol Microbiol **14**(5): 917-928.
Joshi, L. and R. J. St Leger (1999). "Cloning, expression, and substrate specificity of MeCPA, a zinc carboxypeptidase that is secreted into infected tissues by the fungal entomopathogen *Metarhizium anisopliae*." J Biol Chem **274**(14): 9803-9811.
Kahmann, R. and C. Basse (2001). "Fungal gene expression during pathogenesis-related development and host plant colonization." Curr Opin Microbiol **4**(4): 374-380.
Kahmann, R., C. Basse and M. Feldbrugge (1999). "Fungal-plant signalling in the *Ustilago maydis*-maize pathosystem." Curr Opin Microbiol **2**(6): 647-650.
Kalderon, D., B. L. Roberts, W. D. Richardson and A. E. Smith (1984). "A short amino acid sequence able to specify nuclear location." Cell **39**(3 Pt 2): 499-509.

Kamoun, S., P. Hraber, B. Sobral, D. Nuss and F. Govers (1999). "Initial assessment of gene diversity for the oomycete pathogen *Phytophthora infestans* based on expressed sequences." Fungal Genet Biol **28**(2): 94-106.
Karlin, S. and S. F. Altschul (1990). "Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes." PNAS **87**(6): 2264-2268.
Kasamatsu, H. and A. Nakanishi (1998). "How do animal DNA viruses get to the nucleus?" Annu Rev Microbiol **52**: 627-686.
Kauzmann, W. (1959). "Some factors in the interpretation of protein denaturation." Adv Protein Chem **14**: 1-63.
Kelly, R. B. (1985). "Pathways of protein secretion in eukaryotes." Science **230**: 25-32.
Kim, J. Y., A. Mahe, J. Brangeon and J. L. Prioul (2000a). "A maize vacuolar invertase, IVR2, is induced by water stress. Organ/tissue specificity and diurnal modulation of expression." Plant Physiol **124**(1): 71-84.
Kim, J. Y., A. Mahe, S. Guy, J. Brangeon, O. Roche, P. S. Chourey and J. L. Prioul (2000b). "Characterization of two members of the maize gene family, Incw3 and Incw4, encoding cell-wall invertases." Gene **245**(1): 89-102.
Kim, Y. K., D. Li and P. E. Kolattukudy (1998). "Induction of Ca^{2+}-calmodulin signaling by hard-surface contact primes *Colletotrichum gloeosporioides* conidia to germinate and form appressoria." J Bacteriol **180**(19): 5144-5150.
Kolattukudy, P. E. (1985). "Enzymatic penetration of the plant cuticle by fungal pathogens." Annu Rev Phytopath **23**: 223-250.
Kolattukudy, P. E., L. M. Rogers, D. Li, C. S. Hwang and M. A. Flaishman (1995). "Surface signaling in pathogenesis." PNAS **92**(10): 4080-4087.
Kostova, Z. and D. H. Wolf (2003). "NEW EMBO MEMBER'S REVIEW: For whom the bell tolls: protein quality control of the endoplasmic reticulum and the ubiquitin-proteasome connection." Embo J **22**(10): 2309-2317.
Kosugi, S. and Y. Ohashi (2002). "Interaction of the *Arabidopsis* E2F and DP proteins confers their concomitant nuclear translocation and transactivation." Plant Physiol **128**(3): 833-843.
Kruger, W. M., C. Pritsch, S. Chao and G. J. Muehlbauer (2002). "Functional and comparative bioinformatic analysis of expressed genes from wheat spikes infected with *Fusarium graminearum*." MPMI **15**(5): 445-455.
Kurosawa, E. (1926). "Experimental studies on the nature of the substance secreted by the "bakanae" fungus." Nat HistSoc Formosa **16**: 213-227.
Laccetti, L., R. C. Staples and H. C. Hoch (1987). "Purification of calmodulin from bean rust uredospores." Experimental Mycology **11**(3): 231-235.
Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of baceriophage T4." Nature **227**: 680-685.
Lang, I., M. Scholz and R. Peters (1986). "Molecular mobility and nucleocytoplamsic flux in hepatoma cells." Journal of Cell Biology **102**: 1183-1190.
Lee, S. A., Y. Mao, Z. Zhang and B. Wong (2001). "Overexpression of a dominant-negative allele of YPT1 inhibits growth and aspartyl protease secretion in *Candida albicans*." Microbiology **147**(7): 1961-1970.
Lee, S. C. and Y. H. Lee (1998). "Calcium/calmodulin-dependent signaling for appressorium formation in the plant pathogenic fungus *Magnaporthe grisea*." Mol Cells **8**(6): 698-704.

Lim, C. O., H. Y. Kim, M. G. Kim, S. I. Lee, W. S. Chung, S. H. Park, I. Hwang and M. J. Cho (1996). "Expressed sequence tags of Chinese cabbage flower bud cDNA." Plant Physiol **111**(2): 577-588.

Link, T., G. Lohaus, I. Heiser, K. Mendgen, M. Hahn and R. T. Voegele (2005). "Characterization of a novel NADP +-dependent D-arabitol dehydrogenase from the plant pathogen *Uromyces fabae*." Biochem J.

Lipman, D. J., W. J. Wilbur, S. T. F. and M. S. Waterman (1984). "On the statistical significance of nucleic acid similarities." Nucl Acids Res **12**: 215-226.

Liu, Z., L. J. Szabo and W. R. Bushnell (1993). "Molecular cloning and analysis of abundant and stage-specific mRNAs from *Puccinia graminis*." MPMI **6**(1): 84-91.

Liu, Z. M. and P. E. Kolattukudy (1999). "Early expression of the calmodulin gene, which precedes appressorium formation in *Magnaporthe grisea*, is inhibited by self-inhibitors and requires surface attachment." J Bacteriol **181**(11): 3571-3577.

Luria, S. E., J. N. Adam and R. C. Teng (1960). "Transduction of lactose utilizing ability among strains of *Escherichia coli* and *Shigella dysenteriae* and the properties of the transducing phage particle." Virology **12**: 348-390.

Macara, I. G. (2001). "Transport into and out of the nucleus." Microbiol Mol Biol Rev **65**(4): 570-594.

MacFaddin, J. F. (1985). Media for isolation-cultivation-identification-maintenance of medical bacteria. Baltimore, Williams & Wilkins.

Maier, W., D. Begerow, M. Weiß and F. Oberwinkler (2003). "Phylogeny of the rust fungi: an approach using nuclear large subunit ribosomal DNA sequences." Can J Bot **81**(1): 12-23.

Mandel, M. and A. Higa (1970). "Calcium dependent bacteriophage DNA infection." Journal of Moleclular Biology **53**: 159-162.

Maniatis, T., E. F. Fritsch and J. Sambrook (1982). Molecular cloning. A laboratory manual. New York, Cold Spring Harbor Laboratory.

Martin, H. (2001). Charakterisierung eines putativen Aminosäuretransporters aus *Uromyces fabae*. Universität Konstanz, Konstanz, Germany

McGeoch, D. J. (1985). "On the predictive recognition of signal peptide sequences." Virus Res **3**(3): 271-286.

Melchior, F., B. Paschal, J. Evans and L. Gerace (1993). "Inhibition of nuclear protein import by nonhydrolyzable analogues of GTP and identification of the small GTPase Ran/TC4 as an essential transport factor." J Cell Biol **123**(6 Pt 2): 1649-1659.

Mellersh, D. G. and M. C. Heath (2003). "An investigation into the involvement of defense signaling pathways in components of the nonhost resistance of *Arabidopsis thaliana* to rust fungi also reveals a model system for studying rust fungal compatibility." MPMI **16**(5): 398-404.

Mendgen, K. (1996). Fungal attachement and penetration. Plant Cuticles. K. G. ed. Oxford, Bios. Scientific Publishers Ltd.**:** 175-188.

Mendgen, K. and M. Hahn (2002). "Plant infection and the establishment of fungal biotrophy." Trends Plant Sci **7**(8): 352-356.

Merkle, T. (2001). "Nuclear import and export of proteins in plants: a tool for the regulation of signaling." Planta **213**(4): 499-517.

Michal, G. (1993). Biochemical Pathways (Wallchart). Mannheim, Boehringer Mannheim GmbH, Biochemica.

Minsavage, G. V., D. Dahlbeck, M. C. Whalen, B. Kearny, U. Bonas, B. J. Staskawicz and R. E. Stall (1990). "Gene for-gene relationships specifying disease resistance in *Xanthomonas campestris pv. vesicatoria* – pepper interactions." MPMI **3**: 41–47.

Miyoshi, S. and S. Shinoda (2000). "Microbial metalloproteases and pathogenesis." Microbes Infect **2**(1): 91-98.

Möll, U. (2001). Charakterisierung der Invertase von *Uromyces fabae* auf molekularbiologischer und biochemischer Ebene. Universität Konstanz, Konstanz, Germany

Moore, M. S. and G. Blobel (1993). "The GTP-binding protein Ran/TC4 is required for protein import into the nucleus." Nature **365**(6447): 661-663.

Moroianu, J., G. Blobel and A. Radu (1995). "Previously identified protein of uncertain function is karyopherin alpha and together with karyopherin beta docks import substrate at nuclear pore complexes." PNAS **92**(6): 2008-2011.

Muesch, A., E. Hartmann, K. Rohde, A. Rubartelli, R. Sitia and T. A. Rapoport (1990). "A novel pathway for secretory proteins?" Trends Biochem Sci **15**(3): 86-88.

Müller, E. (2001). Molekulare Charakteriesierung und Sequnzierung des Aminosäuretransportergens *AAT3* (Syn. *PIG34*) des *Uromyces fabae*. Universität Konstanz, Konstanz, Germany

Müller, K. O. and H. Börger (1940). "Experimentelle Untersuchungen über die *Phytophthora*-Resistenz der Kartoffel - zugleich ein Beitrag zum Problem der "erworbenen Resistenz" im Pflanzenreich." Arbeiten Biol Reichsanst Land- Forstwirtsch Berlin-Dahlem **23**: 189-231.

Murphy, J. M. and J. D. Walton (1996). "Three extracellular proteases from *Cochliobolus carbonum*: cloning and targeted disruption of *ALP1*." MPMI **9**(4): 290-297.

Murray, N. E., W. J. Brammar and K. Murray (1977). "Lambdoid phages that simplify the recovery of in vitro recombinants." Mol Gen Genet **150**(1): 53-61.

Mysore, K. S., B. Bassuner, X. B. Deng, N. S. Darbinian, A. Motchoulski, W. Ream and S. B. Gelvin (1998). "Role of the *Agrobacterium tumefaciens* VirD2 protein in T-DNA transfer and integration." MPMI **11**(7): 668-683.

Mysore, K. S. and C. M. Ryu (2004). "Nonhost resistance: how much do we know?" Trends Plant Sci **9**(2): 97-104.

Nagai, K., C. Oubridge, N. Ito, J. Avis and P. Evans (1995). "The RNP domain: a sequence-specific RNA-binding domain involved in processing and transport of RNA." Trends Biochem Sci **20**(6): 235-240.

Nakai, K. and P. Horton (1999). "PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization." Trends Biochem Sci **24**(1): 34-36.

Nakai, K. and M. Kanehisa (1991). "Expert system for predicting protein localization sites in gram- negative bacteria." Proteins **11**(2): 95-110.

Nakai, K. and M. Kanehisa (1992). "A knowledge base for predicting protein localization sites in eukaryotic cells." Genomics **14**(4): 897-911.

Nelson, M. A., S. Kang, E. L. Braun, M. E. Crawford, P. L. Dolan, P. M. Leonard, J. Mitchell, A. M. Armijo, L. Bean, E. Blueyes, T. Cushing, A. Errett, M. Fleharty, M. Gorman, K. Judson, R. Miller, J. Ortega, I. Pavlova, J. Perea, S. Todisco, R. Trujillo, J. Valentine, A. Wells, M. Werner-Washburne and D. O. Natvig (1997). "Expressed sequences from conidial, mycelial, and sexual stages of *Neurospora crassa*." Fungal Genet Biol **21**(3): 348-363.

Nielsen, H., S. Brunak and G. von Heijne (1999). "Machine learning approaches for the prediction of signal peptides and other protein sorting signals." Protein Eng **12**(1): 3-9.

Nielsen, H., J. Engelbrecht, S. Brunak and G. von Heijne (1997a). "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites." Protein Eng **10**(1): 1-6.

Nielsen, H., J. Engelbrecht, S. Brunak and G. von Heijne (1997b). "A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites." Int J Neural Syst **8**(5-6): 581-599.

Nielsen, H. and A. Krogh (1998). Prediction of signal peptides and signal anchors by a hidden Markov model. Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology, AAAI Press, Menlo Park, California.

Norkin, L. C. (1999). "Simian virus 40 infection via MHC class I molecules and caveolae." Immunol Rev **168**: 13-22.

Nugent, K. G., K. Choffe and B. J. Saville (2004). "Gene expression during *Ustilago maydis* diploid filamentous growth: EST library creation and analyses." Fungal Genet Biol **41**(3): 349-360.

Osbourn, A. E. (1999). "Antimicrobial phytoprotectants and fungal pathogens: a commentary." Fungal Genet Biol **26**(3): 163-168.

Palmeri, D. and M. H. Malim (1999). "Importin beta can mediate the nuclear import of an arginine-rich nuclear localization signal in the absence of importin alpha." Mol Cell Biol **19**(2): 1218-1225.

Paschal, B. M., C. Delphin and L. Gerace (1996). "Nucleotide-specific interaction of Ran/TC4 with nuclear transport factors NTF2 and p97." PNAS **93**(15): 7679-7683.

Paschal, B. M. and L. Gerace (1995). "Identification of NTF2, a cytosolic factor for nuclear import that interacts with nuclear pore complex protein p62." J Cell Biol **129**(4): 925-937.

Pauling, L., D. H. Campbell and D. Pressman (1943). "The nature of the forces between antigen and antibody and of the precipitation reaction." Physiol Rev **23**: 203-219.

Peery, T., T. Shabat-Brand, R. Steinlauf, Y. Koltin and J. Bruenn (1987). "Virus-encoded toxin of *Ustilago maydis*: two polypeptides are essential for activity." Mol Cell Biol **7**(1): 470-477.

Pieterse, C. M., E. P. Risseeuw and L. C. Davidse (1991). "An *in planta* induced gene of *Phytophthora infestans* codes for ubiquitin." Plant Mol Biol **17**(4): 799-811.

Pieterse, C. M., H. M. Verbakel, J. H. Spaans, L. C. Davidse and F. Govers (1993). "Increased expression of the calmodulin gene of the late blight fungus *Phytophthora infestans* during pathogenesis on potato." MPMI **6**(2): 164-172.

Plinius, G. S. (ca. 77). Naturalis historia, liber XVIII.

Poussereau, N., S. Creton, G. Billon-Grand, C. Rascle and M. Fevre (2001a). "Regulation of *acp1*, encoding a non-aspartyl acid protease expressed during pathogenesis of *Sclerotinia sclerotiorum*." Microbiology **147**(3): 717-726.

Poussereau, N., S. Gente, C. Rascle, G. Billon-Grand and M. Fevre (2001b). "*aspS* encoding an unusual aspartyl protease from *Sclerotinia sclerotiorum* is expressed during phytopathogenesis." FEMS Microbiol Lett **194**(1): 27-32.

Prell, H. H. and P. R. Day (2001). Plant-Fungal Pathogen Interactions. A Classical and Molecular View. Berlin, Heidelberg, New York, Barcelona, Hong Kong, London, Milan, Paris, Singapore, Tokyo, Springer.

Prusky, D., J. L. McEvoy, R. Saftner, W. S. Conway and R. Jones (2004). "Relationship Between Host Acidification and Virulence of *Penicillium spp.* on Apple and Citrus Fruit." Phytopathology **94**(1): 44-51.

QIAexpressionist (2000). The QIAexpressionist™ A handbook for high-level expression and purification of 6xHis-tagged proteins, Qiagen.

Radelof, U., S. Hennig, P. Seranski, M. Steinfath, J. Ramser, R. Reinhardt, A. Poustka, F. Francis and H. Lehrach (1998). "Preselection of shotgun clones by oligonucleotide fingerprinting: an efficient and high throughput strategy to reduce redundancy in large-scale sequencing projects." Nucl Acids Res **26**(23): 5358-5364.

Radu, A., G. Blobel and M. S. Moore (1995). "Identification of a protein complex that is required for nuclear protein import and mediates docking of import substrate to distinct nucleoporins." PNAS **92**(5): 1769-1773.

Rauscher, M., A. L. Adam, S. Wirtz, R. Guggenheim, K. Mendgen and H. B. Deising (1999). "PR-1 protein inhibits the differentiation of rust infection hyphae in leaves of acquired resistant broad bean." Plant J **19**(6): 625-633.

Rauscher, M., K. W. Mendgen and H. Deising (1995). "Extracellular proteases of the rust fungus *Uromyces viciae-fabae*." Experimental Mycology **19**: 26-34.

Reichard, U., G. T. Cole, T. W. Hill, R. Ruchel and M. Monod (2000). "Molecular characterization and influence on fungal development of ALP2, a novel serine proteinase from Aspergillus fumigatus." Int J Med Microbiol **290**(6): 549-558.

Reifenberger, E., K. Freidel and M. Ciriacy (1995). "Identification of novel *HXT* genes in *Saccharomyces cerevisiae* reveals the impact of individual hexose transporters on glycolytic flux." Mol Microbiol **16**(1): 157-167.

Reinhardt, A. and T. Hubbard (1998). "Using neural networks for prediction of the subcellular location of proteins." Nucl Acids Res **26**(9): 2230-2236.

Rentsch, D., M. Laloi, I. Rouhara, E. Schmelzer, S. Delrot and W. B. Frommer (1995). "NTR1 encodes a high affinity oligopeptide transporter in *Arabidopsis*." FEBS Lett **370**(3): 264-268.

Robbins, J., S. M. Dilworth, R. A. Laskey and C. Dingwall (1991). "Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence." Cell **64**(3): 615-623.

Roelfs, A. P. (1982). "Effects of barberry eradication on stem rust in the United States." Plant Disease **66**: 171-181.

Rogers, L. M., M. A. Flaishman and P. E. Kolattukudy (1994). "Cutinase gene disruption in *Fusarium solani* f sp *pisi* decreases its virulence on pea." Plant Cell **6**(7): 935-945.

Saikkonen, K., S. H. Faeth, M. Helander and T. S. Sullivan (1998). "Fungal Endophytes: A Coninuum of Interactions with Host Plants." Annu Rev Ecol Syst **29**: 319-341.

Sambrook, J., E. F. Fritsch and T. Maniatis (1989). Molecular Cloning: a laboratory manual. New York, Cold Spring Harbor Laboratory Press.

Sbaghi, M., P. Jeandet, R. Bessis and P. Leroux (1996). "Degradation of stilbene-type phytoalexins in relation to the pathogenicity of *Botrytis cinerea* to grapevines." Plant Pathology **45**(1): 139-144.

Schulze Gronover, C., C. Schorn and B. Tudzynski (2004). "Identification of *Botrytis cinerea* genes up-regulated during infection and controlled by the Galpha subunit BCG1 using suppression subtractive hybridization (SSH)." MPMI **17**(5): 537-546.

Schumann, G. L. (1991). Plant Diseases: Their Biology and Social Impact, American Phytopathological Society.

Scott, R. R. E. and E. S. E. Delfose (1996). Biological Control of Weeds, CSIRO Publishing.

Sharp, P. A., B. Sugden and J. Sambrook (1973). "Detection of two restriction endonuclease activities in *Haemophilus parainfluencae* used in analytical agarose-Ethidium Bromide Electrophoresis." Biochemystry **12**: 3055-3063.

Shermann, F. (1991). Getting started with yeast. San Diego, CA, A. Press.

Shoichet, B. K., W. A. Baase, R. Kuroki and B. W. Matthews (1995). "A relationship between protein stability and protein function." PNAS **92**(2): 452-456.

Siewers, V., J. Smedsgaard and P. Tudzynski (2004). "The P450 monooxygenase BcABA1 is essential for abscisic acid biosynthesis in *Botrytis cinerea*." Appl Environ Microbiol **70**(7): 3868-3876.

Smith, T. F., M. S. Waterman and C. Burks (1985). "The statistical distribution of nucleic acid similarities." Nucl Acids Res **13**(2): 645-656.

Soanes, D. M., W. Skinner, J. Keon, J. Hargreaves and N. J. Talbot (2002). "Genomics of phytopathogenic fungi and the development of bioinformatic resources." MPMI **15**(5): 421-427.

Sohn, J. (2000). Untersuchungen pflanzen-induzierter Gene der Vitamin B1 Biosynthese beim Rostpilz *Uromyces fabae*. University of Konstanz, Konstanz, Germany

Sohn, J., R. T. Voegele, K. Mendgen and M. Hahn (2000). "High level activation of vitamin B1 biosynthesis genes in haustoria of the rust fungus *Uromyces fabae*." MPMI **13**(6): 629-636.

Southern, E. M. (1975). "Detection of specific sequences among DNA fragments seperated by gel electrophoresis." Journal of Moleclular Biology **98**: 503:517.

St Leger, R. J., L. Joshi and D. W. Roberts (1997). "Adaptation of proteases and carbohydrates of saprophytic, phytopathogenic and entomopathogenic fungi to the requirements of their ecological niches." Microbiology **143 (Pt 6)**: 1983-1992.

Stahl, D. J. and W. Schafer (1992). "Cutinase is not required for fungal pathogenicity on pea." Plant Cell **4**(6): 621-629.

Sterky, F., S. Regan, J. Karlsson, M. Hertzberg, A. Rohde, A. Holmberg, B. Amini, R. Bhalerao, M. Larsson, R. Villarroel, M. Van Montagu, G. Sandberg, O. Olsson, T. T. Teeri, W. Boerjan, P. Gustafsson, M. Uhlen, B. Sundberg and J. Lundeberg (1998). "Gene discovery in the wood-forming tissues of poplar: analysis of 5, 692 expressed sequence tags." PNAS **95**(22): 13330-13335.

Strack, D., T. Fester, B. Hause, W. Schliemann and M. H. Walter (2003). "Arbuscular mycorrhiza: biological, chemical, and molecular aspects." J Chem Ecol **29**(9): 1955-1979.

Strange, R. N. (1998). "Plants under attack II." Sci Prog **81**(1): 35-68.

Struck, C., M. Ernst and M. Hahn (2002). "Characterization of a developmetally regulated amino acid transporter (AAT1p) of the rust fungus *Uromyces fabae*." Mol Plant Pathol **3**(1): 23-30.

Struck, C., M. Hahn and K. Mendgen (1996). "Plasma Membrane H^+-ATPase Activity in Spores, Germ Tubes, and Haustoria of the Rust Fungus *Uromyces viciae-fabae*." Fungal Genet Biol **20**(1): 30-35.

Struck, C., E. Mueller, H. Martin and G. Lohaus (2004). "The *Uromyces fabae UfAAT3* gene encodes a general amino acid." Mol Plant Pathol **5**(3): 183-189.

Struck, C., C. Siebels, O. Rommel, M. Wernitz and M. Hahn (1998). "The plasma membrane H^+-ATPase from the biotrophic rust fungus *Uromyces fabae*: molecular characterization of the gene (*PMA1*) and functional expression of the enzymes in yeast." MPMI **11**(6): 458-465.

Studier, F. W. and B. A. Moffatt (1986). "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes." J Mol Biol **189**(1): 113-130.

Studier, F. W., A. H. Rosenberg, J. J. Dunn and J. W. Dubendorff (1990). "Use of T7 RNA polymerase to direct expression of cloned genes." Methods Enzymol **185**: 60-89.

Sweigard, J. A., A. M. Carroll, L. Farrall, F. G. Chumley and B. Valent (1998). "*Magnaporthe grisea* pathogenicity genes obtained through insertional mutagenesis." MPMI **11**(5): 404-412.

Szurek, B., O. Rossier, G. Hause and U. Bonas (2002). "Type III-dependent translocation of the Xanthomonas AvrBs3 protein into the plant cell." Mol Microbiol **46**(1): 13-23.

Tagu, D., B. Nasse and F. Martin (1996). "Cloning and and characterization of hydrophobins-encoding cDNAs from the ectomycorrhizal basdiomycete *Pisolithus tinctorius*." Gene **168**(1): 93-97.

Thomas, S. W., S. W. Rasmussen, M. A. Glaring, J. A. Rouster, S. K. Christiansen and R. P. Oliver (2001). "Gene identification in the obligate fungal pathogen *Blumeria graminis* by expressed sequence tag analysis." Fungal Genet Biol **33**(3): 195-211.

Tian, D., M. B. Traw, J. Q. Chen, M. Kreitman and J. Bergelson (2003). "Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*." Nature **423**(6935): 74-77.

Tobias, J. W., T. E. Shrader, G. Rocap and A. Varshavsky (1991). "The N-end rule in bacteria." Science **254**(5036): 1374-1377.

Truant, R. and B. R. Cullen (1999). "The arginine-rich domains present in human immunodeficiency virus type 1 Tat and Rev function as direct importin beta-dependent nuclear localization signals." Mol Cell Biol **19**(2): 1210-1217.

Tzfira, T., M. Vaidya and V. Citovsky (2001). "VIP1, an *Arabidopsis* protein that interacts with *Agrobacterium* VirE2, is involved in VirE2 nuclear import and *Agrobacterium* infectivity." Embo J **20**(13): 3596-3607.

Veena, H. Jiang, R. W. Doerge and S. B. Gelvin (2003). "Transfer of T-DNA and Vir proteins to plant cells by *Agrobacterium tumefaciens* induces expression of host genes involved in mediating transformation and suppresses host defense gene expression." Plant J **35**(2): 219-236.

Viaud, M. C., P. V. Balhadere and N. J. Talbot (2002). "A *Magnaporthe grisea* cyclophilin acts as a virulence determinant during plant infection." Plant Cell **14**(4): 917-930.

Vivian, A. and D. L. Arnold (2000). "Bacterial effector genes and their role in host–pathogen interactions." J Plant Pathol **82**: 163–178.

Voegele, R. T., M. Hahn, G. Lohaus, T. Link, I. Heiser and K. Mendgen (2005). "Possible roles for mannitol and mannitol dehydrogenase in the biotrophic plant pathogen *Uromyces fabae*." Plant Physiol **137**(1): 190-198.

Voegele, R. T. and K. Mendgen (2003). "Rust haustoria: nutrient uptake and beyond." New Phytol **159**: 93-100.

Voegele, R. T., C. Struck, M. Hahn and K. Mendgen (2001). "The role of haustoria in sugar supply during infection of broad bean by the rust fungus *Uromyces fabae*." PNAS **98**(14): 8133-8138.

Voiblet, C., S. Duplessis, N. Encelot and F. Martin (2001). "Identification of symbiosis-regulated genes in *Eucalyptus globulus-Pisolithus tinctorius* ectomycorrhiza by differential hybridization of arrayed cDNAs." Plant J **25**(2): 181-191.

Warshel, A. (1978). "Energetics of enzyme catalysis." PNAS **75**(11): 5250-5254.

Warshel, A., F. Sussman and J. K. Hwang (1988). "Evaluation of catalytic free energies in genetically modified proteins." J Mol Biol **201**(1): 139-159.

Werck-Reichhart, D. and R. Feyereisen (2000). "Cytochromes P450: a success story." Genome Biol **1**(6): 3003.

Wirsel, S. G., R. T. Voegele and K. W. Mendgen (2001). "Differential regulation of gene expression in the obligate biotrophic interaction of *Uromyces fabae* with its host *Vicia faba*." MPMI **14**(11): 1319-1326.

Wood, V., R. Gwilliam, M. A. Rajandream, M. Lyne, R. Lyne, A. Stewart, J. Sgouros, N. Peat, J. Hayles, S. Baker, D. Basham, S. Bowman, K. Brooks, D. Brown, S. Brown, T. Chillingworth, C. Churcher, M. Collins, R. Connor, A. Cronin, P. Davis, T. Feltwell, A. Fraser, S. Gentles, A. Goble, N. Hamlin, D. Harris, J. Hidalgo, G. Hodgson, S. Holroyd, T. Hornsby, S. Howarth, E. J. Huckle, S. Hunt, K. Jagels, K. James, L. Jones, M. Jones, S. Leather, S. McDonald, J. McLean, P. Mooney, S. Moule, K. Mungall, L. Murphy, D. Niblett, C. Odell, K. Oliver, S. O'Neil, D. Pearson, M. A. Quail, E. Rabbinowitsch, K. Rutherford, S. Rutter, D. Saunders, K. Seeger, S. Sharp, J. Skelton, M. Simmonds, R. Squares, S. Squares, K. Stevens, K. Taylor, R. G. Taylor, A. Tivey, S. Walsh, T. Warren, S. Whitehead, J. Woodward, G. Volckaert, R. Aert, J. Robben, B. Grymonprez, I. Weltjens, E. Vanstreels, M. Rieger, M. Schafer, S. Muller-Auer, C. Gabel, M. Fuchs, C. Fritzc, E. Holzer, D. Moestl, H. Hilbert, K. Borzym, I. Langer, A. Beck, H. Lehrach, R. Reinhardt, T. M. Pohl, P. Eger, W. Zimmermann, H. Wedler, R. Wambutt, B. Purnelle, A. Goffeau, E. Cadieu, S. Dreano, S. Gloux, V. Lelaure, S. Mottier, F. Galibert, S. J. Aves, Z. Xiang, C. Hunt, K. Moore, S. M. Hurst, M. Lucas, M. Rochet, C. Gaillardin, V. A. Tallada, A. Garzon, G. Thode, R. R. Daga, L. Cruzado, J. Jimenez, M. Sanchez, F. del Rey, J. Benito, A. Dominguez, J. L. Revuelta, S. Moreno, J. Armstrong, S. L. Forsburg, L. Cerrutti, T. Lowe, W. R. McCombie, I. Paulsen, J. Potashkin, G. V. Shpakovski, D. Ussery, B. G. Barrell and P. Nurse (2002). "The genome sequence of *Schizosaccharomyces pombe*." Nature **415**(6874): 871-880.

Wubben, J. P., M. H. Joosten and P. J. De Wit (1994). "Expression and localization of two in planta induced extracellular proteins of the fungal tomato pathogen *Cladosporium fulvum*." MPMI **7**(4): 516-524.

Xu, H. and M. C. Heath (1998). "Role of calcium in signal transduction during the hypersensitive response caused by basidiospore-derived infection of the cowpea rust fungus." Plant Cell **10**(4): 585-598.

Xu, H. and K. Mendgen (1991). "Early events in living epidermal cells of cowpea and broad bean during infection with basidiospores of the cowpea rust fungus." Can J Bot **69**: 2279-2285.

Xuei, X., S. Bhairi, R. C. Staples and O. C. Yoder (1992). "Characterization of INF56, a gene expressed during infection structure development of *Uromyces appendiculatus*." Gene **110**(1): 49-55.

Yamada, M. and H. Kasamatsu (1993). "Role of nuclear pore complex in simian virus 40 nuclear targeting." J Virol **67**(1): 119-130.

Yang, B., W. Zhu, L. B. Johnson and F. F. White (2000). "The virulence factor AvrXa7 of *Xanthomonas oryzae pv. oryzae* is a type III secretion pathway-dependent nuclear-localized double-stranded DNA-binding protein." PNAS **97**(17): 9807-9812.

Yao, C., W. S. Conway and C. E. Sams (1996). "Purification and characterization of a polygalacturonase produced by *Penicillium expansum* in apple fruit." Phytopathology **86**: 1160-1166.

Yoneda, Y., M. Hieda, E. Nagoshi and Y. Miyamoto (1999). "Nucleocytoplasmic protein transport and recycling of Ran." Cell Struct Funct **24**(6): 425-433.

Zhang, J. (2003). "Evolution by gene duplication: an update." Trends Ecol Evol **18**(6): 292-298.

Zhang, R., M. J. Dickinson and A. Pryor (1994). "Double-stranded RNAs in the rust fungi." Annu Rev Phytopathol **32**: 115-133.

Ziemienowicz, A., D. Gorlich, E. Lanka, B. Hohn and L. Rossi (1999). "Import of DNA into mammalian nuclei by proteins originating from a plant pathogenic bacterium." PNAS **96**(7): 3729-3733.

8 Annex

8.1 EST Contigs Displaying Similarity to Genes in the NCBI Database

In the following table, all contigs that were compared to the NCBI database during the EST project and which produced significant results are listed. A short description of the best hit is given. In the description, the accession number of the gene that produced the best hit (the lowest E value) and its name is listed. In doing so, the notation of the source organism was adopted. Therefore it is possible to find both the teleomorph and the anamorph name of the same fungus listed e.g. *Emericella nidulans* and *Aspergillus nidulans*. In cases where the function of an *Uromyces fabae* gene has already been verified, this has been explicitly mentioned. The assignment to functional categories was performed according to Kamoun et al. (1999), and an extra category for contigs of putative viral origin was added.

Table A-1 EST contigs displaying similarity to genes in the NCBI database, grouped according to probable function.

Contig	No of Clones	'Best hit' GI*	Description	BLASTX E value ‡
01 Metabolism: Σ 37 Contigs				
01.01 Amino acid metabolism: Σ 11 Contigs				
C037	1	399897	Imidazoleglycerol-phosphate Dehydratase (IGPD) (*Saccharomyces kluyveri*)	1e-54
C040	1	19115618	Delta 1-Pyrroline-5-Carboxylate Reductase (*Schizosaccharomyces pombe*)	6e-64
C069	3	1346101	4-Aminobutyrate Aminotransferase (Gamma-Amino-N-Butyrate Transaminase) (*Ustilago maydis)*	4e-35
C111	1	2507455	Formate—Tetrahydrofolateligase/ 5,10-Methylenetetrahydrofolate Dehydrogenase (*Spinacia oleracea*)	3e-44
C127 // C257	1	19113203	Saccharopine Dehydrogenase (*Schizosaccharomyces pombe*)	1e-55
C170	2	2494742	Glutamine Synthetase (*Agaricus bisporus*)	4e-73
C206	1	15077830	Homoserine Dehydrogenase (*Filobasidiella neoformans*)	3e-29
C308	1	6319901	Threonine Synthase (*Saccharomyces cerevisiae*)	2e-10
C398	2	19076029	Put. Prephenate Dehydrogenase (*Schizosaccharomyces pombe*)	2e-47
C462	1	19112372	Put. Adenosylhomocysteinase (*Schizosaccharomyces pombe*)	9e-59
C554	1	2494072	Delta-1-Pyrroline-5-Carboxylate Dehydrogenase Precursor P5C (*Agaricus bisporus*)	4e-60

Contig	No of Clones	'Best hit' GI*	Description	BLASTX E value ‡
01.02 Nitrogen and sulfur metabolism: Σ 1 Contigs				
C504	1	16127220	Conserved hypothetical protein/ predicted 2-nitropropane dioxygenase (*Caulobacter crescentus*)	3e-60
01.03 Nucleotide metabolism: Σ 6 Contigs				
C073	1	15616326	Uracil Phosphoribosyltransferase (*Bacillus halodurans*)	5e-40
C082	2	11135375	Thioredoxin (*Coprinus comatus*)	2e-19
C083	1	133242	Ribonuclease RNAse (*Trichoderma viride*)	5e-07
C084	1	6321920	Fur1p/ uracil phosphoribosyl transferase UPRTase (*Saccharomyces cerevisiae*)	4e-27
C171	1	13787042	Guanylate Kinase (*Saccharomyces cerevisiae*)	1e-38
C229	1	19743612	Nucleoside Diphosphate Kinase (*Emericella nidulans*)	2e-51
01.04 Phosphate metabolism: Σ 0 Contigs				
01.05 Carbohydrate metabolism: Σ 7 Contigs				
C008	17	12328513	Put. Trehalose-phosphate synthase / *PIG15* (*Oryza sativa*)	5e-08
C023	13	AAB39878	*PIG8*/ Probable NADP-dependent mannitol dehydrogenase (MtDH) (Mannitol 2-dehydrogenase [NADP+]) (*Uromyces fabae*)	1e-140
C045	1	8050874	Probable subunit of 1,3-Beta-Glucan Synthase (*Mortierella alpina*)	5e-43
C064	1	493580	Beta-D-Glucoside Glucohydrolase/ extracellular Beta Glucosidase (*Hypocrea jecorina*)	4e-40
C098	8	21220388	Polyol Dehydrogenase (*Streptomyces coelicolor*) *D-arabitol dehydrogenase (*Uromyces fabae*)	2e-48 *verified (Link et al. 2005)
C370	2	15620806	Invertase (*Aspergillus niger*)	5e-18
C490	1	18376290	Related to 6-Phosphofructo-2-Kinase/ Fructose-2,6-Bisphosphatase Fructose-6-Phosphate 2-Kinase (*Neurospora crassa*)	4e-28
01.06 Lipid and sterol metabolism: Σ 5 Contigs				
C075	1	135757	Acetyl-CoA Acetyltransferase (*Rattus norvegicus*)	3e-51
C253	1	13542175	3-Oxoacyl-[acyl-carrier-protein] Reductase (*Thermoplasma volcanium*)	5e-14
C377	1	1346032	Farnesyl-diphosphate Farnesyltransferase (*Kluyveromyces lactis*)	1e-47
C415	1	1332597	Triglyceride Lipase (*Saccharomyces cerevisiae*)	1e-24
C541	1	19115257	Put. Acyl carrier Protein, Mitochondrial Protein (*Schizosaccharomyces pombe*)	2e-17

Contig	No of Clones	'Best hit' GI*	Description	BLASTX E value ‡
01.07 Biosynthesis of vitamins, cofactors and prosthetic groups: Σ 4 Contigs				
C003	36	1764131	*THI1* (*PIG1*) (*Uromyces fabae*)	0.0 (verified)
C051	17	12230775	*PIG4*/ THI2p / Thiazole Biosynthetic Enzyme (*Uromyces fabae*)	0.0 (verified)
C367	1	3123310	Put. Flavoprotein (Monodehydroascorbate Reductase) (*Schizosaccharomyces pombe*)	1e-20
C581	1	1764131	*PIG1* similar (*Uromyces fabae*)	1e-06
01.08 Aminosugar metabolism: Σ 2 Contigs				
C538	1	21357745	Glucosamine--Fructose-6-Phosphate Aminotransferase (*Drosophila melanogaster*)	4e-84
C011	10	31087953	Chitinase *PIG18* (*Puccinia triticina*)	9e-157
02 Energy: Σ 37 Contigs				
02.01 Glycolysis & Pyruvat Metabolism Σ 12 Contigs				
C034	3	19112484	Fructose-Bisphosphate Aldolase (*Schizosaccharomyces pombe*)	1e-110
C036	1	12643655	Pyruvate Kinase (*Agaricus bisporus*)	1e-70
C054	5	15722244	Glyceralaldehyde 3-Phosphate Dehydrogenase (*Mucor racemosus*)	1e-114
C056	5	12230023	Enolase 1 (2-Phosphoglycerate Dehydratase) (*Cunninghamella elegans*)	1e-166
C146	2	17988592	Phosphoglycerate Mutase (*Brucella melitensis*)	2e-44
C176	1	266757	Phosphoglycerate Kinase (*Yarrowia lipolytica*)	4e-58
C193	1	21300334	Put. Dihydrolipoamide Acetyltransferase (*Anopheles gambiae*)	1e-32
C259	1	20067633	Glucose-6-Phosphate Isomerase (gpi) (*Bufo melanostictus*)	4e-61
C401	1	18376030	Related To Pyruvate Dehydrogenase Kinase Isoform (*Neurospora crassa*)	4e-35
C443	1	11260590	Pyruvate Dehydrogenase kinase (*Arabidopsis thaliana*)	1e-07
C498	1	136053	Triosephosphate Isomerase (*Emericella nidulans*)	2e-32
C520	1	171429	Pyruvate Dehydrogenase E1-beta subunit (*Saccharomyces cerevisiae*)	2e-36
02.02 Gluconeogenes: Σ 1 Contig				
C542	1	11268307	Citrate Synthase, Mitochondrial Precursor (*Neurospora crassa*)	3e-54
02.03 Pentose phosphate pathway: Σ 2 Contigs				
C002	5	21954534	6-Phosphogluconate Dehydrogenase (*Aspergillus oryzae*)	1e-169
C458	1	3334110	6-Phosphogluconate Dehydrogenase, Decarboxylating (*Candida albicans*)	4e-11
02.04 TCA pathway: Σ 8 Contigs				
C095	7	19702487	Malate Dehydrogenase (*Talaromyces*	1e-105

Contig	No of Clones	'Best hit' GI*	Description	BLASTX E value ‡
			emersonii)	
C128	1	2706555	Succinate Dehydrogenase Iron-Sulfur Subunit (*Agaricus bisporus*)	1e-28
C142	1	1942164	Fumarase, Chain A, Fumarase With Bound Pyromellitic Acid (*Escherichia coli*)	1e-26
C149	6	4029338	Malate Dehydrogenase (*Piromyces* sp. E2)	5e-71
C386	1	11252276	Alpha-Ketoglutarate Dehydrogenase / 2-Oxoglutarate Dehydrogenase (*Neurospora crassa*)	3e-28
C413	1	CAC18164.2	Related To Malate Dehydrogenase/ Put. Cytosolic NADP-Malic Enzyme (*Neurospora crassa*)	2e-34
C418	1	NP_298170.1	Carbonic Anhydrase (*Xylella fastidiosa*)	1e-24
C489	1	6325122	Iron-Sulfur Cluster nifU-like Protein (*Saccharomyces cerevisiae*)	3e-26
02.05 Respiration: Σ 15 Contigs				
C112	1	19112733	Ubiquinol-Cytochrome c Reductase Iron-Sulpher Subunit (*Schizosaccharomyces pombe*)	3e-71
C117	4	P49377	ATP Synthase Gamma Chain/ mitochondrial (*Kluyveromyces lactis*)	4e-40
C118	1	P00011	Cytochrome c (*Canis familiaris*)	9e-34
C138	3	NP_015247.1	ATP Synthase F0 Sector Subunit 4 (*Saccharomyces cerevisiae*)	8e-46
C143	2	19115831|	ATP Synthase Alpha Chain, Mitochondrial (*Schizosaccharomyces pombe*)	1e-135
C199	3	17506915	ATPase (Caenorhabditis elegans)	2e-31
C246	1	4028029	Cytochrome b (*Pneumocystis carinii f. sp. carinii*)	4e-64
C295	1	128862	NADH-Ubiquinone Oxidoreductase 30.4 kD Subunit Precursor (*Neurospora crassa*)	9e-27
C306	1	20137714	Cytochrome c Oxidase Subunit V (*Aspergillus niger*)	2e-24
C315	1	15088723	NADH-Ubiquinone Oxidoreductase Chain 5 (*Schizophyllum commune*)	2e-53
C318	1	6689660	Subunit NUKM of Protein NADH:Ubiquinone Oxidoreductase (*Yarrowia lipolytica*)	3e-45
C371	4	NP_593151.1	ATP Synthase Beta Chain, Mitochondrial (*Schizosaccharomyces pombe*)	6e-80
C408	1	128868	NADH-Ubiquinone Oxidoreductase, 29/21K Chain Precursor (*Neurospora crassa*)	1e-07
C444	1	3288188	Cytochrome c Oxidase Subunit I (*Diospyros virginiana*)	5e-58
C568	1	2493049	ATP Synthase Delta Chain, Mitochondrial (*Agaricus bisporus*)	3e-32
02.06 Fermentation: Σ 0 Contigs				

Contig	No of Clones	'Best hit' GI*	Description	BLASTX E value ‡
03 Cell growth/ Cell division/ DNA synthesis: Σ 8 Contigs				
03.01 Cell growth: Σ 0 Contigs				
03.02 Meiosis: Σ 0 Contigs				
03.03 DNA synthesis/ replication: Σ 3 Contigs				
C158	1	18376276	Probable DNA-Directed DNA Polymerase III (*Neurospora crassa*)	1e-34
C202	1	11282390	Helicase Homolog (*Schizosaccharomyces pombe*)	3e-19
C281	1	19075945	Helicase (*Schizosaccharomyces pombe*)	5e-07
03.04 Recombination/ DNA repair: Σ 0 Contigs				
03.05 Cell cycle/ Cell cycle control: Σ 1 Contig				
C445	1	19111914	Ran-specific GTPase activating protein (*Schizosaccharomyces pombe*)	9e-36
03.06 Cytokinesis: Σ 4 Contigs				
C027	1	1729837	Tubulin Alpha-1a Chain (*Schizophyllum commune)*	3e-37
C239	1	13398364l	Septin (*Aspergillus nidulans*)	2e-79
C375	1	15230956	Put. Mitotic Checkpoint Protein (*Arabidopsis thaliana*)	6e-13
C389	1	5725417	Put. Septin (*Pyrenopeziza brassicae)*	6e-75
03.07 Growth regulators: Σ 0 Contigs				
03.08 Other proteins: Σ 0 Contigs				
04 Transcription : Σ 15 Contigs				
04.01 rRNA synthesis : Σ 0 Contigs				
04.02 tRNA synthesis : Σ 0 Contigs				
04.03 RNA transport : Σ 0 Contigs				
04.04 mRNA synthesis: Σ 4 Contigs				
C221	2	19919850	Putative Multi-Protein Binding Factor 1 Transcriptional Activation (*Yarrowia lipolytica*)	2e-28
C333	1	5679035l	BcDNA.GM05329 Transcription factor (*Drosophila melanogaster*)	8e-17
C470	1	19114415	Put. Component Of RNA Polymerase II Holoenzyme and Mediator Subcomplex; RNA Polymerase Transcriptional Regulation Mediator (*Saccharomyces cerevisiae*)	1e-07
C492	1	15829017	DNA-Directed RNA Polymerase Beta Chain (*Mycoplasma pulmonis*)	6e-13
04.05 General transcription factors: Σ 4 Contigs				
C104	1	AAA79367.1	TATA Binding Protein//Transcription Initiation Factor TFIID (*Pneumocystis carinii*)	9e-79
C132	1	5531330	Putative Transcription Factor (*Nicotiana tabacum*)	9e-09
C181	3	NP_594523.1	Regulatory Protein // Positive Sulphur Transcription Regulator	2e-16

Contig	No of Clones	'Best hit' GI*	Description	BLASTX E value ‡
			(*Schizosaccharomyces pombe*)	
C427	1	AAF57272.2	Zinc Finger Protein (*Drosophila melanogaster*)	4e-07
04.06 Transcriptional control: Σ 2 Contigs				
C400	2	NP_197262.1	Zinc Finger Protein (*Arabidopsis thaliana*)	4e-09
C331	1	NP_197262.1	Zinc Finger Protein (*Arabidopsis thaliana*)	4e-09
04.07 Chromatin modifier: Σ 0 Contigs				
04.08 mRNA processing: Σ 3 Contigs				
C189	1	NP_594207.1	Protein with 3 RNA binding domains (*Schizosaccharomyces pombe*)	1e-07
C486	2	NP_651322.1	RNA splicing-related protein (*Arabidopsis thaliana*)	4e-60
C488	1	NP_595715.1	MLO3 Protein, Nuclear Protein; RNA-binding (*Schizosaccharomyces pombe*)	3e-23
04.09 Other proteins involved in transcription: Σ 1 Contig				
C395	1	XP_058547.5	Similar to Zinc Finger Protein (*Homo sapiens*)	9e-13
05 Protein synthesis: Σ 29 Contigs				
05.01 Ribosomal proteins: Σ 19 Contigs				
C044	1	14285761	Ribosomal Protein S28 (*Neurospora crassa*)	1e-66
C049	1	19075843	40s Ribosomal Protein (*Schizosaccharomyces pombe*)	3e-16
C100	1	2507312	60s Ribosomal Protein (*Puccinia graminis f. sp. avenae*)	2e-67
C130	2	19075214	60s Ribosomal Protein L24 (*Schizosaccharomyces pombe*)	4e-25
C168	1	17865559	Ribosomal Protein L39 (*Trichoderma hamatum*)	4e-20
C169	1	17865559	60s Ribosomal Protein (*Trichoderma hamatum*)	4e-20
C226	1	19115094	60s Ribosomal Protein L32-A (*Schizosaccharomyces pombe*)	1e-38
C293	1	6322847	Homology To Mammalian L14; Rpl14ap Ribosomal Protein L14 (*Saccharomyces cerevisiae*)	3e-26
C294	1	6321659	Mitochondrial Ribosomal Protein MRPL9 (*Saccharomyces cerevisiae*)	9e-32
C360	2	19114915	40s Ribosomal Protein S11 (*Schizosaccharomyces pombe*)	2e-49
C394	1	730576	60s Ribosomal Protein L8 (*Xenopus laevis*)	1e-100
C410	1	6831665	40S Ribosomal Protein S5 (*Cicer arietinum*)	6e-60
C425	1	5757449	L41 Ribosomal Protein (*Coprinus cinereus*)	8e-54
C436	1	12329977	Ribosomal Protein L13A (*Xanthophyllomyces dendrorhous*)	2e-70
C438	1	19112528	60s Ribosomal Protein (*Schizosaccharomyces	2e-58

Contig	No of Clones	'Best hit' GI*	Description	BLASTX E value ‡
			pombe)	
C465	1	15236171	Putative S-Phase-Specific Ribosomal Protein (*Arabidopsis thaliana*)	2e-38
C482	1	AAL08563.1	Putative Ribosomal Protein L35 (*Paracoccidioides brasiliensis*)	3e-29
C531	1	AAM09311.1	Hypothetical Protein 60s Ribosomal Protein L15 (*Dictyostelium discoideum*)	7e-62
C547	1	NP_595832.1	60s Ribosomal Protein L7a (*Schizosaccharomyces pombe*)	2e-40
05.02 Translational factors: Σ 2 Contigs				
C303	2	T42745	Hypothetical Protein Eukaryotic Translation Initiation Factor (*Schizosaccharomyces pombe*)	8e-61
C572	1	P56329	Translation Initiation Factor 2 (*Schizosaccharomyces pombe*)	3e-61
05.03 Translational control: Σ 0 Contigs				
05.04 tRNA synthesis: Σ 2 Contigs				
C165	1	CAA26666.1	Threonyl-tRNA synthetase (*Saccharomyces cerevisiae*)	9e-50
C507	1	Q9HGT6	Seryl-tRNA synthetase, cytoplasmic (*Candida albicans*)	1e-26
05.05 Other proteins involved in protein synthesis: Σ 4 Contigs				
C328	1	17981723	Protein-L-Isoaspartate(D-Aspartate) O-Methyltransferase (*Drosophila melanogaster*)	4e-27
C340	1	19114887	Elongation Factor 2 (*Schizosaccharomyces pombe)*	3e-40
C502	1	17149842	FK506-Binding Protein 2 Rapamycin-Binding Protein (*Homo sapiens*)	2e-31
C524	2	19075803	Elongation Factor 1 (*Schizosaccharomyces pombe*)	9e-31
06 Protein destination: Σ 16 Contigs				
06.01 Folding and stabilization of proteins: Σ 1 Contigs				
C548	1	2501203	Protein Disulfide Isomerase Precursor (*Aspergillus oryzae*)	5e-40
06.02 Targeting/ sorting/ translocation of proteins: Σ 2 Contigs				
C523	1	5006903	Mitochondrial Processing Peptidase Beta Subunit (*Lentinula edodes*)	4e-48
C573	1	13324618	TOM7 (*Neurospora crassa*)	5e-07
06.03 Modification of proteins: Σ 0 Contigs				
06.04 Assembly of protein complexes: Σ 0 Contigs				
06.05 Proteolysis: Σ 13 Contigs				
C086	1	11359605	Related to 26s Proteasome Subunit P28 (*Neurospora crassa*)	5e-10
C123	1	6015089	Extracellular Elastinolytic Metalloproteinase Precursor (*Aspergillus fumigatus*)	6e-23

Contig	No of Clones	'Best hit' GI*	Description	BLASTX E value ‡
C155	1	18152941l	Proteinase A (*Pichia angusta*)	4e-46
C210	1	12718381	Put. Zinc Metallo-Protease (*Neurospora crassa*)	2e-42
C212	1	20843105	Proteasome Subunit C2 (*Mus musculus*)	1e-53
C256	1	15238924	Ubiquitin-like Protein (*Arabidopsis thaliana*)	1e-20
C279	1	6319430	Proteasome Subunit (*Saccharomyces cerevisiae*)	1e-50
C437	1	19115284	Proteasome Component (*Schizosaccharomyces pombe*)	9e-88
C450	1	19112827	Ubiquitin-like Modifier (*Schizosaccharomyces pombe*)	6e-08
C537	1	19909972	Leucine Aminopeptidase (*Coprinopsis cinerea*)	1e-48
C545	1	2143220	Cellular Serine Proteinase (*Aspergillus fumigatus*)	e-104
C569	1	6015089	Metalloproteinase (*Aspergillus fumigatus*)	6e-23
06.06 Other proteins involved in protein destination: Σ 0 Contigs				
07 Transport facilitators: Σ 12 Contigs				
07.01 Ion channels: Σ 0 Contigs				
07.02 Ion transporters: Σ 0 Contigs				
07.03 Sugar/ carbohydrate transporters: Σ 2 Contigs				
C224	3	CAC41332 14348970	HXT1p (Hexose Transporter) (*Uromyces vicia-fabae*)	2e-79 (verified)
C243	1	NP_493723.1	Hypothetical Protein C03H5.2 (*Caenorhabditis elegans*) / (UDP-Galactose Transporter Homologue (*Schizosaccharomyces pombe*))	2e-20 (8e-20)
07.04 Amino Acid Transporters: Σ 3 Contigs				
C013	4	1764098	Permease/ Amino Acid Transporter *PIG2* (*Uromyces fabae*)	0.0 (verified)
C015	2	CAC67419.1	Put. Permease, *PIG27* Amino Acid Transporter (*Uromyces fabae*)	2e-77 (verified)
C217	1	CAF32328	Amino Acid Transporter (*PIG34*) (*Uromyces viciae-fabae*)	e-134 (verified)
07.05 Lipid transporters: Σ 0 Contigs				
07.06 Purine/ pyrimidine transporters: Σ 1 Contig				
07.07 Transport ATPases: Σ 4 Contigs				
C144// C258	1	Q40635	Vacuolar ATP Synthase (*Oryza sativa*)	4e-22
C160	1	11267279	Na+/K+-exchanging ATPase (P-type cation-transporting) (*Blastocladiella emersonii*)	9e-18
C161	4	P02723	ADP,ATP Carrier Protein (*Neurospora crassa*)	3e-72
C551	2	CAA05841.1	PM (H+) ATPase PMA1 (*Uromyces fabae*)	e-138
07.08 ABC transporters: Σ 0 Contigs				

Contig	No of Clones	'Best hit' GI*	Description	BLASTX E value ‡
07.09 Other transport facilitators: Σ 3 Contigs				
C030	1	6716399	Peptide Transporter MTD1 (*Schizosaccharomyces pombe*)	6e-12
C247	4	20942368	Similar to put. Integral Membrane Transporter (*Mus musculus*)	6e-20
C419	1	6321354	Involved in active glycerol uptake; Gup1p (*Saccharomyces cerevisiae*)	1e-52
08 Intracellular traffic: Σ 8 Contigs				
08.01 Nuclear transport: Σ 1 Contig				
C115	1	Q9P926	Nuclear Transport Factor (*Saccharomyces cerevisiae*)	2e-33
08.02 Mitochondrial transport: Σ 6 Contigs				
C115	1	15236269	Frataxin Homolog (Arabidopsis thaliana)	5e-21
C116	1	NP_594211.1	Yeast Putative Mitochondrial Carrier// Probable Membrane Protein (*Schizosaccharomyces pombe*)	2e-26
C265	1	19112220	Putative Mitochondrial Protein Import Protein – DNA J Protein (*Schizosaccharomyces pombe*)	3e-54
C289	1	130684	Outer Mitochondrial Membrane Porin (*Neurospora crassa*)	9e-08
C312	3	NP_596208.1	Mitochondrial Phosphate Transporter (*Schizosaccharomyces pombe*)	8e-51
C578	1	6324198	Translocase of outer Mitochondrial Membrane; TOM22p (*Saccharomyces cerevisiae*)	4e-10
08.03 Vesicular transport: Σ 1 Contig				
C430	1	6325247	Vesicle coat component (*Saccharomyces cerevisiae*)	3e-32
08.04 Extracellular transport: Σ 0 Contigs				
08.05 Cellular import: Σ 0 Contigs				
08.06 Other proteins involved in intracellular traffic: Σ 0 Contigs				
09 Cellular organization and biogenesis: Σ 9 Contigs				
09.01 Cell wall/ plasma membrane: Σ 0 Contigs				
09.02 Cytoplasm: Σ 0 Contigs				
09.03 Cytoskeleton: Σ 3 Contigs				
C020	8	284667	Heavy Neurofilament Subunit *PIG9* (*Rattus norvegicus*)	2e-06
C429	1	15865475	Beta1-Tubulin (*Uromyces fabae*)	9e-60
C449	1	584708	Actin Binding Protein (*Saccharomyces exiguous*)	1e-15
09.04 ER/ golgi: Σ 0 Contigs				
09.05 Nuclear: Σ 3 Contigs				
C296	2	11121216	Histone H3 (*Mortierella alpina*)	6e-29
C358	4	462244	Histone H4 (*Phanerochaete chrysosporium*)	1e-37
C275	1	462244	Histone H4 (*Phanerochaete chrysosporium*)	3e-38

Contig	No of Clones	'Best hit' GI*	Description	BLASTX E value ‡
09.06 Chromosome structure: Σ 0 Contigs				
09.07 Mitochondrial: Σ 0 Contigs, 15 Contigs are listed in other categories				
09.08 Peroxisomal: Σ 0 Contigs				
09.09 Vacuolar: Σ 0 Contigs				
09.10 Other proteins involved in cellular organization and biogenesis: Σ 3 Contigs				
C283	1	19112281	Yeast Reduced Viability upon Starvation Protein 161 Homolog, Implicated in Cell Growth and Cytoskeletal org. (*Schizosaccharomyces pombe*)	4e-56
C314	1	13359451	Put. Senescence-Associated Protein (*Pisum sativum*)	3e-45
C324	1	19113894	Negative Regulator of Differentiation (*Schizosaccharomyces pombe*)	8e-08
10 Signal transduction: Σ 8 Contigs				
10.01 Receptors: Σ 0 Contigs				
10.02 Mediators: Σ 1 Contig				
C287	1	17647231	Calmodulin (*Drosophila melanogaster*)	5e-55
10.03 Kinases: Σ 1 Contig				
C540	1	1438945	Serine/Threonine Protein Kinase FSK (*Nectria haematococca*)	5e-27
10.04 Phosphatases: Σ 1 Contig				
C188	1	19075920	Serine/Threonine Protein Phosphatase (*Schizosaccharomyces pombe*)	2e-48
10.05 G proteins: Σ 2 Contigs				
C216	1	9754765	Guanine Nucleotide-Binding Protein (*Aspergillus nidulans*)	2e-72
C527	1	19075931	Rho1 GTP-Binding Protein (*Mucor rouxii*)	7e-66
10.06 Other proteins involved in signal transduction: Σ 2 Contigs				
C070	1	19075931	Putative receptor-associated protein (*Schizosaccharomyces pombe*)	5e-16
C269	1	19112217	Similarity to Mammalian Signal Transducing Adaptor (*Schizosaccharomyces pombe*)	9e-31
11 Pathogenicity factors, avirulence factors and elicitors: Σ 0 Contigs				
11.01 Pathogenicity factors: Σ 0 Contigs				
11.02 Avirulence factors: Σ 0 Contigs				
12 Cell defense: Σ 15 Contigs				
12.01 Stress response: Σ 5 Contigs				
C209	2	7549229	Heat Shock Protein 90 (*Candida tropicalis*)	4e-18
C244	1	19113023	nucleotide excision repair protein yeast rad23/ human HHR23A homolog (*Schizosaccharomyces pombe*)	2e-16
C282	1	21950718	Small Heat Shock Protein (*Laccaria bicolor*)	2e-20
C417	2	4099014	Heat Shock Protein (*Coccidioides immitis*)	5e-71
C511	1	6320464	Similar to E. coli ClpB Protein; involved in folding of some mitochondrial proteins;	2e-25

Contig	No of Clones	'Best hit' GI*	Description	BLASTX E value ‡
			Hsp78p (*Saccharomyces cerevisiae*)	
12.02 Detoxification: Σ 11 Contigs				
C009	1	12229734	*PIG16*/ Cytochrome P450 Monooxygenase (*Uromyces fabae*)	0.0
C024	4	6014892	*PIG28*/Peptidyl Prolyl Cis/Trans-Isomerase (*Uromyces fabae*)	5e-84
C096	1	11096026	cAMP dependent protein Kinase catalytic subunit (*Cryptococcus neoformans var. grubii*)	8e-81
C162	1	4138171l	Peroxiredoxin (Malassezia furfur)	4e-36
C238	2	11256324	Superoxide Dismutase (*Hevea brasiliensis*)	6e-28
C364	2	13310484	Sor-like Protein (*Ginkgo biloba*)	2e-70
C384	1	21555230	Transporter-Like Protein/ Tetracycline Resistance Efflux Protein (*Arabidopsis thaliana*)	2e-07
C495	1	19113943	Glutamate--Cysteine Ligase (*Schizosaccharomyces pombe*)	1e-74
C505	1	20150376	Cytochrome-c Peroxidase / Ascorbate Peroxidase (*Saccharomyces cerevisiae*)	6e-10
C534	1	7705704	Glutathione S-Transferase Subunit 1 (*Homo sapiens*)	2e-11
C555	2	18645077	FAD Monooxygenase, Phenol 2-Monooxygenase (uncultured Proteobacterium)	8e-06
13 Viral proteins: Σ 3 Contigs				
C228	1	2015339	189 kDa Replicase [Kyuri Green Mottle Mosaic Virus]	2e-13
C341	2	12018228	Replicase Large Component [Cucumber Fruit Mottle Mosaic Virus]	8e-27
C404	2	9629353	186K Protein [Cucumber Green Mottle Mosaic Virus]	8e-09
14 Similarity to proteins of unknown function: Σ 35 Contigs				
C010	1	39946757	(*Magnaporthe grisea*) 70-15 chromosome	1e-15
C053	2	EAA01885.1	ebiP24 (*Anopheles gambiae* str. PEST)	2e-11
C062	2	19114723	Hypothetical Protein (*Schizosaccharomyces pombe)*	3e-23
C077	1	21296559	Hypothetical Protein B1D1.160 [imported] (*Neurospora crassa*)	5e-30
C080	1	21281373	Hypothetical 178.4 kDa Protein (*Dictyostelium discoideum*)	1e-08
C099	1	T48633	Hypothetical Protein T15N1.90 (*Arabidopsis thaliana*)	1e-14
C120	2	4760440	DyP (peroxidase gene dyp) (*Galactomyces geotrichum*)	7e-05
C150	1	19114725	Hypothetical Protein (*Schizosaccharomyces pombe*)	5e-35

Contig	No of Clones	'Best hit' GI*	Description	BLASTX E value ‡
C172	1	21299610	agCP6112 (*Anopheles gambiae* str. PEST)	7e-11
C196	1	7305487	SH3 (*Mus musculus*)	2e-11
C223	2	8922762	SIPL protein (*Homo sapiens*)	5e-15
C254	1	20976608	Hypothetical Protein (*Dictyostelium discoideum*)	4e-08
C298	1	19115720	Conserved Hypothetical Protein (*Schizosaccharomyces pombe*)	2e-07
C307	1	646206	Contains similarity to gb|D13630 KIAA0005 gene from *Homo sapiens* (*Arabidopsis thaliana*)	1e-06
C313	1	6319587	Protein with weak homology to D. melanogaster serendipity protein and X. laevis basis fibroblast growth factor; Ysa1p (*Saccharomyces cerevisiae*)	2e-11
C332	1	13359203	KIAA1665 protein (*Homo sapiens*)	1e-25
C337	1	19113431	Hypothetical Protein (*Schizosaccharomyces pombe*)	2e-25
C344	2	6319267	Yeast Hypothetical 27.1 kD Protein In Acs1-Gcv3 Intergenic Region (*Saccharomyces cerevisiae*)	4e-17
C346	8	UFU81791_1	*PIG11* similar (*Uromyces fabae*)	1e-08
C355	1	19114169	Vip1 protein (*Schizosaccharomyces pombe*)	1e-20
C357	1	6323761	Hypothetical Protein (*Saccharomyces cerevisiae*)	2e-19
C382	1	7301549	CG5880 Gene Product (*Drosophila melanogaster*)	5e-09
C387	1	11282577	Pumilio-like protein	1e-10
C392	1	124428	Infection Structure-Specific Protein (*Uromyces appendiculatus*)	2e-12
C420	1	21296559	agCP5091 (*Anopheles gambiae* str. PEST)	1e-09
C426	1	21299893	agCP9060 (*Anopheles gambiae* str. PEST)	4e-22
C441	1	21648337	YLR154W-A; Mit1p (*Saccharomyces cerevisiae*)	4e-16
C483	1	21299441	agCP5686 (*Anopheles gambiae* str. PEST)	3e-37
C484	1	15241349	Put. Protein (*Arabidopsis thaliana*)	1e-11
C493	1	14091856	Conserved Hypothetical Protein (*Oryza sativa*)	1e-08
C519	1	15028454	88 kDa Immunoreactive Mannoprotein (*Filobasidiella neoformans*)	1e-06
C549	1	19113122	Weak Similarity to Chick Phosphatidylcholine-Sterol Acetyltransferase (*Schizosaccharomyces pombe*)	1e-18
C550	1	6321790	NC_001140) Hypothetical ORF; Yhr003cp (*Saccharomyces cerevisiae*)	2e-11
C566	1	10433948	Unnamed protein product (*Homo sapiens*)	4e-11
15 Unclassified proteins: Σ 308 Contigs				

* Genebank Identifier Number of the “best hit” at NCBI
‡ BLASTX E value for a random hit

8.2 PSORT II Algorithms

All of the algorithms described below can be found at http://psort.nibb.ac.jp/ (Nakai and Kanehisa 1992).

MTOP: This program is used for the prediction of membrane topology. It predicts whether the N-terminus is located on the cytoplasmic or peripheral site of the membrane. MTOP assumes that the overall topology of eukaryotic membrane proteins is determined by the net charge difference of 15 residues flanking the most N-terminal transmembrane segment on both sides. The applied rule assumes that the difference in the charges of the 15 residues flanking the first internal signal-anchor determines its orientation. The more positive charged part is facing the cytosol. If the protein in question has more than one membrane-spanning region, the orientation of the most N-terminal one determinates the orientation of all subsequent membrane segments (Hartmann et al. 1989).

MITDISC: The method employed by PSORT II to recognize mitochondrial targeting signals is a discriminant analysis whose variables are the amino acid compositions of the N-terminal 20 residues (Nakai and Kanehisa 1992).

Gavel: This program is employed for the prediction of cleavage sites for mitochondrial pre-sequences (Gavel and von Heijne 1990). It must be noted that the authors of PSORT II question its reliability (http://psort.nibb.ac.jp/psort/helpwww2.html).

NUCDISC: This program is used for the discrimination of nuclear localization signals. Many nuclear proteins have their own nuclear localization signal (NLS). Presently, NLSs are classified into three categories: the classical-NLS, the biparatite-NLS , and an N-terminal signal found in yeast protein (Hicks and Raikhel 1995a; Hicks and Raikhel 1995b).

The classical type of NLSs is that of SV40 large T antigen. PSORT II uses the following two rules to detect it: The four-residue pattern (called “pat4”) is composed of four basic amino acids (K or R), or of three basic amino acids (K or R). The other (called ”pat7”) is a pattern starting with P and followed within three residues by a basic segment containing three K/R residues out of four (Hicks and Raikhel 1995a; Hicks and Raikhel 1995b).

Another type of NLS is the bipartite NLS, first found in *Xenopus* nucleoplasmin by Robbins et al. (Robbins et al. 1991 cited in http://psort.nibb.ac.jp/psort/helpwww2.html; Holt et al. 2000) The pattern is: two basic residues, a ten residue spacer, and another basic region consisting of at least three basic residues out of five residues.

Another type of NLS is the N-terminal signal found in yeast protein, Mat alpha2. PSORT II does not use an application to find this signal because this type has not been well studied (http://psort.nibb.ac.jp/psort/helpwww2.html).

The individual scores for the NLS types mentioned above are combined by a discriminant function to generate the “NLS score”. In addition, PSORT II examines the presence of an RNP (ribonucleoprotein) consensus motif (called “RNA-binding motif”), because some RNPs are transported to the nucleus by signals existing in the bound RNAs (Nagai et al. 1995). In the version of NUCDISC used in this thesis, ribosomal proteins are classified as cytoplasmic proteins, although some of them have NLSs and are transported into the nucleus (http://psort.nibb.ac.jp/psort/helpwww2.html).

ER Membrane Retention Signals: PSORT II postulates that the proteins with N-terminal signal sequence will be transported to the cell surface by default unless they have other signals for specific retrieval, retention, or commitment. The retrieval signal for ER luminal proteins from the bulk flow is the consensus motif, KDEL (HDEL in yeast), in the C-terminus. In addition, these proteins should have a cleavable signal sequence in their N-terminus, but the existence of KDEL / HDEL is in practice often sufficient. Although PSORT II only recognizes the KDEL / HDEL pattern, it is known that variations of this motif exists in some organisms and/or cell types (http://psort.nibb.ac.jp/psort/helpwww2.html).

The retrieval signals for ER membrane proteins appear more complex. Two kinds of signals are known; one is the di-lysine motif (the KKXX motif), which exists near the C-terminus of type Ia proteins, and the other is the di-arginine motif (the XXRR motif), which exists near the N-terminus of type II proteins. Both of these motifs exist close to the terminus of the cytoplasmic tail. However, for practical prediction, the existence of these motifs themselves is neither necessary nor sufficient for the localization at the ER membrane. Thus, the reliability of prediction is not high in this stage (http://psort.nibb.ac.jp/psort/helpwww2.htm).

NNCN: In the yeast genome, nuclear proteins are the majority. Since the precise discrimination of NLS is presently difficult, the prediction of nuclear proteins affects much to the total prediction accuracy. PSORT II uses Reinhardt's method for Cytplasmic/Nuclear discrimination. This is a heuristic approach, assuming that nuclear proteins are generally rich in basic residues: If the sum of K and R compositions are higher than 20%, then the protein is considered to have a higher possibility of becoming a nuclear instead of a cytoplasmic protein. Moreover, the score "NNCN", which discriminates the tendency to be in either the nucleus or the cytoplasm, is calculated based on the amino acid composition according to the neural network constructed by Reinhardt (Reinhardt and Hubbard 1998; , A. Reinhardt and T. Hubbard, Nucl. Acids Res. 26, 2230, 1998, http://psort.nibb.ac.jp/psort/helpwww2.htm).

<u>k-Nearest Neighbors Classifier:</u> PSORT II uses the k-nearest neighbor (k-NN) algorithm for assessing the probability of localizing at each candidate site (Horton and Nakai 1997). If for example these k data points contain nuclear proteins with a 50% probability, the protein is predicted to be located in the nucleus with the probability of 50%. Because the data size for each localization site varies greatly, the sites with smaller samples would be hard to be predicted if a large k value is used. Therefore, PSORT II employs an experimental modification: two-fold k-NN. Namely, two different k values (k1 < k2) are used and the localization sites are classified into two categories according to their data size. First, a prediction is performed using the smaller k1 value. If its predicted site belongs to the smaller category, the algorithm terminates; otherwise, the prediction is redone using the larger k2. Currently, k1 and k2 are set to 9 and 23, respectively. Unfortunately, since the two-fold k-NN algorithm used by PSORT II does not seem to be very accurate, it been declared to be a beta version whose performance has not been explored (http://psort.nibb.ac.jp/psort/helpwww2.htm).

8.3 Further Sequence Specific Characteristics

The following table displays the complete results generated by the prediction programs:

Table A-2 PROSITE and PSORT II Sequence Analysis

	PIG5	*PIG7*	*PIG9*	*PIG14*	*PIG15*	*PIG23*
PSORT II						
MTOP	N-terminal side inside	N-terminal side inside	N-terminal side inside	C-terminal side inside MTOP result inconsistent with existence of signal peptide	C-terminal side inside MTOP result inconsistent with existence of signal peptide	N-terminal side inside
NUCDISC	None	Pat4: RRHK (3) at 65 Pat7: PGHRRHK (3) at 62 Bipartite: KRELDQDA NPGHRRHK S at 53 Content of basic residues: 9.1% NLS Score: 0.47	Pat4: RPKK (4) at 161 Pat7: none Bipartite: KRWLAFIK DYFSRLRPK at 147 Content of basic residues: 11.6% NLS Score: 0.27	None	None	None
GAVEL	None	R-2 motif at 40 VRT|SH	R-2 motif at 31 GRP|LI	R-2 motif at 39 SRL|AP	R-2 motif at 37 NRI|PI	R-10 motif at 28 VRS FS

	PIG5	*PIG7*	*PIG9*	*PIG14*	*PIG15*	*PIG23*
ER	None	None	KKXX-like motif in the C-terminus: PSKP	None	None	XXRR-like motif in the N-terminus: RSYL
NNCN:	Nuclear (76.7)	Nuclear (70.6)	Nuclear (89)	Cytoplasmic (76.7)	Cytoplasmic (89)	Nuclear (89)
k-NN:	44.4% extracellular including cell wall 33.3% cytoplasmic 11.1% mitochondrial 11.1% nuclear	34.8%: mitochondrial 30.4 %: nuclear 26.1%: extracellular, including cell wall 4.%: cytoskeletal 4.3%: vacuolar	33.3%: extracellular including cell wall 33.3%: nuclear 11.1%: cytoskeletal 11.1%: vacuolar 11.1%: mitochondrial	77.8%: extracellular, including cell wall 11.1%: mitochondrial 11.1%: vacuolar	44.4%: extracellular, including cell wall 22.2%: plasma membrane 11.1%: cytoplasmic 11.1%: mitochondrial	34.8%: mitochondrial 34.8 %: nuclear 26.1%: extracellular, including cell wall 4.3%: cytoskeletal
PROSITE						
N-end rule Estimated half-life, in vivo	>20 h, yeast, >10 h, *E. coli*	>20 h, yeast, >10 h, *E. coli*	>20 h, yeast, >10 h, *E. coli*	>20 h, yeast, >10 h, *E. coli*	>20 h, yeast, >10 h, *E. coli*	>20 h, yeast, >10 h, *E. coli*
Stability index (Cut off value 40)	36.98 Stable	41.50 Unstable	56.71 Unstable	43.76 Unstable	34.83 Stable	57.37 Unstable