UPTEC X07 026
Examensarbete 20 p
April 2007

Affected protein synthesis in barley upon pathogen attack

Maria Asplund

UPPSALA UNIVERSITET

Molecular Biotechnology Programme

Uppsala University School of Engineering

UPTEC X 07 026	**Date of issue 2007-04**

Author

Maria Asplund

Title (English)

Affected protein synthesis in barley upon pathogen attack

Title (Swedish)

Abstract

Powdery mildew is a wide-spread and economically important plant disease. The aim of this work was to validate a protein synthesis feature seen in infected barley cells. Promoter sequences of candidate genes were isolated with the intention to identify potential regulatory elements and expression profiles for the same genes were determined. Generality of the protein synthesis feature was investigated in *Arabidopsis*. Results showed extensive upregulation of protein synthesis genes probably related to the formation of the fungal feeding organ inside barley cells. Only one promoter sequence was isolated, and no significant conclusions could be drawn. The upregulation of protein synthesis genes was also observed in *Arabidopsis*, which indicate that this might be a general plant response to powdery mildew infection.

Keywords

powdery mildew, plant-pathogen interaction, promoter sequencing, expression profile, quantitative reverse transcription PCR, protein synthesis

Supervisors

Peter Hagedorn

Technical University of Denmark, Risø National Laboratory

Scientific reviewer

Erik Söderbäck

Uppsala universitet

Project name	Sponsors
Language **English**	Security
ISSN 1401-2138	Classification
Supplementary bibliographical information	Pages **47**

Biology Education Centre Biomedical Center Husargatan 3 Uppsala

Box 592 S-75124 Uppsala Tel +46 (0)18 4710000 Fax +46 (0)18 555217

Acknowledgments

I would like to thank the Cell Biology Department at RISØ for including me in their very positive and stimulating atmosphere. It has been a pleasure to work by your side. I would especially like to thank my supervisor Peter Hagedorn, always full of interesting ideas and guiding me into new angles resulting very fruitful for my work. In addition I would like to thank Pernille Olssen, guiding me in the lab and always being very helpful and optimistic. Jakob Skov, for sharing his extensive knowledge about qRT-PCR as well as teaching me how to interpret the results with the biology behind experiments in mind. Michael F. Lyngkjær, for letting me be a part of his group and for sharing his wide knowledge about plants and plant-pathogens. Last but not least I would like to thank "the family", my sister Sofie Asplund and her boyfriend Knud Jønsson for being so kind to let me stay in their home and supporting me in times of joy as well as stress.

Affected protein synthesis in barley upon pathogen attack

Maria Asplund

Populärvetenskaplig sammanfattning

Precis som människor drabbas också växter av sjukdomar. Dessa sjukdomar påverkar ofta människan hårt i form av stora skördeförluster. Ett exempel på en sådan sjukdom som drabbar många växter, däribland korn, vete, ärtväxter, äpple, sockerbeter och druvor, är mjöldagg. Sjukdomen orsakas av svampar, vars sporer lätt sprids med vinden. Svampen lever som en parasit på växten och utnyttjar denna för näringsupptag utan att ta död på sin värd.

Det är önskvärt att försöka förstå samlevnaden mellan svamp och växt för att kunna hindra spridning av mjöldagg och liknande sjukdomar. Många års forskning har lagts ner i ämnet men fortfarande kvarstår många frågor. I detta examensarbete tittar jag närmare på en av dessa frågor; hur genuttrycket i växtcellen påverkas under svampinfektion.

Genuttrycket av ett antal gener har visats uppregleras i specifikt kornceller som infekterats av mjöldagg. Utav dessa har vissa gener undersökts närmare genom att titta på genuttrycks-profilen i tiden. Genom mitt arbete har jag kunnat påvisa att särskilt gener relaterade till proteinsyntes tycks uppregleras och att det ökade uttrycket tycks höra samman med bildandet av svampens näringsupptagsorgan i växtcellen. Jag har därtill lyckats visa att en ökad proteinsyntes tycks vara en generell reaktion som sker även i andra växter som infekterats av mjöldagg. Ökad proteinsyntes hör normalt samman med celltillväxt under goda tillväxtförhållanden. Under rådande omständigheter då växtcellen är under attack av en svamp, och således under stress, är detta inte en väntad reaktion. Min hypotes blir således att uppregulering av proteinsyntes är en reaktion influerad av svampen. Svampen tycks styra växtcellen i den riktning den önskar.

Examensarbete 20 p
Civilingenjörsprogrammet Molekylär Bioteknik
Uppsala universitet
2007

Table of contents

1. Introduction

1.1 Thesis outline

The work described in this report was done at the Technical University of Denmark, Risø National Laboratory in the Biosystems Department between October 2006 and March 2007. It was done as a part of an ongoing project investigating: How protein synthesis is regulated in plant cells after fungal infection. This thesis can be divided into three parts more or less separated in time, experimental technique and aim. I call them; Promoter Sequencing, Expression Profile and Powdery Mildew Infecting *Arabidopsis*. The chapters Aim, Materials and Methods and Results and Discussion have been divided into these three parts. The division was done to make it easier to explain my work in a structured manner and hopefully makes it easier for the reader to follow in the report.

1.2 Powdery mildew – a plant disease

Powdery mildew is very common and widespread plant disease that affects a wide variety of plants species including common crop plants like barley, wheat, pea, apple, sugar beet and grape. It can reduce crop yields by as much as 20-40% and is one of the most economically important groups of plant pathogens [1, 2]. A typical disease symptom is white fluffy superficial fungal growth on the surface of leaves and its fungal spores are spread to other plants by the wind. The powdery mildew fungus is an obligate biotroph which means it requires a living host to survive. In contrast to many other pathogens it does not use 'brute force' by killing host cells in order to access nutrients. Instead a more subtle way of operation is used where the powdery mildew fungus develops a feeding organ, called a haustorium, in close cooperation with the host plant cell (see fig. 1). To be able to do this the powdery mildew possesses mechanisms to escape and suppress the plants defense systems and thereby keeping its host alive [3].

Several species of powdery mildew exist and the species are often subdivided into different *formae speciales* (ff.spp.). A given formae specialis (f.sp.) reflects the ability of a powdery mildew species to infect and reproduce only on a particular host or group of hosts [4]. An example of this is the powdery mildew fungus, *Blumeria graminis* f.sp. *hordei*, hence forth referred to as barley powdery mildew, whose host plant is barley, *Hordeum vulgare*. Only attack by an appropriate f.sp of powdery mildew will cause a high percentage of successful infection and establishment of disease.

The disease cycle of powdery mildew is very characteristic. When a barley powdery mildew spore (conidium) land on a host plant cell it will go through several developmental steps. Within 12-15 h the fungus will try to penetrate the plant cell by the formation of a structure called the appressorium. Two different paths are then possible in the fungus-plant development: (i) the penetration is successful and the fungus will infect or (ii) the penetration is resisted via cell wall fortification by formation of a papilla structure around the penetration site and the fungus will die. In the first case, where the plant cell is infected, a feeding organ (haustorium) will develop within the plant cell for transfer of nutrients, including sugars and

amino acids, to the fungus [5]. The haustorium will be fully developed within 48 h. Hyphae will grow superficially over leaf surface from the infecting spore and from these hyphae new appressoria will develop, around 40 h, with which the fungus will penetrate and infect surrounding cells. This will lead to additional haustorium formation and an increase in fungal uptake of nutrition. Finally, around 5 days after inoculation, new spores will have formed from the newly developed fungal colony (see fig. 2) [4, 6].

However, not all host plants are susceptible to all powdery mildew races. Some plants possess race-specific resistance. This type of resistance reflects the ability of a host plant cell to recognize an intruding pathogen. The resistance is controlled by gene-for-gene interactions in which corresponding genes in the host and parasite determine whether the two organisms, fungus and plant, are compatible. When any of the many genes for resistance (R genes) in the host is matched by a specific, corresponding gene for avirulence (Avr genes), the host and pathogen are incompatible. An incompatible interaction will result in a localized and rapid programmed cell-death at attempted infection sites. Another type of resistance is the non-host resistance, taking place when a plant is attacked by an inappropriate ff. spp. of powdery mildew, e.g. when *Arabidopsis* is attacked by barley powdery mildew or barley is attacked by wheat powdery mildew. The fungus often fails to penetrate the attacked plant cell, or if the penetration does succeed, and a haustorium is formed, the attacked plant cell will die soon afterwards. In both types of resistance the fungus will not be able to form a colony and spread to other plants.

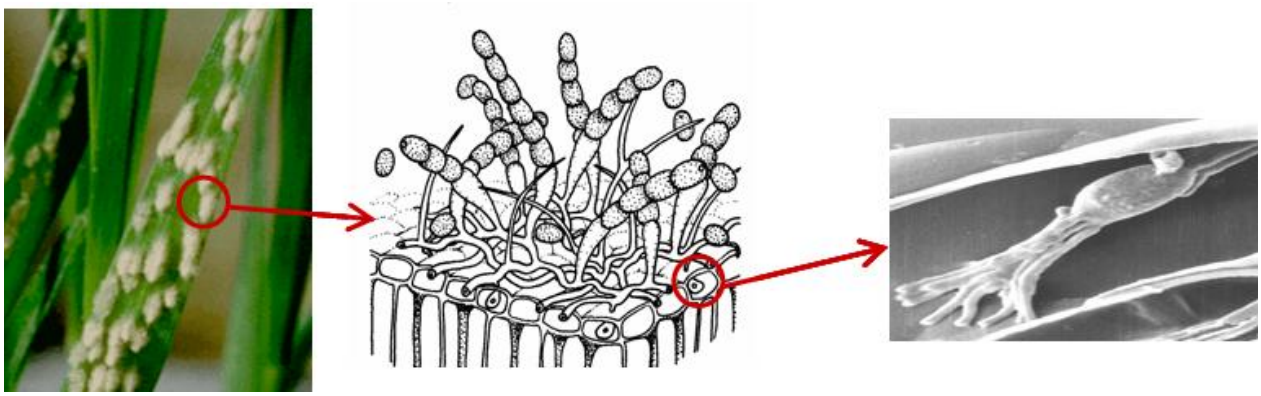


Figure 1. Powdery mildew infecting barley
Left picture is showing barley leafs infected by powdery mildew, followed by an illustration of a fungal colony in the middle and to the right a picture of the fungal feeding organ (haustorium) formed inside the plant cell. Left picture is an approximate enlargement of 1:2. Middle illustration is an approximate enlargement of 1:200. Right picture is a scanning electron microscope picture and the enlargement is approximately 1:1000.

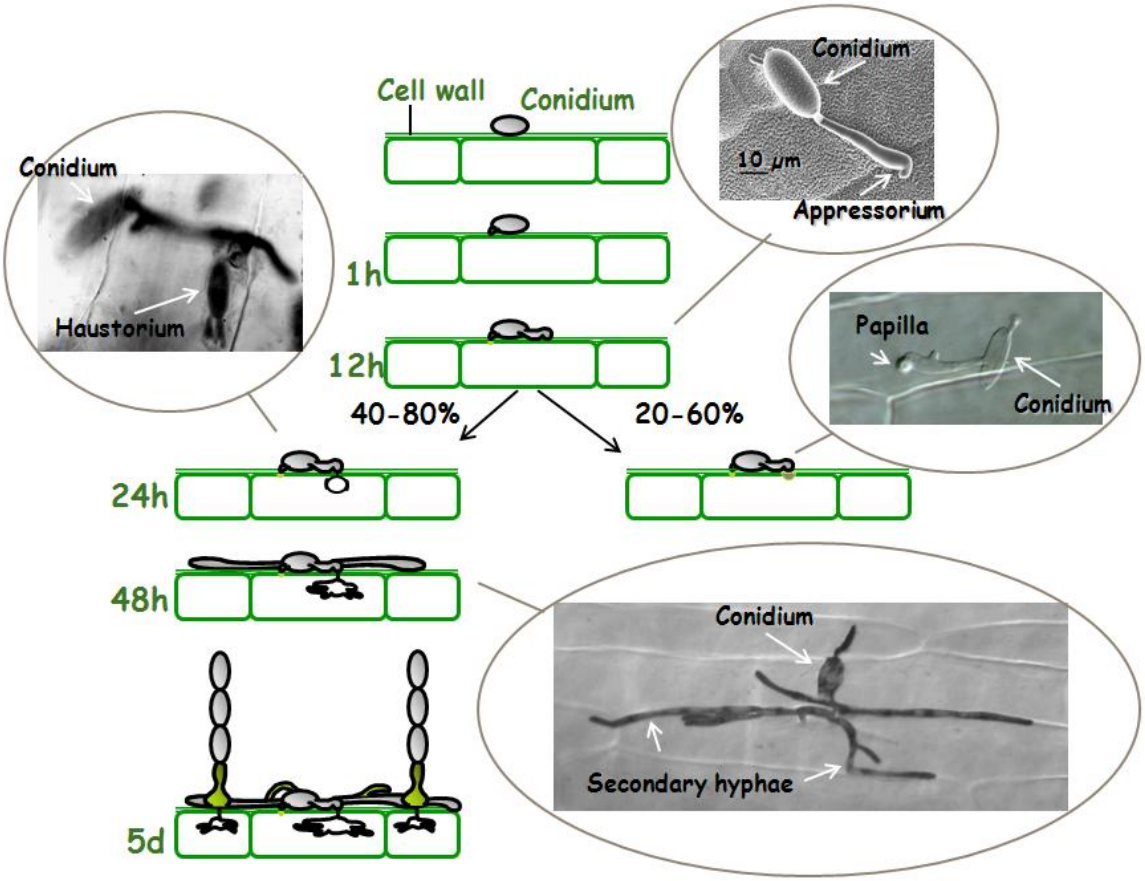


Figure 2. The disease cycle of powdery mildew on barley
When a fungal conidum has landed on a host plant cell two different paths are possible in the plant-fungus interaction. Either the fungal penetration is successful and the fungus will infect the plant cell, or the penetration is resisted by the plant cell by the formation of papilla and the conidium will die. If the fungus succeeds in infecting the plant cell it will start forming the haustorium and if not recognized by the plant cell a fungal colony will finally form. However, if the plant possesses race-specific resistance or if the plant is not a natural host of the invading fungus the infection will be recognized by the plant cell. As a defence response the plant cell will die soon after infection and no colony will be allowed to form. Top right picture is a scanning electron microscope picture. Middle right and bottom right pictures are taken using differential interference contrast (DIC) optics. Top left picture is a bright light microscope picture. In the top left and bottom right picture the fungus has been stained with Evans blue. The conidium, seen in all pictures, is approximately 25 µm long.

2. Background

2.1 Microarray study on single-cell material

When inoculating the barley plant with powdery mildew the leaf as a whole is inoculated. However, the plant cells can be divided into different types depending on what happens in the individual cell. The epidermal cells can be divided as follows: (i) cells where a powdery mildew spore has landed on the cell and the cell has been infected (haustorium formation), (ii) cells where a powdery mildew spore has landed on the cell but the cell resisted the attack (papilla formation) and (iii) cells where no powdery mildew spore has actually landed on the plant cell. The development in these cells would most likely be differentiated and a study was done previous to this work to investigate and compare expression levels between cells. This was done by extracting single-cell material of specific type and analyzing samples in a microarray study. From these experiments genes were found that were upregulated (compared to noninoculated control cells) in infected and resistant cells and also genes that were upregulated in only infected cells and not resistant cells. The hypothesis was formed that genes upregulated in both infected and resistant cells are involved in a stress response to the fungal attack, while genes upregulated in infected but not resistant cells are involved in the actual infection of the plant cell by the powdery mildew, including haustorium formation. The powdery mildew most probably affects the expression profile in barley and the question is how and why?

2.2 Protein synthesis feature

Having found genes upregulated in infected but not resistant cells the first question was: what is the function of these genes? A classification of genes was done where each gene was assigned a class according to what is known about the gene product. Having done this a feature appeared. Many of the genes encoded products involved in protein synthesis. Another microarray study was then done looking at the development over time in powdery mildew infected barley cells. From this study a common expression profile among genes related to protein synthesis was seen. Using clustering analysis other genes with the same expression profile were found and put into the same cluster. Hence forth these genes, upregulated in infected cells and showing a common expression profile over time, will be referred to as the candidate genes, where some of them but not all are putative genes related to protein synthesis.

2.3 Motif in upstream sequence of rice homologous genes

A gene embedded in random DNA is inert. For a gene to be active it needs to be embedded in sequence motifs where proteins capable of directing transcription can bind, otherwise the protein it encodes will never be synthesized. All genes with an influence on phenotype have contiguous regulatory sequences that, together with the expression and activity of proteins encoded elsewhere, regulate when expression occurs, in what cells or tissues, under which conditions and to what extent [7]. One could say that transcriptional regulatory sequences are

as important for a gene's function as its coding sequence.
Genes showing a common expression profile over time indicates that these genes might be co-regulated. However, eukaryotic genes are not organized in co-expressed operons as many prokaryotic genes, but are individually regulated by complex promoters. Hence a common expression profile is most probably coordinated by one or a few trans-acting factors, or transcription factors, regulating by binding to common expression elements or motifs in genes promoter regions [8].

Knowing this, it would be interesting to look for a common motif among the promoter regions of the candidate genes. However, since the whole genome of barley has not been sequenced analyses of promoter regions can not be done using only bioinformatics. As an alternative, ortologous rice genes were identified and their respective upstream regions collected for an in silico analyses. The motif GCGGCGGCG was found in most of the upstream sequences. This highly repetitive and palindromic sequence could be a motif to which a common transcription factor binds and might explain the general expression profile seen among the candidate genes.

2.4 *Arabidopsis* studies

Two distinct studies of particular interest to this work have earlier been done in the same department. Results from these experiments contain data about expression levels in *Arabidopsis* after infection with powdery mildew. In both experiments *Arabidopsis* is the investigated plant and RNA samples were extracted from plants treated in different ways. In Experiment 1 (Exp1) samples are RNA extractions made 18 hours after inoculation (hai) from; (i) *Arabidopsis* inoculated with barley powdery mildew (ii) *Arabidopsis* inoculated with *Arabidopsis* powdery mildew and (iii) noninoculated *Arabidopsis*. Experiment 2 (Exp2) contains another set of data from (i) *Arabidopsis* inoculated with barley powdery mildew and (ii) noninoculated *Arabidopsis*, this time extracted 12 hai. The development when powdery mildew attacks arabidopsis progresses faster than when powdery mildew infects barley. Already at 8 hai the fungus will try and penetrate the plant cell and at 12 hai the haustorium has started to form if penetration was successful (see fig 3). Hence both samples are extracted at a time point when haustorium should have started to form inside plant cell. *Arabidopsis* powdery mildew on *Arabidopsis* is a pathogen host interaction, while barley powdery mildew on *Arabidopsis* is a pathogen non-host interaction. Sample material where *Arabidopsis* has been inoculated with *Arabidopsis* powdery mildew hence includes more infected cells (fungal haustorium formed inside plant cell) than *Arabidopsis* inoculated with barley powdery mildew.

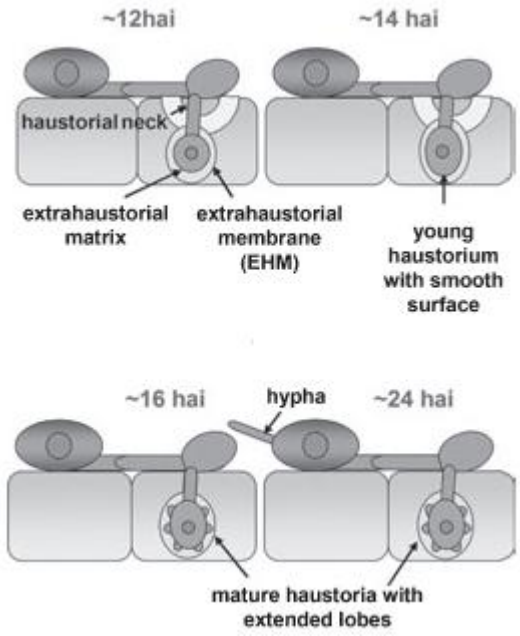


Figure 3. Infection of *arabidopsis*
Part of the infection sequence of *Arabidopsis* powdery mildew on host plant cells. Picture modified after [9].

2.5 PUT-sequences as gene references

The barley genome has not been fully sequenced. However, the PlantGDB (Plant Genome Data Base) contain PUT-sequences representing tentative unique genes of many plants, including barley. PUT stands for PlantGDB-assembled Unique Transcripts and the sequences are a collection of mRNA sequences extracted from NCBI [10, 11]. PUT-sequences were downloaded and used throughout this work as gene references [12].

3. Aim

3.1 Promoter sequencing

As described in the background chapter, the candidate genes display a common expression profile over time. Assuming that this is due to a common transcription factor, influencing the expression of the genes, a common motif in the promoter region of the gene is likely present. In the background studies the putative motif GCGGCGGCG was found in many orthologous genes in rice. However it would probably be more informative to look at the actual promoter sequences of candidate genes in barley. In the Promoter Sequencing part of this work I tried to isolate and sequence upstream regions of the candidate genes in barley. The goal of the isolation was to reach upstream promoter sequences, and this was done using a technique referred to as Walking Into The Unknown (WITU) [13]. The technique amplifies unknown flanking genomic DNA upstream of a known sequence, the known sequence in this case being a PUT sequence.

3.2 Expression profile

In this part of my work I wanted to validate the expression profile over time seen earlier in microarray experiment. I did so using quantitative PCR (qPCR) to make a quantitative analysis of a specific transcript, referred to as expression level, in samples collected over time. RNA was extracted from leaf samples and reverse transcribed into complementary DNA (cDNA). Samples comprising cDNA were then used as template in qPCR reactions amplifying specific genes. A quantitative analysis was made by comparing expression levels with stably-expressed reference genes. The technique is referred to as quantitative reverse transcription PCR (qRT-PCR).

While the microarray experiment was done on single-cell material, I looked at whole-leaf material. Both experiments however investigate barley leafs and development of expression levels over time after infection with powdery mildew.

3.3 Powdery mildew infecting *Arabidopsis*

An interesting question to ask is; can the upregulation of protein synthesis genes after powdery mildew infection be observed also in other plants species, e.g. *Arabidopsis*? In other words, after powdery mildew has infected *Arabidopsis* can I see an upregulation of genes related to protein synthesis? Knowing the answer to this question one could draw a preliminary conclusion about how general the effect of powdery mildew infection is in plants. In this study I did not obtain any experimental data myself, but instead I analyzed data from two experiments done earlier in the same department. The experiments are explained in the background chapter. The aim of this in silico experiment was to investigate how expression levels in *Arabidopsis* is affected by powdery mildew infection, focusing especially on genes related to protein synthesis.

4. Materials and methods

4.1 Promoter sequencing

4.1.1 Theory polymerase chain reaction

The polymerase chain reaction (PCR) is a reoccurring technique in this work, as well as in most molecular biology labs of today, because of its vast possibilities as an analytical tool. It is used to amplify a specific nucleic acid sequence in a cyclic process and generates a large number of identical copies that can readily be analyzed. The basic ingredients of the PCR reaction is; the DNA template to be amplified, a heat-stable DNA polymerase, two oligonucleotide primers complementary to the template and opposite in direction and dNTPs. The reaction is performed by temperature cycling and consists of three main steps: (i) *denaturation*, which separate the double stranded DNA (ii) *annealing* at a temperature generally a few degrees below the melting temperature of the primers, which makes primers base pair with template, and (iii) *elongation* at a temperature optimal for the DNA polymerase which extends the primers by incorporating dNTPs [14].

In this work two different heat-stable polymerases were used; Hot Start DNA Polymerase (Finnzymes) and HotStarTaq® DNA Polymerase (Qiagen). These polymerases need an initial incubation step at high temperature to activate the enzyme before actual PCR starts, which reduces non-specific amplification.

4.1.2 WITU in short

The WITU technique is used to amplify unknown DNA adjacent to a known sequence. Aliquots of genomic DNA are restricted with different restriction enzymes to create short DNA fragments. To each DNA molecule an adaptor is ligated at each end. The adaptors of known sequence contain binding sites for two distinct forward primers. I.e., if one wants to amplify the unknown upstream sequence of a specific barley gene. The putative gene sequence is known through the PUT-sequence, but its location in the genome is unknown. In aliquots with restricted genomic DNA, gene and adjacent upstream sequence will hopefully be present on the same fragment. However the size of adjacent upstream sequence depends on the location of the specific restriction sites in relation to the gene. Two reverse primers are designed from known sequence. PCR is then run in two nested reactions. First PCR is run with forward primer closest to the 5’ end of adaptor and reverse primer most downstream in known sequence. Specific amplified products will contain the binding sites for second forward and reverse primers and second PCR reaction is run with these primers. Note that the forward primers are specific to adaptor present on all fragments. Hence the specificity of the amplification is derived from the reverse primers only. Running two nested PCR reactions increases the specificity. Using the WITU technique it is possible to isolate a specific unknown sequence, as long as a contiguous known sequence is present.

4.1.3 Selecting PUT-sequences

PUT-sequences do not always cover the whole corresponding mRNA. Therefore, if a PUT-sequence starts far downstream of ATG start-site, given that WITU has a limit to its reach, the chance of isolating sequence all the way up to the promoter region is small. Before choosing which genes' promoters to try and isolate, PUT-sequences were blasted using tBLASTx against the non-redundant (nr) database at NCBI (http://www.ncbi.nlm.nih.gov/BLAST/) to find candidate ATG-site (translational start site). Eleven PUT-sequences agreed with the demand that sequence could not start too far downstream of supposed ATG-site and were selected for the experiment.

4.1.4 Primer design

Two reverse primers were designed for each PUT-sequence. Primers were located close downstream, all situated within +110 bp of the candidate ATG-site in the PUT-sequences. In one selected PUT-sequence the ATG start-site was not present and primers were then designed close to the 5'-end of the sequence. Primers were named Gene Specific Primer (GSP) 1 – 11 and indexed 1 and 2 since there were two primers per gene. E.g. the two primers designed for reference gene 1 were called GSP1.1 and GSP1.2. The primer most downstream in PUT-sequence corresponded to GSP1.1. All primers were designed so that no hairpin or primer-dimer formation would disturb the reaction and their melting temperature was 66 +1.5 ºC, to suit the PCR program.

4.1.5 Preparation of template

Genomic DNA from barley had earlier been restricted with 6 different restriction nucleases, which generates DNA-fragments of about 4 kb. These fragments were then ligated to adaptors to which the primers PP1for and dirPP2 will hybridize [13] (see reference for additional explanations). Only one restriction nuclease is added to each sample, meaning six different templates for each gene specific amplification and so the same primer pair will be included in 6 different PCR reactions, where the same promoter is the target. However, depending on where the specific restriction nuclease has cut the genome in relation to the specific gene, the amplification results would vary.

4.1.6 PCR reaction

A 20 µl PCR reaction contained: 200 µM dNTP mix, 0.02 U/µl Phusion Hot Start DNA Polymerase (Finnzymes) and 0.5 µM of each primer in Phusion HF Buffer (Finnzymes). In the first PCR reaction the primers PP1for and GSPi.1 (i =1, 2, ... ,11) were used and 1 µl template (described above) was added. In nested PCR the primers dirPP2 and GSPi.2 were used and 1 µl 1:50 diluted product from the first PCR reaction served as template.

To increase the specificity WITU includes two succeeding PCR reactions. This is necessary since only the reverse primer in the PCR reaction is specific for the gene, forward primer being complementary to adaptor present on all genome fragments. The amplification program of first and nested PCR are the same and comprised: (i) initial denaturation step at 98ºC for 30 s, (ii) first thermocycle repeated ten times with a denaturation step at 98ºC for 5 s and an annealing and elongation step at 72ºC for 3 min, (iii) second thermocycle repeated six times with a denaturation step at 98ºC for 5 s, an annealing step at 68ºC for 20 s and an elongation step at 72ºC for 3 min and (iv) final elongation at 72ºC for 3 min. The program is referred to as a step down PCR and is specified by a high annealing temperature (more than T_m of primers) in the first set of cycles followed by a lower annealing temperature (T_m of primers)

in the second set of cycles. A high annealing temperature in the first set of cycles reduces non-specific primer annealing.

4.1.7 Gel electrophoresis and staining

Product from the nested PCR reaction was separated using gel electrophoresis on 1% agarose gel (w/v agarose solution in TAE buffer). Cresol red was used as loading buffer and PCR products were separated at 50 V. Gels were stained in ethidium bromide bath for 10 min and destained in water bath for 10 min before examined under UV-light.

4.1.8 Purification of fragments and expression in E-coli

Fragments indicating specific amplification were cut out of gel and purified with Geneclean® II Kit (Q•BIOgene) according to protocol. Purified fragments were cloned using the pENTR™ Directional TOPO® Cloning Kit (Invitrogen), with the pENTRTM/D-TOPO® vector and TOP10 chemically competent *E. coli* cells according to protocol. *E. coli* were then grown on selective LB-plates, containing 50 μg/ml kanamycin, for approximately 18 h.

4.1.9 False positive test

For each transformation the 4 biggest colonies were selected. Each colony was tested for inserted PCR fragment by colony PCR. A 15 μl PCR reaction contained: 320 μM of each dNTP, 0.04 U/μl Taq Polymerase (Promega), 2 mM $MgCl_2$, 0.8 μM M13forward and 0.8 μM M13reverse in DNA polymerase Thermophilic buffer (Promega). The same amplification program as above was used and products were separated on 1% agarose gel as above. Each colony was inoculated to a new selective LB-plate. Colonies with verified insert, amplified product in the colony PCR, were selected as true positives. The true positives were inoculated in 5 ml LB kanamycin medium and incubated over night at 37ºC. Cells were then harvested by centrifugation at 8000 rpm for 3 min.

4.1.10 Plasmid DNA-purification and sequencing

Plasmids were purified using the QIAprep® Spin Miniprep Kit (Qiagen) according to protocol. Two 8 μl samples out of each plasmid sample were vacuum dried and sent to sequencing, one for forward and one for reverse sequencing of insert. The sequencing reactions generated up to 1100 bases, but generally around 800 bases were of acceptable sequencing quality. Expecting some inserts of size greater than 1000 kb, both forward and reverse sequencing of insert was necessary to cover its whole length.

4.1.11 Analysis of cloned fragments

Vector sequence (pENTR™/D-TOPO®) was removed from the sequencing results by finding the dirPP2 and or GSPx.2 sequences, removing them and any sequence beyond. All sequences corresponding to the upstream sequence of a specific gene were then aligned to find a consensus sequence, which was stored including the ATG start-site in the 3’ end.

4.2 Expression profile

4.2.1 Theory quantitative PCR (qPCR)

In comparison to ordinary PCR, where only the final product is analyzed, qPCR is a technique where the progress of the PCR reaction is followed in time. To make this possible a fluorescent reporter that binds to the product and reports its presence by a fluorescence signal is necessary. The signal generated reflects the amount of amplified product and the typical short amplified sequences are referred to as amplicons. Signal is measured after each thermocycle and since this is done in real-time the technique is also called real-time PCR.

In the beginning of the PCR process the resulting signal is weak and can not be distinguished from background fluorescence. However, as the amount of amplicon accumulates the signal increases exponentially until it finally levels off and saturates. The levelling off is an effect of reaction running out of some critical component, such as primers, dNTPs or reporter. Noteworthy is that at the end all response curves (fluorescence plotted against thermocycle) have reached more or less the same level, why these measurements tell very little about the original concentration of template in samples, only saying if template was present or not. On the other hand, response curves are separated during the exponential phase of amplification, reflecting the difference in the initial amount of template molecules. A relative quantification between two samples can be done by comparing the number of amplification cycles needed before the samples' response curves reach a specific threshold fluorescence signal value. The cycle when sample reaches this threshold value is called the threshold cycle or C_t. Note that samples with a high concentration of template will reach the threshold fluorescence value faster and hence their C_t-value will be lower than a sample with a low concentration of template which requires more cycles before fluorescence reaches the threshold value [15, 16]. The threshold should be set at a level where response curves are parallel. However, exactly where is somewhat arbitrary since it does not significantly affect the difference between C_t values. The threshold can be selected by the Rotor-Gene software or manually and was selected manually in the work presented here (see fig. 3).

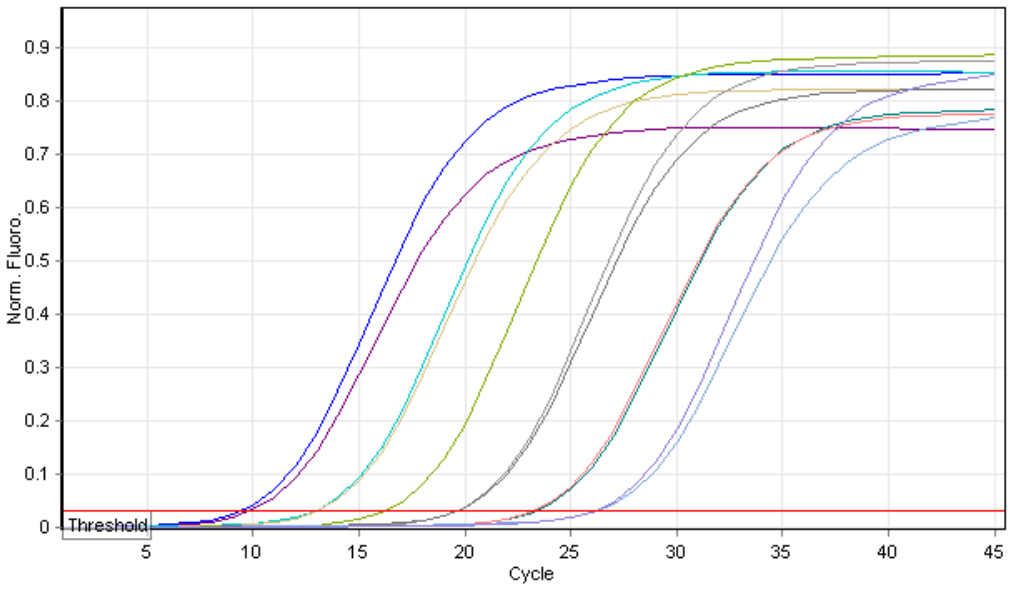


Figure 3. Response curves Fluorescence signal plotted against thermocycle. Graph showing five different samples run in duplicate and threshold set at 0.03.

Several reporters exist for quantifying PCR products, including different fluorescent probes and dyes. The reporter used in this work was the fluorescent dye SYBR®Green I. SYBR®Green I emits virtually no fluorescence when in solution but when it binds to double stranded DNA it becomes greatly fluorescent (see fig. 4). The fluorescence increases with the amount of amplicon formed and even though the relationship is not strictly proportional a certain amount of a particular amplicon always gives rise to the same fluorescence [15]. These qualities make the dye excellent for use in qPCR. SYBR®Green I is a non-specific reporter, i.e. it will incorporate into any double stranded DNA and give rise to fluorescence. For this reason the PCR reaction's specificity is determined entirely by its primers.

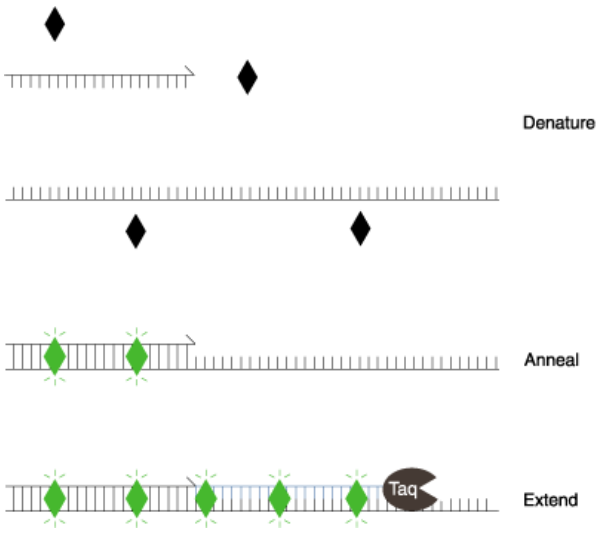


Figure 4. SYBR®Green I
Fluorescent dye used in qPCR becoming greatly fluorescent as it binds to double-stranded DNA.

4.2.2 Running qPCR

All qPCR experiments performed in this work were done in a Rotor-Gene 2000 Real-Time Cycler. On this equipment a PCR run make up maximum 72 PCR reactions which are separated in 100 µl plastic tubes with lids.

A 20 µl PCR reaction contained: 0.5 µM forward primer, 0.5 µM reverse primer in 1x QuantiTect SYBR green PCR Master Mix (Qiagen). The contents of the 2x QuantiTect SYBR green PCR Master Mix being HotStarTaq DNA Polymerase, QuantiTect Green PCR Buffer (Tris·Cl, $(NH_4)_2SO_4$, 5mM $MgCL_2$), dNTP mix and the fluorescent dyes SYBR Green I and ROX (passive reference dye). 3 µl template was added to each PCR-reaction.

The amplification program comprised: initial HotStarTaq DNA Polymerase activation step at 95°C for 15 min and a 45 times repeated thermocycle with a denaturation step at 95°C for 15 sec, an annealing step at 60°C for 30 sec and an elongation step at 72°C for 30 sec.

4.2.3 Melting curve analysis

It is important to verify that the PCR reaction is specific, only amplifying the desired product, since the SYBR®Green I will emit fluorescence for any double-stranded DNA. Such verification can be done by doing a *melting curve analysis* and is done at the end of each PCR run. Temperature is set at a low temperature where also unspecific DNA-binding will occur and is followed by a gradual increase in temperature while fluorescence is measured between every step. Double stranded DNA denaturates (melts) at a characteristic temperature called

the melting temperature (T_m), which depends on size and composition of the DNA molecule [17]. Gradually increasing the temperature, DNA molecules will melt around their specific T_m. As a consequence when amplified PCR product melts, fluorescent dye is released and fluorescence from sample drops significantly. A melting-curve analysis is made by calculating the first negative derivative (-dF/dT) which is plotted versus temperature. This graph will show peaks where specific melting occurs. Specific PCR reactions result in a peak at a characteristic temperature in melting curve analysis that will generally distinguish itself from undesired products (see fig. 5).

Another common problem in PCR reactions are primer-dimer formation. Primer-dimers are formed when primers hybridize to each other because of complementarity, in particular in the 3'end, and is amplified by the DNA polymerase. Primer-dimer formation interferes with the formation of specific products and may give incorrect readouts [15], for that reason it is desirable to detect when primer-dimers are present. As primer-dimers are shorter than the desired amplicon, they will melt at a lower temperature in the melt-curve analysis than amplicon. This makes it possible to determine their presence and avoid being misled by incorrect results.

A melt-curve analysis was done after all qPCR reactions and comprised an initial binding step at 50°C for 30 s thereafter gradually increasing temperature 0.5°C per step until 95°C was reached.

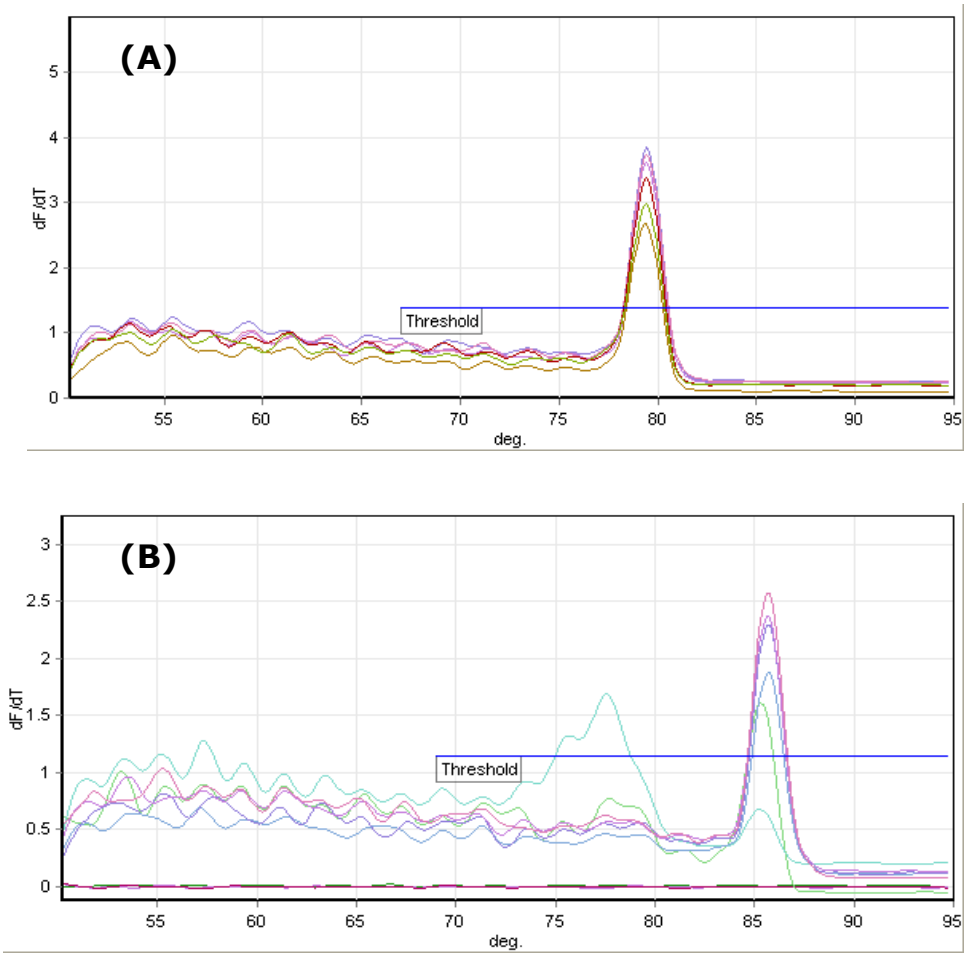


Figure 5. ***Melting curve analysis***
A) Different samples in different PCR reactions with the same specific amplification. **B)** Turquoise line showing primer-dimer formation with consequent decrease in specific product formation.

4.2.4 Determine Expression Levels

To be able to look at expression levels by means of qPCR, samples of RNA must first be reverse transcribed into complementary DNA (cDNA), since the PCR reaction only amplifies DNA. The combined technique with reverse transcription of RNA samples followed by qPCR is called quantitative Reverse Transcription PCR (qRT-PCR).

If one wants to compare the expression level of a specific gene under certain conditions it is necessary to collect samples under these conditions. The conditions in this work were inoculated sample and noninoculated sample at distinct time points after inoculation. These conditions represent the interesting sources of variance in expression levels and are referred to as *interesting variance*. However, samples will also be affected by other sources of variance. These sources, referred to as *obscuring variation*, can be caused by difference in efficiency of RNA extraction and of cDNA synthesis, handling, etc. As a consequence results from qPCR will not necessarily reflect the original concentration of transcripts in samples and comparing concentrations in between samples will not be informative. This is why the obscuring variance needs to be removed. Fortunately, the obscuring variance is sample specific and can hence be removed by the use of *reference genes*. Reference genes are genes which have been shown to be stably-expressed under the experiment's specific conditions, meaning that their RNA levels should be the same in different samples taken under the varying conditions.

What is measured in a qRT-PCR reaction is a specific cDNA concentration in a sample. The cDNA concentration is proportional to the original transcript concentration according to

$$[cDNA_a] = \Delta x \cdot A$$

where *a* denotes the gene of interest (GOI), Δx is the factor of obscuring variance and A is GOI's transcript concentration. The factor of obscuring variance is sample specific and independent of which gene that has been measured. Hence by dividing *a*'s cDNA concentration with the cDNA concentration of a reference gene (*r*) both measured in sample *i* the obscuring variance can be removed

$$[cDNA_a]_i = \Delta x_i \cdot A_i$$

$$[cDNA_r]_i = \Delta x_i \cdot R_i$$

$$\frac{[cDNA_a]_i}{[cDNA_r]_i} = \frac{\Delta x_i \cdot A_i}{\Delta x_i \cdot R_i} = \frac{A_i}{R_i}$$

where the result A_i/R_i is the transcript concentration of *a* relative to the transcript concentration of *r* in a specific sample *i*. This value is referred to as the expression level of *a* and is a relative quantification.

One is usually interested in comparing the expression level of a specific gene between two samples (e.g. two inoculated samples collected at two different time points). The transcription concentration of a reference gene should be the same in all samples *i*. Imagining two different samples: i=1 and i=2; R_1=R_2=R and the expression level of *a* computes

sample1: $\frac{A_1}{R_1} = \frac{A_1}{R}$

sample2: $\frac{A_2}{R_2} = \frac{A_2}{R}$

where the division between the cDNA concentrations of *a* and *r* hence equals the expression level of *a*, in relation to *r*, according to

$$\frac{[cDNA_a]_1}{[cDNA_r]_1} = \frac{A_1}{R_1} = \frac{A_1}{R}$$
$$\frac{[cDNA_a]_2}{[cDNA_r]_2} = \frac{A_2}{R_2} = \frac{A_2}{R}$$

The expression level of *a* is expected to vary between samples and comparisons between samples can be made since R_i=R and dividing A_i with R hence corresponds to a normalization. Comparing the expression level of *a* in two samples computes

$$\frac{A_1}{R} / \frac{A_2}{R} = \frac{[cDNA_a]_1}{[cDNA_r]_1} / \frac{[cDNA_a]_2}{[cDNA_r]_2} = \frac{A_1}{A_2}$$

where

$$[cDNA_r]_1 \neq [cDNA_r]_2$$

A value greater than one would indicate an upregulation in sample1 compared to sample2 and a value less than one would indicate a downregulation.

One can use as many reference genes as one wants and using e.g. three reference genes instead of one gives more reliable data. In these experiments the genes; NDH1, PP2C and UBC2 were used as reference genes. These genes were selected from the result of an investigation conducted to find stably-expressed genes in barley upon pathogen attack (unpublished data). The most stably-expressed genes indicated by the microarray experiment described in background as well as some generally used reference genes were selected for further investigation by qRT-PCR. The three selected genes showed constant expression over time after inoculation with mildew in inoculated samples and noninoculated samples.

Representing the average expression of the three reference genes the geometric mean of their cDNA concentrations is calculated by

$$\overline{[cDNA_r]}_i = \sqrt[3]{[cDNA_{r_1}]_i \cdot [cDNA_{r_2}]_i \cdot [cDNA_{r_3}]_i} = \sqrt[3]{\Delta x_i^{\,3}\ R_1 \cdot R_2 \cdot R_3} = \Delta x_i \cdot \sqrt[3]{R_1 \cdot R_2 \cdot R_3} = \Delta x_i \cdot \overline{R}$$

and is used as the term to remove obscuring variance according to

$$\frac{[cDNA_a]_i}{\overline{[cDNA_r]}_i} = \frac{\Delta x_i \cdot A_i}{\Delta x_i \cdot \overline{R}} = \frac{A_i}{\overline{R}}$$

4.2.5 qRT-PCR preparations

Sample preparation

Barley plants were grown in light chambers with 12 hours of light and 12 hours of darkness in cycles. Approximately 10 days after sowing seeds, 6 hours into a dark period, half of the plants were inoculated with barley powdery mildew. A large number of spores per unit leaf area, 100/mm^2, were used to maximize the percentage of epidermal cells that are challenged by appressorium, giving up to one appressorium per epidermal cell. Plants were harvested at eight different time points: 12, 15, 18, 21, 24, 30, 36, 48 hai. At each time point three leafs from plants inoculated with powdery mildew and three leaves from noninoculated plants, i.e. three biological replicas for each condition, were harvested. The leaves (samples) were placed separately into packages of aluminium foil, dipped in liquid nitrogen (-70°C) and stored at -80°C. The low temperature stops the cell development and inhibits the activity of RNAases ensuring intact RNA. Sampled leaves were used for RNA extraction and following cDNA synthesis. The leaf samples in total included 48 different samples, comprising 16 different conditions with three biological replicas per condition, which were converted into 48 different cDNA samples.

Primer design and primer test

The same 11 PUT-sequences, selected among candidate genes, as in Upstream Sequencing experiment were analyzed in the Expression Profile experiment. Primers were designed using the software Primerselect from Lasergene with PUT-sequences as references. Primers were ordered from MWG and their specificity was tested by qPCR, reaction run as described in 4.2.2. Each primer pair was combined with three different templates (each in a different reaction): (i) cDNA from barley, (ii) genomic DNA from barley and (iii) H_2O (no template control). To verify specific amplification of desired product, end PCR reactions were separated on 2.5% agarose gels, confirming a fragment of the expected size.

Cloning and plasmid purification

There were two main purposes for cloning the amplified product of PCR reaction: (i) to confirm specific and correct product amplification, and (ii) to use cloned and purified plasmids as template when generating standard curves.

PCR reactions which showed specific amplification, i.e. agarose gel containing one band of the expected size, were selected for transformation into *E. coli*. 4 µl of product from qRT-PCR was cloned according to protocol, using the TOPO TA Cloning® Kit (Invitrogen), with the pCR4 TOPO® vector and TOP10 chemically competent *E. coli*. E-coli were then grown on selective LB-plates (50 µg/ml kanamycin) for approximately 18 h at 37ºC. The three biggest colonies were inoculated in 5 ml LB kanamycin medium, incubated over night at 37ºC and then harvested by centrifugation at 8000 rpm for 3 min. Plasmids were purified using the QIAprep® Spin Miniprep Kit (Qiagen) according to protocol.

Validation of insert

The pCR4 TOPO® vector holds two recognition sites for the restriction enzyme *Eco*RI. The sites are situated close to and on both sides of cloning site in vector (see fig. 6) and makes it possible to check for insert by digesting plasmid with *Eco*RI. Digested sample was separated on 2.5% agarose gel together with 100 bp ladder. Size of restricted fragment, was estimated using ladder and confirmed to coincide with size of putative inserted PCR product. 2 µl

plasmid sample was vacuum dried, sent to sequencing and results were compared to expected amplicon from target PUT-sequence. This way both size and sequence agreement between amplified product and expected product was verified.

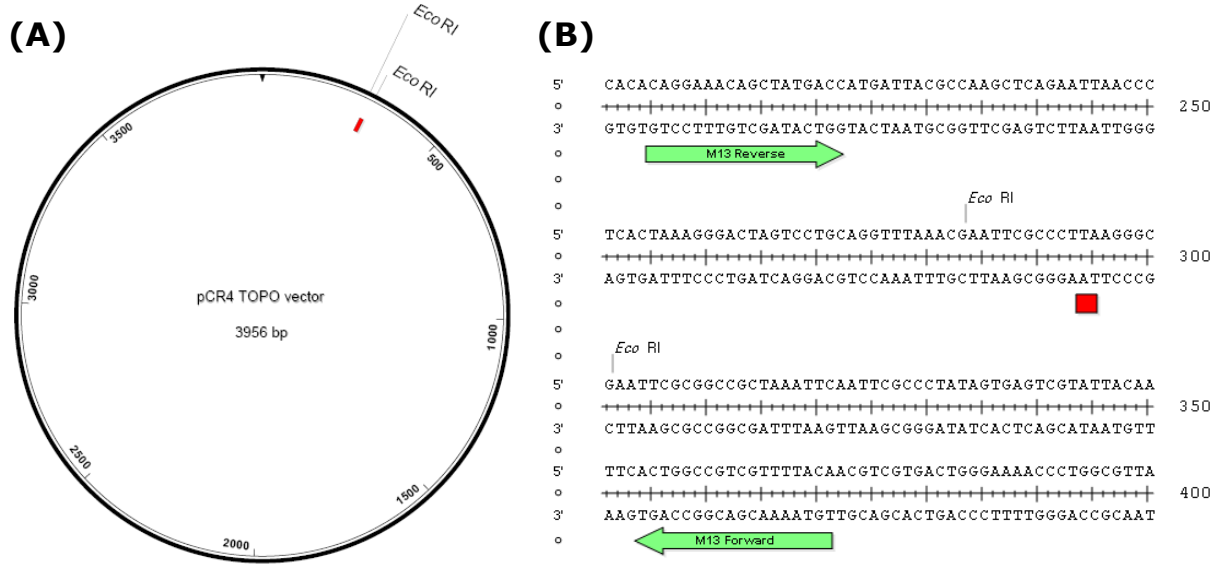


Figure 6. pCR4 TOPO® vector
A) Vector with *Eco*RI restriction sites. **B)** Partial vector base sequence showing TOPO cloning site (red) and M13for and M13rev primer sites for sequencing.

Preparation of template for constructing standard curve

Plasmid samples with verified insert were used as template in qRT-PCR for generating standard curves. However, DNA polymerases works better on linear DNA than circular and hence plasmids were linearized. This was done by digesting either with BstXI or NcoI, which both digest once in vector sequence (see fig 7). An approximate plasmid concentration was determined by running samples on 1% agarose gel together with λ/BstEII marker, with known concentration. The intensity of the band corresponds to amount (mass) of DNA, making it possible to estimate concentration (molecules/µl) of vector in sample. This was done by: (i) comparing unknown sample bands with known marker bands and determine the amount of DNA in band (m_s), (ii) calculating the concentration (g/l) by dividing by added volume (m_s/v), (iii) finding the number of base pares in vector (vector + insert) (n) and (iv) using the formula

$$C_s\,(molecules \cdot \mu l^{-1}) = \frac{\frac{m_s}{v} \cdot N_A}{660(g \cdot mol^{-1} \cdot bp^{-1}) \cdot n}$$

where $N_A = 6.022 \cdot 10^{23}$ molecules per mol (Avogadros constant) and 660 $g \cdot mol^{-1} \cdot bp^{-1}$ is the average weight of one mole of one pair of nucleotides.

Each linearized plasmid sample was then diluted to an approximate concentration of $10^6 \cdot \mu l^{-1}$ and from this a dilution series with concentrations 10^6, 10^5, 10^4, 10^3, 10^2 and 10 (μl^{-1}) was prepared.

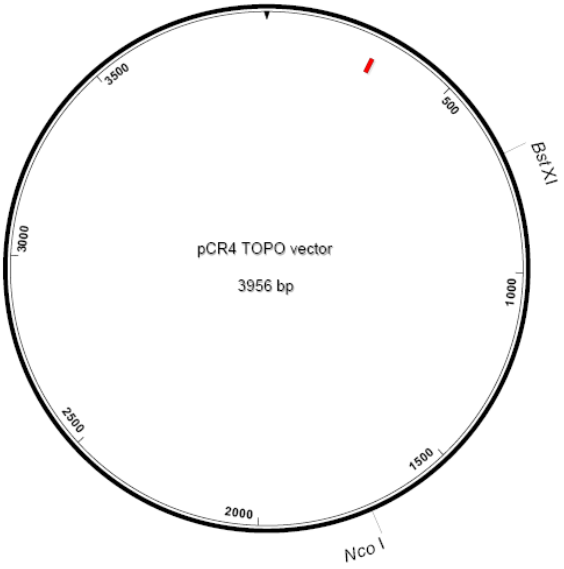


Figure 7. Linearizing pCR4 TOPO® vector
Vector with BstXI och NcoI restriction sites, M13reverse and M13forward priming sites in green and TOPO cloning site in red.

Generating target specific standard curves

The standard curve was based on a serial dilution of a standard, in this case the purified plasmid containing the target sequence. The C_t values of the diluted plasmid were read out and plotted versus the logarithm of the sample's concentration [15]. Linear regression analysis was then used to find the slope (k) [18], which corresponds to the amplification efficiency by

$$E = 10^{-\frac{1}{k}} - 1$$

A PCR-run was made for each dilution series with reactions containing the different concentrations of plasmid, containing a specific amplicon. Primers complementary to specific amplicon were used and plasmid hence served as template. All samples were run in duplicate, and a standard curve was generated using the Rotorgene software (see fig. 8).

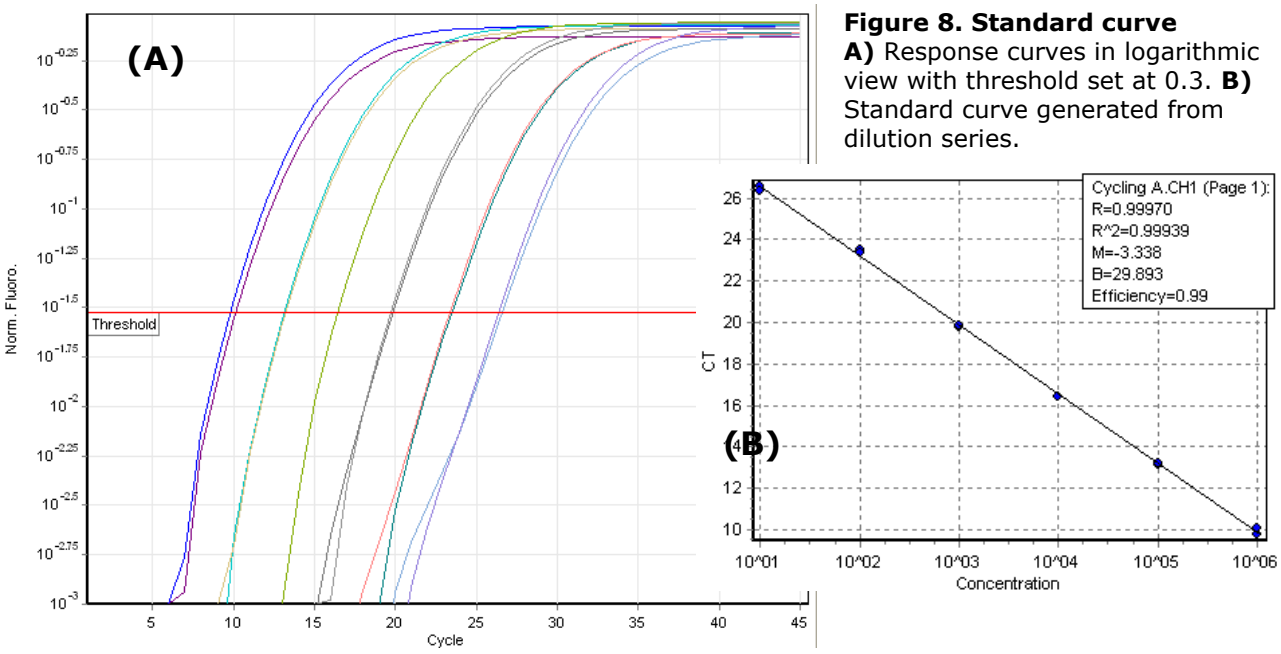


Figure 8. Standard curve
A) Response curves in logarithmic view with threshold set at 0.3. **B)** Standard curve generated from dilution series.

4.2.6 qRT-PCR on time-course samples

The gene specific qPCR makes up 48 PCR reactions (see set up in fig. 9), one per cDNA sample. PCR was run 9 times, 6 times on genes from PSC and 3 times on reference genes. Each PCR-run was set up to amplify one specific PCR product, hence using one gene specific primer pair in all reactions. The concentration of added cDNA template was equivalent to approximately 35 ng/µl RNA.

After each run the generated standard curve, specific to amplified product, was used to calculate concentration of template in respective sample.

		A	B	C	D	E	F	G	H	I
		12hai	15 hai	18hai	21hai	24hai	30hai	36hai	48hai	
1	mildew+	12/+ A	15/+ A	18/+ A	21/+ A	24/+ A	30/+ A	36/+ A	48/+ A	empty
2	mildew+	12/+ B	15/+ B	18/+ B	21/+ B	24/+ B	30/+ B	36/+ B	48/+ B	empty
3	mildew+	12/+ C	15/+ C	18/+ C	21/+ C	24/+ C	30/+ C	36/+ C	48/+ C	empty
4	mildew-	12/- A	15/- A	18/- A	21/- A	24/- A	30/- A	36/- A	48/- A	empty
5	mildew-	12/- B	15/- B	18/- B	21/- B	24/- B	30/- B	36/- B	48/- B	empty
6	mildew-	12/- C	15/- C	18/- C	21/- C	24/- C	30/- C	36/- C	48/- C	empty
7		water	water	empty	empty	empty	Empty	empty	empty	empty
8		empty	empty	empty	empty	empty	Empty	empty	empty	empty

Figure 9. Set up for time-course study
48 PCR reactions containing 48 different cDNA samples, specified in picture: hai/ + (inoculated with mildew) or (-) noninoculated with mildew A, B or C (biological replica).

4.3 Powdery mildew infecting *Arabidopsis*

4.3.1 Data

Data used in this in silico experiment were collected from the results of the two experiments described in section 2.4 of background. The data consisted of hybridization intensities which were obtained by hybridizing RNA samples on *Arabidopsis* Chip ATH1 (Affymetrix). The chip is used for measurement of expression levels in *Arabidopsis*. The oligonucleotides attached to the chip are called probes and for each gene the chip contains 11 probe pairs, which make up a probe set. Several probes for each transcript are needed to maximize sensitivity and specificity of detection. Extracted RNA samples were labelled and one sample was hybridized to each array. Arrays were then scanned and images produced and analyzed to obtain an intensity value for each probe. High intensity corresponds to a lot of binding of sample to probe, which indicate high levels of the specific transcript in the sample. The intensities were used as raw data in this in silico experiment to find significantly regulated genes in inoculated samples compared to noninoculated samples.

4.3.1 Extracting regulation from microarray data

Interesting and obscuring variance

The goal of most microarray experiments is to learn how expression levels of different genes differ in response to genetic or environmental differences. In these two data sets there was an environmental change from noninoculated control plants to inoculated plants attacked by powdery mildew. This source of variation represents the interesting variation. However, as in the qRT-PCR experiment, microarray experiments also result in obscuring variance. In this case the obscuring variance is caused by variation during sample preparation, manufacture of the array and the processing of the array (labelling, hybridization and scanning). The obscuring sources of variation can have many different effects on data. Unless arrays are properly normalized, comparing data from different arrays can lead to misleading results [19].

Software

For making statistical computing on the microarray data the programming language R was used. R is a combination of software facilities for data manipulation, calculation and graphical display. Available for installation in R are the software packages GCRMA and Limma. GCRMA (GeneChip Robust Multi-array Analysis) is used for extracting expression levels from array data, while Limma is used for making differential expression analysis [20].

Procedure

First objective was to combine the 11 probe pair intensities for a given gene to find a measurement of expression that represents the amount of the corresponding mRNA species. This was done by implementing the justGCRMA function from the GCRMA package. The data from Affymetrix experiments contain intensity information about each probe on the chip, extracted from the image data. justGCRMA uses the intensity information from each probe and return a single expression value per gene, and removes obscuring variance. The experiments additionally contain biological replicas, i.e. gene chips have been hybridized to different samples extracted under the same conditions. Expression values from inoculated samples were compared to noninoculated samples, including all biological replicas, to find

the interesting variance or significantly regulated genes. This was done by implementing different functions available in the Limma package. The output was a list containing all probe set IDs and their regulation in inoculated samples compared to noninoculated samples. The regulation is annotated by: 1 = significant upregulation, -1 = significant downregulation and 0 = no significant regulation. The False Discovery Rate (FDR) was set to FDR<0.05, meaning that 5% of genes defined as being significantly regulated are expected to be false positives (not differentially regulated).

4.3.2 Extracting upregulated transcripts

To be able to compare results from this in silico experiment with results obtained from barley experiments, and hopefully make conclusions about common features, the focus was set at upregulated genes in infected material, as has been the focus of barley experiments. Consequently, all genes with regulation equal to 1 were collected and their probe set IDs were assembled in a text file. However, a file containing a lot of probe set IDs does not tell us much. What is interesting is the actual transcript matching the probe set and in the end, the function of the corresponding gene product. Each probe set ID was hence assigned to its corresponding transcript ID.

4.3.3 Assigning functions to transcripts by gene ontology

Upregulated genes in inoculated samples

To assign functions to transcripts the Gene Ontology (GO) database (http://www.geneontology.org) was used. The GO database is the result of a project whose aim has been to facilitate consistent descriptions of gene products in different databases. This has been done by the construction of a controlled vocabulary describing gene products in terms of (i) their associated biological processes, (ii) cellular components and (iii) molecular functions. The building blocks of the GO database are the GO-terms which are the gene descriptions. Each GO-term is associated with a unique numerical identifier, GO-ID. Different gene databases use this database to associate qualities to genes [21].

GOstat (http://gostat.wehi.edu.au/cgi-bin/goStat.pl) is a program available online, which associates different gene identifiers, e.g. transcript IDs, with the corresponding gene's GO-IDs. The list of upregulated genes (List1) was put into GOstat and the program was asked to find the GO-ID of the molecular function associated to the genes. Additionally GOstat also analyses the list of associated GO-IDs, by finding statistically overrepresented GO-IDs. This is done by passing a second list of gene identifiers (List2) to the program. List2 will also be assigned to GO-IDs and is then used as reference. List2 contained all transcript IDs with corresponding probes attached to the *Arabidopsis* Chip ATH1. The program counts the number of appearances of each GO-ID in List1, referred to as the group count, and then in List2, referred to as the total count. Fisher's Exact Test is performed to judge whether the observed difference is significant or not. This will result in a p-value for each GO-ID that the observed counts could have been due to chance. A low p-value signifies that the result is statistically relevant. A list is generated which show how specific the GO-IDs are in the list of upregulated genes [22]. In this experiment only overrepresented (and not underrepresented) GO-IDs were investigated. The top value in the list contains the group of genes with the lowest p-value, i.e. the result most unlikely due to chance.

The GO-IDs were then assigned to their corresponding GO-terms, in this case the quality describing the gene's molecular function. The final output is a list of overrepresented

molecular functions among upregulated genes in inoculated samples compared to noninoculated samples. The 25 most overrepresented GO-IDs were listed in a table showing GO-ID, genes associated to respective GO-ID, group count, total count, p-value and GO-term.

Regulation of homologs to candidate genes

We also wanted to examine the regulation of *Arabidopsis* homologs to the 11 candidate genes in barley. The search for homologs was done using the Find Your Gene resource at PLEXdb (Plant Expression database) (http://www.plexdb.org/modules/PD_dataSelection/blast2expression.php). PLEXdb is a unified public resource for plants' and plant pathogens' expression data with a search engine to map sequences against example sequences included in the database. The PUT-sequences of the 11 barley genes were put in and blasted, using tBLASTx, against the *Arabidopsis* ATH1 gene chip, which has been used in experiments. tBLASTx, is used to find homologous proteins that are not expected to be very conserved at nucleotide sequence level. When implementing tBLASTx both the query sequence and sequences in database are translated into amino acid sequences and aligned at amino acid level. This is convenient for two as distantly related species as *Arabidopsis* and barley. The corresponding regulation of the best hit was then extracted from the two experiments.

5. Results and discussion

5.1 Promoter sequencing

After having implemented the WITU technique on all 11 genes the results were scarce. Out of the 11 PUT-sequences only two (reference gene 3 and 4) indicated specific amplification. The cause of this was hard to determine. No specific quality, e.g. T_m, could be related to the varied results. Additionally when analyzing sequencing results only one out of the two genes, reference gene 3, was really a specific amplification. Results from reference gene 4 could not be aligned into one consensus sequence. Gene 3 resulted in a consensus sequence measuring 830 bp upstream of gene, including ATG-site. This sequence was examined closer. The exact motif found in rice could not be found, but a motif very similar was found. The found motif contained an extra inserted C at position seven and was GCGGCGCGCG (see fig. 10). This indicated that the motif seen in rice might be conserved to some degree. However we would have needed more upstream sequence of genes to really evaluate the presence of motifs.

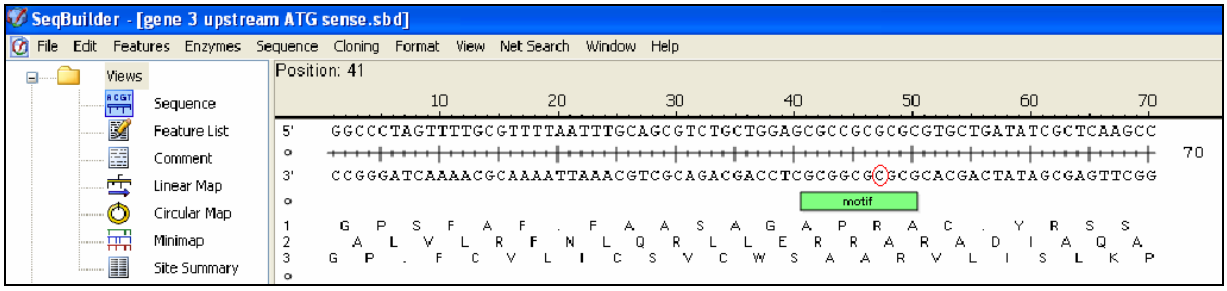


Figure 10. promoter region gene 3
Picture showing part of isolated promoter region from gene 3. A similar motif as the putative motif found in rice (unknown function), was found where the sequence was the same except for an extra inserted C (circled in figure).

5.2 Expression profile

5.2.1 Limited by primers

The intention was to look at the expression profile of 11 genes selected among the candidate genes. However, when trying out the primers designed for the corresponding PUT-sequences only 6 out of the 11 reactions gave amplified product; reference gene 3-8. New primers were designed for the five PUT-sequences not working, but neither the new primers gave any specific amplification. Without further trouble shooting, due to lack of time, experiments continued with six genes.

5.2.2 Expression levels

The expression levels were obtained by taking the relative cDNA concentration of GOI to reference genes (as described in 4.2.4). Hence, the values obtained (referred to as the expression levels) do not have any unit but is a relative quantification compared to reference genes. Values of different samples can be compared since they are relative quantifications to the same reference genes with supposedly the same expression in all samples.

It is important that the data obtained from the experiments give a good general picture of transcription levels in inoculated samples compared to noninoculated samples. To do this it is essential to show that values obtained from inoculated samples are significantly different from noninoculated samples. This was done by collecting three biological replicas (A, B and C) for each condition. The replicas were used to calculate the mean value as well as standard deviation of the expression level for each specific condition. If it was possible to show that two conditions' standard deviations were separated, the conclusion was drawn that they showed significant difference in expression level (see fig. 11). All genes illustrated an upregulation of expression levels in inoculated samples compared to noninoculated samples (see appendix A).

Graphs in figure 12-17 show the expression level versus hai. The values were calculated at eight distinct time points, indicated in graph, and these discrete values were connected by a line to illustrate the pattern of the expression level, or expression profile. This is a common way of illustrating the results since, even though the measurements were done at discrete time points, the regulation of the expression level is in reality a continuous process taking place in cells.

The expression profiles shown in a linear view can be hard to compare given that some genes show a significantly lower expression profile and hence "disappear" in the lower part of the graph (see fig. 12). On the other hand this view is informative since it illustrates in an easily interpreted way how expression levels change over time and the expression level of distinct genes in relation to each other. A feature present among all genes appears to be a significant increase in expression at the 48 hai time point.

In a logarithmic view it was possible to distinguish all the expression profiles (see fig 13). Focusing at the expression profiles of inoculated samples it was possible to observe distinct magnitude of expression levels; gene 5, 7 and 4 showed the highest expression levels (order depending on time point), followed by gene 6 and 3 and last gene 8 with the lowest value of the six. This graph additionally indicates that all genes were upregulated in inoculated

samples (indicated by circles) compared to noninoculated samples (indicated by squares) and that the upregulation occured from first time point. The standard deviation was not included in graph since it disturbed the picture, but can be seen in the genes' individual histograms illustrated in appendix A.

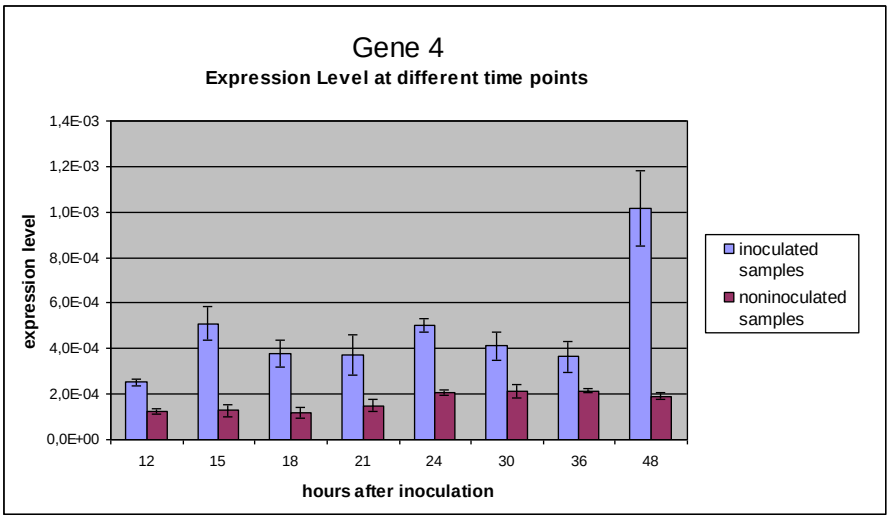


Figure 11. Expression levels gene 4
For each condition three biological replicas have been collected. The biological replicas' expression levels are used to calculated condition's mean value and standard deviation. Graph is showing a significant difference in expression level between inoculated samples and noninoculated samples.

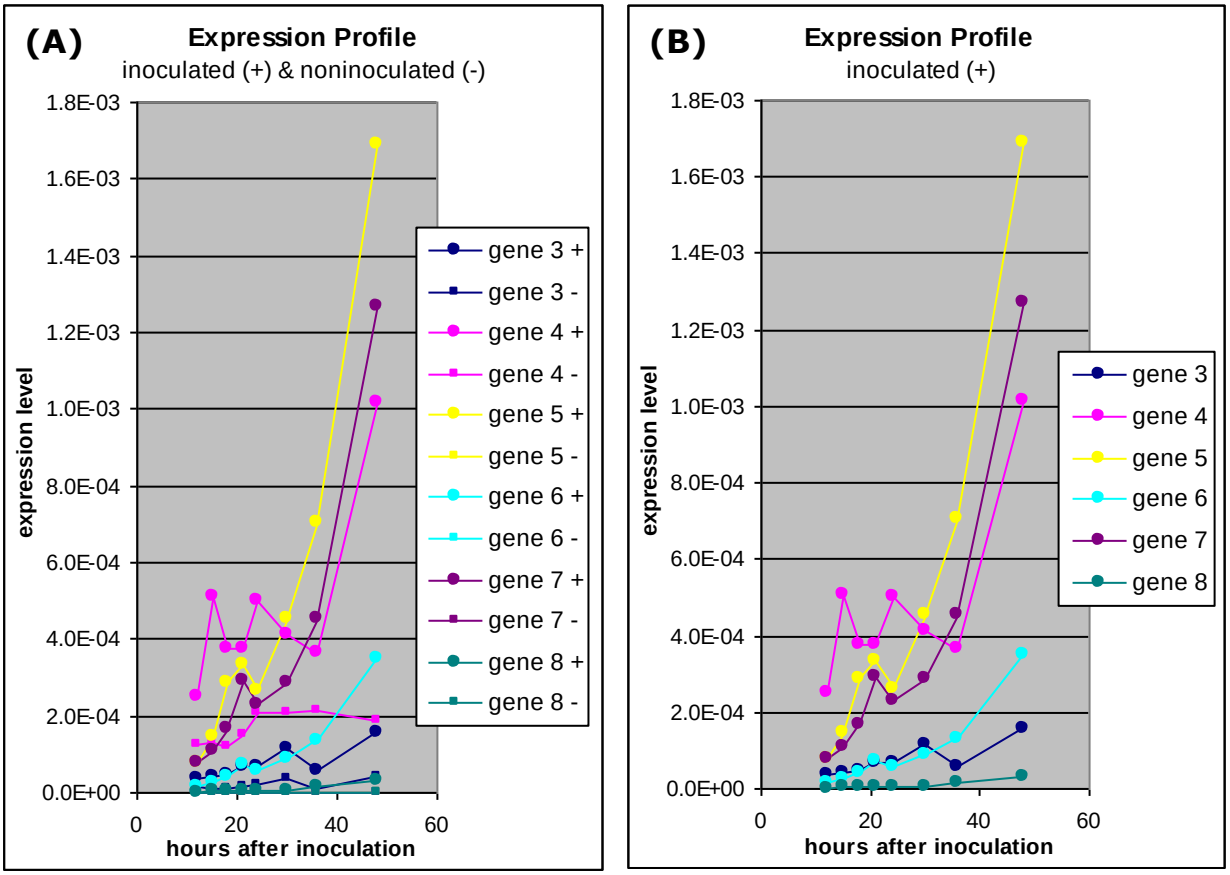


Figure 12. Expression profiles in linear view
A) In linear view it was hard to distinguish the expression profiles of many of the noninoculated samples.
B) Expression profiles of only inoculated samples. A significant increase in expression levels was seen at 48 hai.

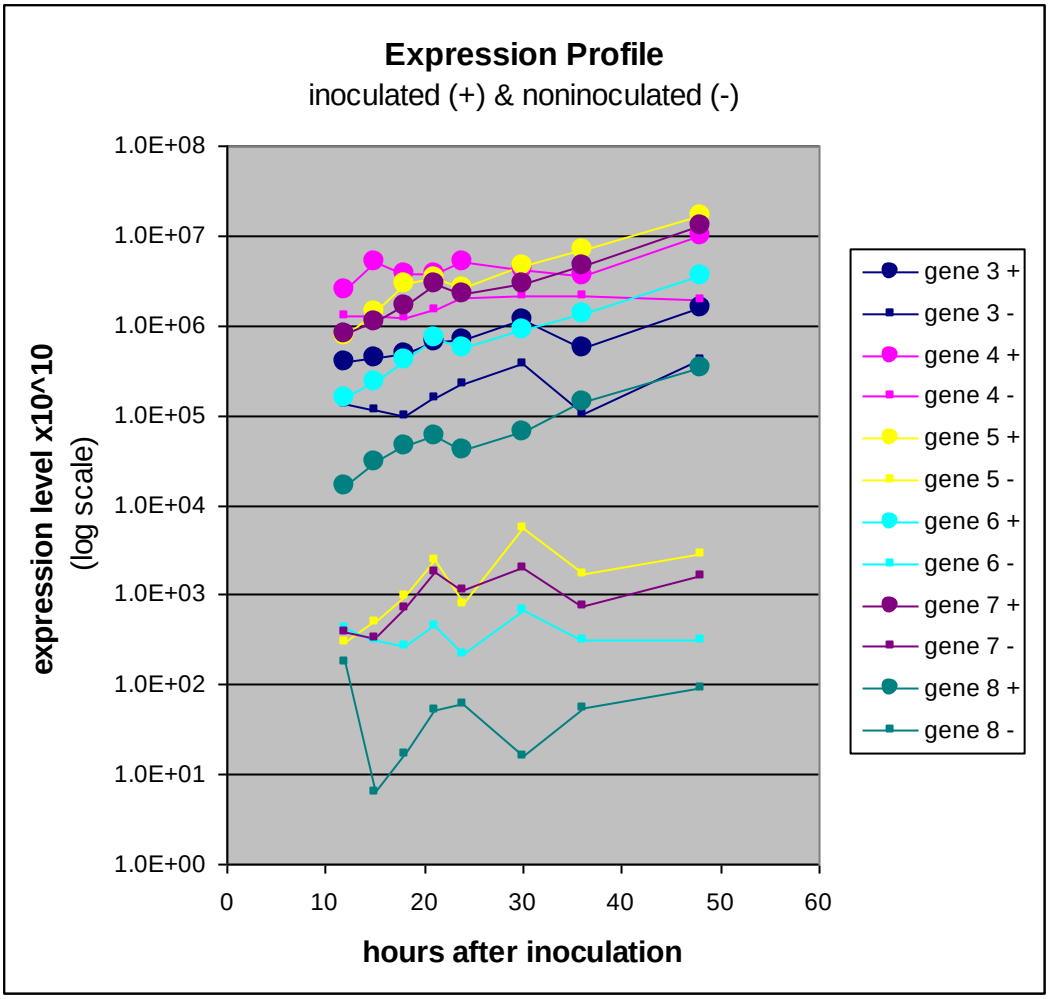


Figure 13. Expression profiles in logarithmic view
In this view all expression profiles were distinguishable. All genes showed significant upregulation in inoculated samples compared to noninoculated samples.

5.2.3 Light influence

Results from gene 3 indicated that its expression level is dependent on light (see fig. 14). The expression level of both inoculated and noninoculated sample rose in the light periods and fell in the dark periods. It was possible to see a clear relationship in the expression level between inoculated sample and noninoculated sample, where inoculated samples and noninoculated samples showed a similar expression profile at different levels.

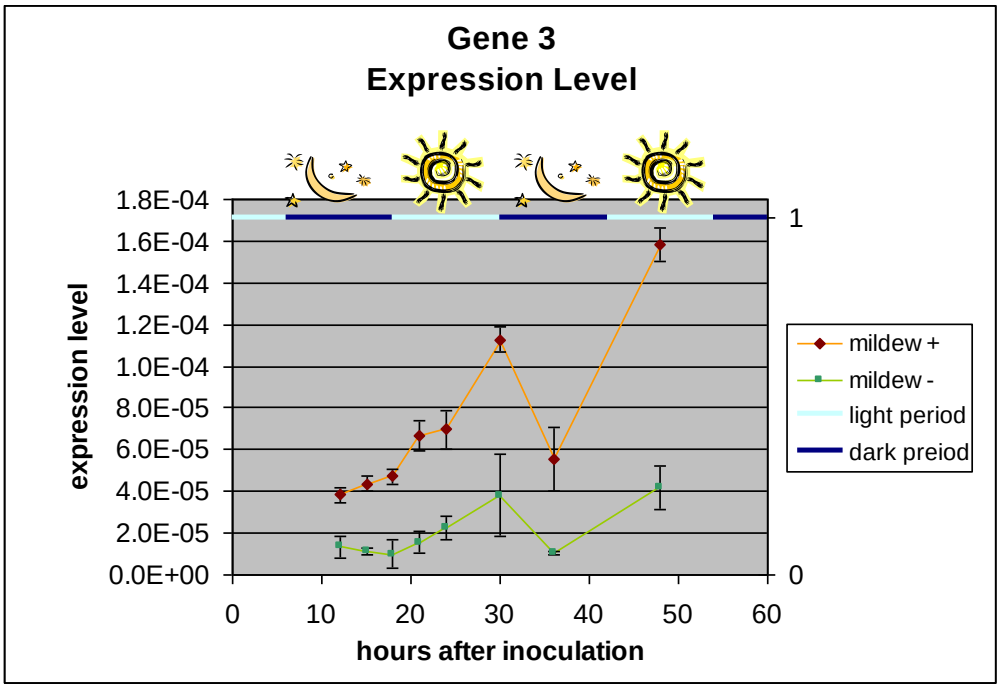


Figure 14. Expression profile of gene 3
Expression level of gene 3 is most likely dependent on light.

5.2.4 Upregulation

When looking at the above graphs (fig. 12-14), showing the expression level versus hai, it was hard to recognize the actual upregulation of genes. However, if the specific genes' expression levels are normalized to its respective expression level at first time point noninoculated, the computed values would indicate the actual upregulation of expression level. This was done and a different picture arose (see fig. 15). Gene 3 and especially gene 4, that showed high expression level in inoculated samples, were actually the genes least upregulated, their expression level being relatively high also in noninoculated samples. On the other hand gene 8 with the lowest expression level was in fact showing strong upregulation; at 48 hai approximately 8300 times (see table 1). Gene 5 showed an extremely high upregulation, about 57000 times at the most. A common pattern could be seen in the expression profiles of gene 5-8. This pattern showed a peak at 21 hai, followed by a dip at 24 hai and then increasing until 48 hai.

Taking a closer look at the actual function of these strongly upregulated genes an interesting feature appeared. Out of the 4 genes which showed the highest upregulation in inoculated samples (gene 5-8), three have been classified as protein synthesis genes and the function of

the fourth is unknown. Additionally, these four genes showed a very similar expression profile. A tight coordination of the expression is typical for genes functioning in a common process and is a wide-spread phenomenon in eukaryotes. The co-expression of related genes is of energetic economical interest to the cell and co-translation may also be necessary for forming a complex [23], as would probably be the case for i.e. components of the ribosomal machinery. It has actually been shown that genes related to protein synthesis show tight correlation between expression profiles [24]. The common expression profile is consistent with the original classification and Gene 8 with unknown function is probably also involved in protein synthesis.

It has been shown in yeast and mammals that the widely investigated protein; target of rapamycin (TOR), plays an important role in adjusting level of protein synthesis to nutrient availability, including regulation of transcription of ribosomal proteins. To conserve resources cells must limit the production of ribosomes, when the need for protein synthesis is reduced, such as occurs when nutrients are limiting [25]. Cells additionally down regulate energy demanding processes such as ribosomal synthesis as a response to stress or low energy, this way arresting growth [26]. Not as much is known about how plants regulate the expression of ribosomal proteins however the current knowledge suggests that TOR is an important component also in a plant pathway regulating cell growth in response to stress [27, 28]. What is intriguing about the results obtained in this study as well as in the background study is that protein synthesis seems to be generally upregulated, indicated by increased expression levels of related genes, an unexpected response in cells under pathogen attack (a stress factor) as well as cells experiencing nutrient depletion, such as occurs when the fungus feed on the plant cell. The plant response is taking an unexpected path indicating that the fungus is somehow governing what is happening in the plant cell.

How would the expression levels be controlled by the fungus? One evaluated hypothesis was that a fungal protein would enter the plant cell and work as a transcription factor. It has been shown that biotrophic parasites can secrete proteins from the haustorium into the plant cell [29]. However the upregulation of transcript levels occur from first time point investigated, 12 hai, where no haustorium has developed why this is an excluded scenario. The upregulation of the protein synthesis must be induced some other way and the question *how* remains.

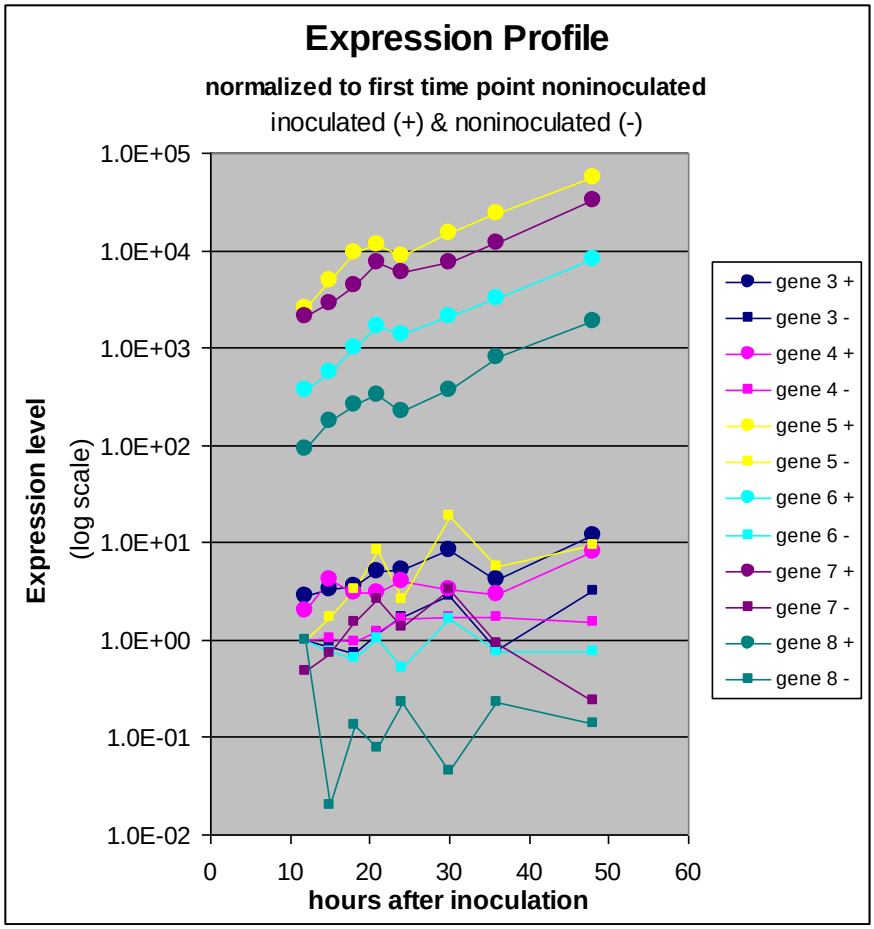


Figure 15. Upregulation: expression level normalized to first time point
Gene 5-8 showed a very high upregulation in inoculated samples compared to noninoculated samples and a similar expression profile. Gene 3 and 4 were also significantly upregulated in inoculated samples but to a much lesser degree than gene 5-8 and their expression profiles are distinct.

Reference gene	blastX against NCBI nr	Protein classification	Upregulation; 48 hai inoculated compared to 12 hai noninoculated	
3	Os03g0297000 [Oryza sativa (japonica cultivar-group)]	unknown	1.2E+01	= 12
4	glutathione transferase [Arabidopsis thaliana]	toxin catabolism	8.1E+00	= 8.1
5	60S ribosomal protein L33-A [Aspergillus terreus NIH2624]	protein synthesis	5.7E+04	= 57000
6	probable ribosomal protein L6.e.B, cytosolic [Neurospora crassa]	protein synthesis	8.3E+03	= 8300
7	translation elongation factor 1a [Hypocrea jecorina]	protein synthesis	3.3E+04	= 33000
8	unnamed protein product [Aspergillus oryzae]	unknown	1.9E+03	= 1900

Table 1. Upregulation of investigated candidate genes
Table showing: Candidate gene, blastX result, classification and approximate upregulation at 48 hai.

5.2.5 Comparing transcription profiles to single-cell study

The transcription profiles found in these qRT-PCR experiments were different from the expression profile found in the microarray-experiments (see fig. 16 and 17) done earlier. However this was not surprising since we were now looking at whole-plant and not single-cell material as in the microarray experiment. Whole-plant material implies that results are illustrating the combined expression profile of different cells, including epidermal cells in direct contact with the powdery mildew spore as well as surrounding cells and mesophyll (internal) cells without direct contact. These cells actually show distinct expression levels if investigated separately, which has been shown in the microarray experiment done earlier.

The expression profile derived from inoculated samples was analyzed in association to what is known about the development after powdery mildew attack on barley. Looking back at the background microarray study, the hypothesis was that the candidate genes are involved in the actual fungal infection (haustorium formation) of the plant cell. In the expression profile derived from that earlier study a peak was seen at 21 hai followed by a decrease in expression level until 48 hai. The peak at 21 hai is coincident with the results derived from this work but on the contrary expression level rose until 48 hai. The microarray study looked at single-cell infected material where a spore has landed on the plant cell. The plant cell will become infected around 24 hai, i.e. haustorium will form, but what happens then? The answer is that fungus will develop and it will start growing hyphe over the plant epidermal cells. This hyphe will then develop new appressorium that try and penetrate surrounding plant cells and new haustorium will form in surrounding cells, starting around 40 hai. However, the impact of the hyphe derived appressorium would not be seen in the microarray single-cell study since it does not include the surrounding cells, but indeed in this whole-leaf study. The results being consistent with this fact the continuous increase in expression level until 48 hai seen in the above results was assigned to the formation of hyphe derived haustorium. These results confirm the hypothesis that candidate genes are involved in the haustorium formation.

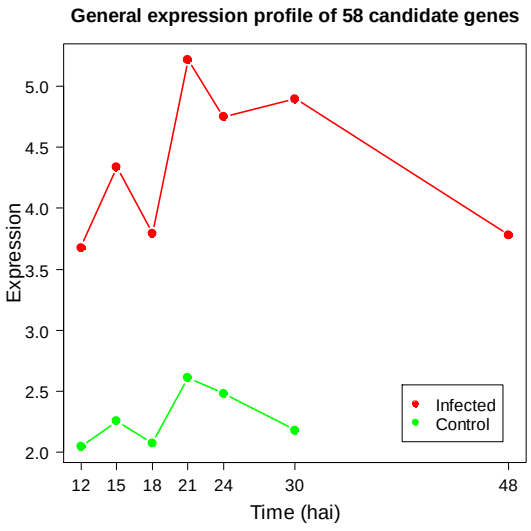


Figure 16. Expression profile from microarray study
The expression profile derived from microarray experiment (background study) done on infected single-cell material.

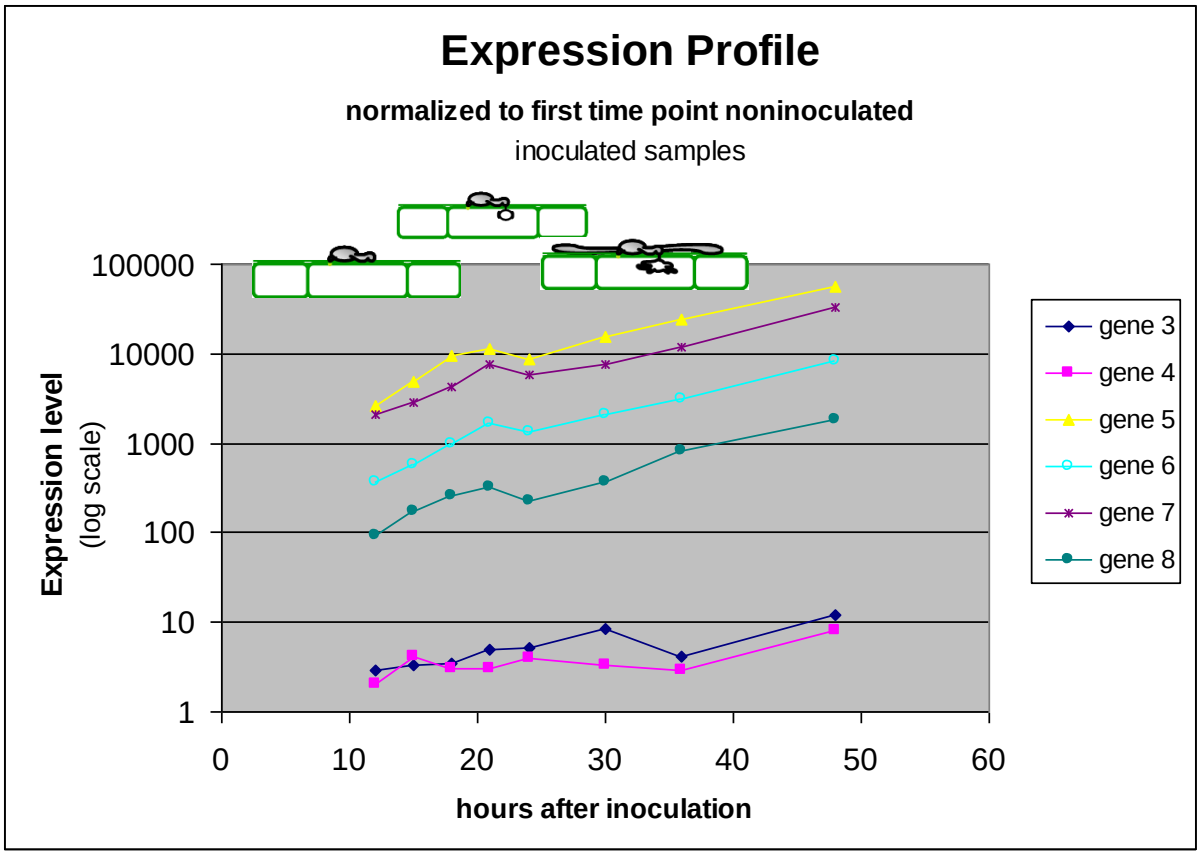


Figure 17. Expression profiles in inoculated samples
Expression profiles derived from qRT-PCR experiments done on inoculated whole-leaf material.

5.3 Powdery mildew infecting *Arabidopsis*

5.3.1 Validation of consistency between Exp1 and Exp2

The data from Exp1 and Exp2 both contain results from *Arabidopsis* inoculated with barley powdery mildew, a pathogen non-host interaction. The results obtained from these data should be similar and could hence be used to evaluate the relevance of this in-silico analysis. To see how many genes that were upregulated in both experiments and how many genes that were specifically upregulated in one of the experiments a Venn diagram was constructed (see fig 18). A Venn diagram shows different sets of data represented by figures, circles in this case, and shared characteristics between data sets are listed where figures are overlapping. This Venn diagram shows how many of the genes from each of the two experiments were upregulated and how many of these upregulated genes are the same in both experiments. Results showed that most genes were upregulated in both experiments though some were also upregulated in only one of the two. That some genes were only upregulated in one of the experiments could be explained by the fact that samples were extracted at two different time points, 12 and 18 hai respectively and even though samples would be collected at the same time point one would probably expect some biological variance.

Figure 18. Venn diagram of upregulated genes Number of genes upregulated in Exp1 (blue) and Exp2 (green), where overlap is showing number of genes upregulated in both experiments.

5.3.2 Protein synthesis feature in *Arabidopsis*

When *Arabidopsis* had been inoculated with barley powder mildew (Exp1 and Exp2) several defence responses were among the list of the 25 most overrepresented GO-terms. On the other hand no GO-terms related to protein synthesis were present in any of the two experiments (see table 2A and 2B). The GO-term *Defence response* appeared as the most overrepresented GO-term in Exp2 (see table 2B).

When *Arabidopsis* had been inoculated with *Arabidopsis* powdery mildew the GO-term *Protein biosynthesis* appeared as the most overrepresented Go-term (see table 2C). Moreover, other functions related to protein synthesis were present in the list while the defence responses were no longer as abundant.

Two features appeared: (i) When barley powdery mildew attacks *Arabidopsis* the response is an upregulation of defence response genes. (ii) When *Arabidopsis* powdery mildew attacks *Arabidopsis* the response is an upregulation of protein synthesis genes. Barley powdery mildew on *Arabidopsis* is a pathogen non-host interaction and, as explained in the introduction, few fungal feeding organs (haustorium) will form within plant cells. Additionally the plant in a non-compatible interaction is better equipped to recognize and stop an intruding pathogen, which would lead to defence responses being turned on. This is consistent with a lot of the defence response genes being upregulated.

On the other hand the *Arabidopsis* powdery mildew on *Arabidopsis* is a pathogen host interaction and will result in abundant haustorium formation within plant cells. The response seen in this compatible interaction was upregulation of protein synthesis related genes, while the defence responses were not as overrepresented as seen in the results from the non-compatible interaction.

The results from this experiment, performed in *Arabidopsis*, hence confirmed the results from the barley experiments saying that a powdery mildew attack, with resulting haustorium formation within plant cell, results in an upregulation of plant genes related to protein synthesis. These are the first results, to my knowledge, indicating that upregulation of genes related to protein synthesis, in a compatible interaction between powdery mildew and plant, is a general feature.

A)

GO-ID	Genes	Groupcount (List 1)	Totalcount (List 2)	Pvalue	GO-term
GO:0009607	ACD1; AT3G50480.1; A	171	411	1,42E-11	response to biotic stimulus
GO:0006952	ACD1; AT1G58220.1; B	141	327	4,04E-11	defense response
GO:0050896	ACD1; AT5G52450.1; A	332	953	7,85E-10	response to stimulus
GO:0009056	ACD1; UBC14; AT5G05	96	220	9,83E-08	catabolism
GO:0006950	ACD1; AT3G50480.1; A	161	420	1,02E-07	response to stress
GO:0044248	ACD1; UBC14; AT5G05	95	219	1,26E-07	cellular catabolism
GO:0044265	UBC14; AT5G05780.1;	75	163	2,02E-07	cellular macromolecule catabolism
GO:0009057	UBC14; AT5G05780.1;	76	167	2,79E-07	macromolecule catabolism
GO:0030163	UBC14; UBC34; AT5G0	56	117	3,13E-06	protein catabolism
GO:0043632	UBC14; UBC34; AT5G0	55	115	3,29E-06	modification-dependent macromolecule catabolism
GO:0019941	UBC14; UBC34; AT5G0	55	115	3,29E-06	modification-dependent protein catabolism
GO:0006511	UBC14; UBC34; AT5G0	55	115	3,29E-06	ubiquitin-dependent protein catabolism
GO:0044257	UBC14; UBC34; AT5G0	55	116	4,09E-06	cellular protein catabolism
GO:0051603	UBC14; UBC34; AT5G0	55	116	4,09E-06	proteolysis during cellular protein catabolism
GO:0006512	UBC14; AT5G05780.1;	67	150	4,09E-06	ubiquitin cycle
GO:0043285	UBC14; UBC34; AT5G0	57	127	3,26E-05	biopolymer catabolism
GO:0006464	AT2G33170.1; BIK1; AT	139	389	0,000106	protein modification
GO:0016192	AT5G50380.1; AT5G504	32	61	0,00011	vesicle-mediated transport
GO:0051707	ACD1; AT3G50480.1; A	74	182	0,000115	response to other organism
GO:0009987	AT3G03110.1; AT2G141	1125	4011	0,000116	cellular process
GO:0006886	AT3G03110.1; AT2G349	41	88	0,000321	intracellular protein transport
GO:0009613	ACD1; AT3G50480.1; P	61	147	0,00035	response to pest, pathogen or parasite
GO:0051234	AT5G50380.1; AT5G413	250	780	0,00035	establishment of localization
GO:0051179	AT5G50380.1; AT5G413	250	781	0,000375	localization
GO:0046903	AT5G50380.1; AT4G315	29	56	0,000384	secretion

B)

GO-ID	Genes	Group count (List 1)	Total count (List 2)	Pvalue	GO-term
GO:0006950	AT3G50480.1; AT2G(	119	420	4,08E-12	response to stress
GO:0050896	AT3G50480.1; AT5G5	212	953	1,70E-08	response to stimulus
GO:0046903	AT5G50380.1; AT5G5	25	56	2,85E-07	secretion
GO:0051234	AT5G50380.1; AT5G4	172	780	1,44E-06	establishment of localization
GO:0051179	AT5G50380.1; AT5G4	172	781	1,44E-06	localization
GO:0045045	AT5G50380.1; AT5G5	22	50	2,34E-06	secretory pathway
GO:0006810	AT5G50380.1; AT5G4	169	772	2,34E-06	transport
GO:0009613	AT3G50480.1; AT2G3	46	147	3,30E-06	response to pest, pathogen or parasite
GO:0050875	AOX1A; AT5G67510.	674	3844	7,06E-06	cellular physiological process
GO:0009628	GSA1; AT2G05710.1	121	524	7,06E-06	response to abiotic stimulus
GO:0009987	AOX1A; AT5G67510.	697	4011	1,53E-05	cellular process
GO:0042829	AT1G74020.1; AT1G5	30	85	1,58E-05	defense response to pathogen
GO:0051707	AT3G50480.1; AT5G5	52	182	1,69E-05	response to other organism
GO:0044271	AT1G35190.1; AT5G5	26	70	2,02E-05	nitrogen compound biosynthesis
GO:0046907	AT1G56590.1; AT5G5	47	161	2,47E-05	intracellular transport
GO:0009607	AT3G50480.1; AT3G4	97	411	2,88E-05	response to biotic stimulus
GO:0006091	AT2G44790; AT3G48	76	304	3,22E-05	generation of precursor metabolites and energy
GO:0042828	AT1G57560.1; GSH2	40	131	3,22E-05	response to pathogen
GO:0051649	AT1G56590.1; AT5G5	47	163	3,22E-05	establishment of cellular localization
GO:0051641	AT1G56590.1; AT5G5	47	164	3,90E-05	cellular localization
GO:0006952	AT3G45290.1; AT2G3	80	327	4,16E-05	defense response
GO:0016192	AT5G50380.1; AT5G5	23	61	4,44E-05	vesicle-mediated transport
GO:0009266	AT1G59860.1; MBF1	33	103	5,86E-05	response to temperature stimulus
GO:0009814	AT1G74020.1; AT1G5	27	78	6,29E-05	defense response to pathogen, incompatible interaction
GO:0000160	ERF11; AT5G07580.1	15	28	6,46E-05	two-component signal transduction system (phosphorelay)

C)

GO-ID	Genes	Group count (List 1)	Total count (List 2)	P-value	GO-term
GO:0006412	AT5G67510.1; AT1G(	166	348	3,08E-22	protein biosynthesis
GO:0009607	ACD1; AT3G50480.1;	176	411	1,51E-16	response to biotic stimulus
GO:0050896	ACD1; AT5G52450.1;	343	953	2,79E-16	response to stimulus
GO:0009059	AT5G67510.1; AT1G(	181	430	2,80E-16	macromolecule biosynthesis
GO:0006952	ACD1; AT1G58220.1;	145	327	1,39E-15	defense response
GO:0044267	AT2G33170.1; AT5G(	377	1115	5,89E-13	cellular protein metabolism
GO:0009987	AT3G03110.1; AT2G1	1148	4011	1,33E-12	cellular process
GO:0019538	AT2G33170.1; AT5G(	383	1144	1,60E-12	protein metabolism
GO:0044249	ATCOAD; AT5G6751	263	731	1,72E-12	cellular biosynthesis
GO:0044260	AT2G33170.1; AT5G(	377	1135	8,64E-12	cellular macromolecule metabolism
GO:0043170	R1; AT2G33170.1; AT	497	1568	1,04E-11	macromolecule metabolism
GO:0009058	ATCOAD; AT5G6751	285	817	1,14E-11	biosynthesis
GO:0050875	AT3G03110.1; AT2G1	1093	3844	3,72E-11	cellular physiological process
GO:0008152	ACD1; AOX1A; AT2G	955	3312	5,01E-11	metabolism
GO:0006950	ACD1; AT3G50480.1;	163	420	7,19E-11	response to stress
GO:0044237	ACD1; AOX1A; AT2G	879	3027	8,88E-11	cellular metabolism
GO:0007028	AT2G20450.1; AT3G1	44	81	2,08E-08	cytoplasm organization and biogenesis
GO:0042254	AT2G20450.1; AT3G1	44	81	2,08E-08	ribosome biogenesis and assembly
GO:0007046	AT2G20450.1; AT3G1	43	80	5,46E-08	ribosome biogenesis
GO:0044248	ACD1; UBC14; AT5G	91	219	8,30E-08	cellular catabolism
GO:0009056	ACD1; UBC14; AT5G	91	220	1,09E-07	catabolism
GO:0044265	UBC14; AT5G05780.1	72	163	1,19E-07	cellular macromolecule catabolism
GO:0009057	UBC14; AT5G05780.1	73	167	1,67E-07	macromolecule catabolism
GO:0044238	AT5G67510.1; AT2G2	771	2727	7,50E-07	primary metabolism
GO:0051707	ACD1; AT3G50480.1;	74	182	8,40E-06	response to other organism

Table 2. Overrepresented GO-IDs
The 25 most overrepresented GO-IDs among upregulated genes are shown. **A)** Results from the pathogen non-host interaction in Exp1. **B)** Results from the pathogen non-host interaction in Exp2. **C)** Results from the pathogen host interaction in Exp1. Many defence related GO-terms (highlighted in green) were overrepresented in (**A**) and (**B**) where the interaction between fungus and plant is incompatible. On the other hand, in (**C**), where the interaction between fungus and plant is compatible, less defence related GO-terms were overrepresented and GO-terms related to protein synthesis (highlighted in yellow) appeared as the most overrepresented as well as in other places in the list.

5.3.3 Regulation of *Arabidopsis* homologs to candidate genes

The 11 PUT-sequences analyzed in the Promoter Sequencing and Expression Profile experiments were blasted against the ATH1 gene chip. Out of these only three produced significant hits. This could be explained by the fact that *Arabidopsis* and barley are two not very closely related plants. Homologous proteins might not exist between the two species. Out of three significant hits two, homologs to gene 2 and 7, were significantly upregulated (see table 3) when *Arabidopsis* has been inoculated with *Arabidopsis* powdery mildew.

Reference Gene	Clonename	PUT-sequence	Target on microarray	Expect value of hit	Transcript ID	Exp1 Regulation *arabidopsis powdery mildew*	Exp1 Regulation *barley powdery mildew*	Exp2 Regulation *barley powdery mildew*
1	HO09N01	PUT-155a-Hordeum_vulgare-25976	254614_at	0,012	At4g19190	1	0	0
2	HO06G08	PUT-155a-Hordeum_vulgare-68197	265805_s_at	3,00E-11	At2g18020	1	1	0
3	HO11M11	PUT-155a-Hordeum_vulgare-209110926	248061_at	0,28	At5g55340	0	0	0
4	HD02A06	PUT-155a-Hordeum_vulgare-671110918	246040_at	0,33	At5g19370	-1	-1	0
5	HO03K13	PUT-155a-Hordeum_vulgare-17979	no hit					
6	HO07B10	PUT-155a-Hordeum_vulgare-48030	263195_at	0,09	At1g36150	0	0	0
7	HO07O11	PUT-155a-Hordeum_vulgare-71020	247644_s_at	7,00E-19	At5g60390	1	1	1
8	HO06F07	PUT-155a-Hordeum_vulgare-61862	258930_at	0,014	At3g10040	0	0	0
9	HO10A15	PUT-155a-Hordeum_vulgare-59930	247654_at	0,046	At5g59850	1	0	1
10	HO12H08	PUT-155a-Hordeum_vulgare-9668	266230_at	0,89	At2g28830	1	0	-1
11	HO13N09	PUT-155a-Hordeum_vulgare-16756	255977_at	2,00E-10	At1g34030	0	0	0

Table 3.
Table showing the 11 genes analysed in the Promoter Sequencing and the Expression Profile experiments. *Target on microarray* indicate the best hit among probes on *Arabidopsis* gene chip ATH1. Three genes resulted in significant hits (highlighted in yellow) and their regulation in inoculated samples compared to noninoculated samples is shown in the last three columns.

6. Conclusions

Attack by powdery mildew results in distinct phenotype in plant cells, i.e. papilla formation or haustorium formation, and the difference is also evident at transcript level, as was shown in the microarray experiment. We have been able to show that all candidate genes investigated by qRT-PCR are significantly upregulated as a consequence of powdery mildew infection. Genes involved in protein synthesis show a particularly extensive increase in transcription and the upregulation of the protein synthesis genes seems to be related to haustorium formation inside plant cells. Additionally we have shown that the protein synthesis feature is also present in *Arabidopsis* infected by powdery mildew indicating that this is a general feature among plants being infected by powdery mildew. Upregulation of protein synthesis is not an expected response in plant cells under pathogen attack and experiencing nutrient depletion, which indicates that the fungus is somehow affecting the expression levels of plant genes. When having infected the plant cell the powdery mildew seems to govern the plant cell, that way using it for its own benefit.

7. Future perspectives

Usually as you investigate one thing, new interesting aspects appear. Many interesting questions have evolved during my work and even though I have intended to answer as many as possible several remain. Here I have stated a few questions that would be interesting to follow up on.

What happens beyond 48 hai?

If the protein synthesis feature is haustorium related one would expect a decrease in transcript levels of these genes beyond 48 hai, as no more haustorium is formed.

Is there a motif present in *Arabidopsis*?

In-silico analysis in *Arabidopsis* resulted in isolation of many protein synthesis related genes upregulated in infected material. The upstream sequences of these could easily be extracted, e.g. from TAIR (http://www.*Arabidopsis*.org), since the whole genome of *Arabidopsis* has been sequenced.

Is their a motif, or cis-regulatory element, present in the upstream region of candidate genes?

We were not really able to answer this question and intents to isolate candidate genes' upstream regions continue.

Does a transcription factor bind to motif in upstream region of candidate genes?

If a motif is found it would be interesting to do a binding analysis to look for proteins binding to motif, which could be transcription factors.

What is the function of the upregulated protein synthesis?

It would be interesting to investigate what the function of the upregulated protein synthesis could be. One hypothesis is that the increased protein synthesis is necessary for the haustorium formation.

8. References

1. Braun, U. (1987) A Monograph of the Erysiphales (Powdery Mildews)

2. Agrios, G. (1988) Plant Pathology. 337-339

3. Panstruga, R. (2002) Live and let live: insights into powdery mildew disease and resistance. Molecular Plant Pathology 3, 495–502

4. Bélanger, R. (2002) The Powdery Mildews, A Comprehensive Treatise

5. Szabo, L. (2001) Hidden robbers: The role of fungal haustoria in parasitism of plants. PNAS 14, 7654-7655

6. Hückelhoven, R. (2005) Powdery mildew susceptibility and biotrophic infection strategies. FEMS Microbiology Letters 245, 9-17

7. Gregory, A. (2003), The Evolution of Transcriptional Regulation in Eukaryotes. Mol. Biol. Evol. 20. 1377-1419

8. Niehrs, C. (1999) Synexpression groups in eukaryotes. Nature 402, 483–487

9. Koh, S. (2005) *Arabidopsis* thaliana subcellular responses to compatible Erysiphe cichoracearum infections. The Plant Journal 44, 516-529

10. Dong, Q. (2006) Comparative Plant Genomics Resources at PlantGDB. Plant Physiology 139, 610-618;

11. "PlantGDB-assembled Unique Transcripts (PUTs) - Motivation and Procedure". Jan. 2006. http://www.plantgdb.org/prj/ESTCluster/PUT_procedure.php (25 mars 2007)

12. version 155a PlantGDB, Plant Genome Database: http://www.plantgdb.org/download/download.php?dir=/Sequence/ESTcontig/Hordeum_vulgare (6 Oct. 2006)

13. Zhang, Z. (2000) Walking into the unknown: a 'step down' PCR-based technique leading to the direct sequence analysis of flanking genomic DNA. Gene 253, 145-150

14. Brown, T. A. (1999) Genomes

15. Kubista, M. (2006) The real time polymerase chain reaction. Molecular Aspects of Medicine 27, 95-125

16. Bustin, S. (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. Journal of Molecular Endocrinology 25, 169-193

17. Nolan, T. (2006) Quantification of mRNA using qRT-PCR. Nature Protocols 1, 1559-1582

18. Rutledge, R. G. (2003) Mathematics of quantitative kinetic PCR and the application of standard curves. Nucleic Acids Res. 31, e93

19. Irizarry, R. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4, 249-264

20. Smyth, G. (2004). Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Statistical Applications in Genetics and Molecular Biology 3, No. 1, Article 3

21. "An Introduction to the Gene Ontology." http://www.geneontology.org/GO.doc.shtml (25 March 2007)

22. Beißbarth, T. "Find statistically overrepresented GO terms within a group of genes." http://gostat.wehi.edu.au/cgi-bin/goStat.pl (25 March 2006)

23. Niehrs, C. (1999) Synexpression groups in eukaryotes. Nature 402, 483-487

24. Eisen, M. (1998) Cluster analysis and display of genome-wide expression patterns. PNAS 95, 14863-14868

25. Mayer, C. (2006) Ribosome biogenesis and cell growth: mTOR coordinates transcription by three classes of nuclear RNA polymerases. Oncogene 25, 6384-6391

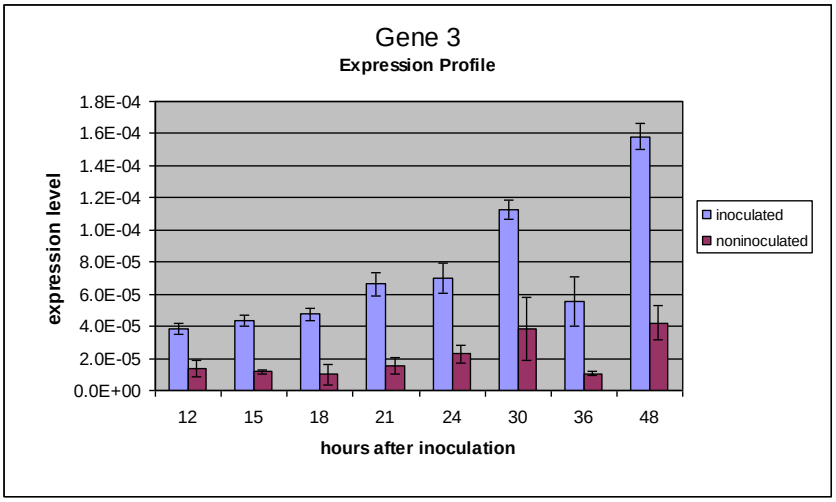
26. Wullschleger, S. (2006) TOR signaling in Growth and Metabolism. Cell 124, 471-484

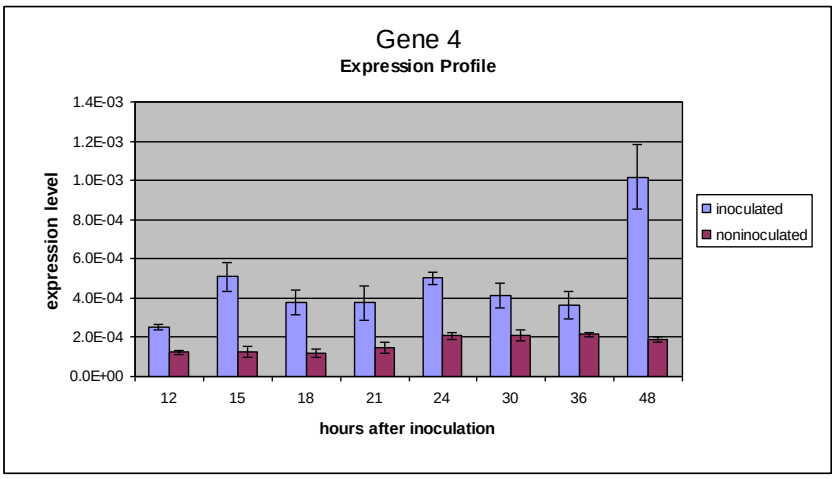
27. Kerri, B. (2006) Ribosomal protein gene regulation: what about plants?. Canadian Journal of Botany 84, 342-362

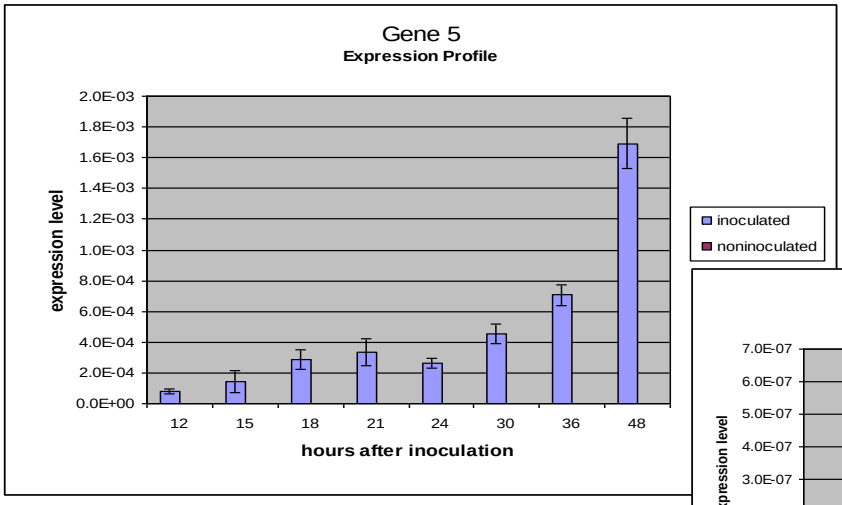
28. Mahfouz, M. (2006) *Arabidopsis* TARGET OF RAPAMYCIN Interacts with RAPTOR, Which regulates the activity of S6 Kinase in Response to Osmotic Stress Signals. The Plant Cell 18, 477-490

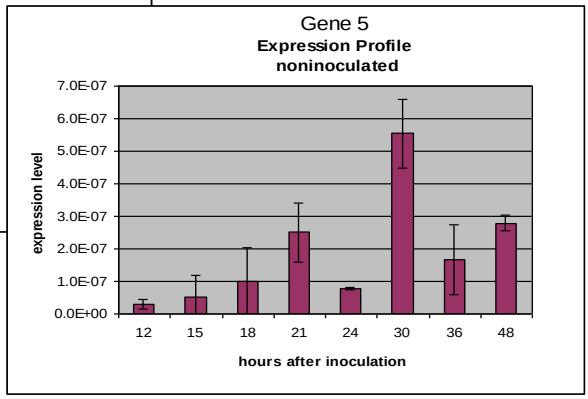
29. Kemen, E. (2005) Identification of a Protein from Rust Fungi Transferred from Hautoria into Infected Plant Cells. Molecular Plant-Microbe Interactions 18, 1130-1139

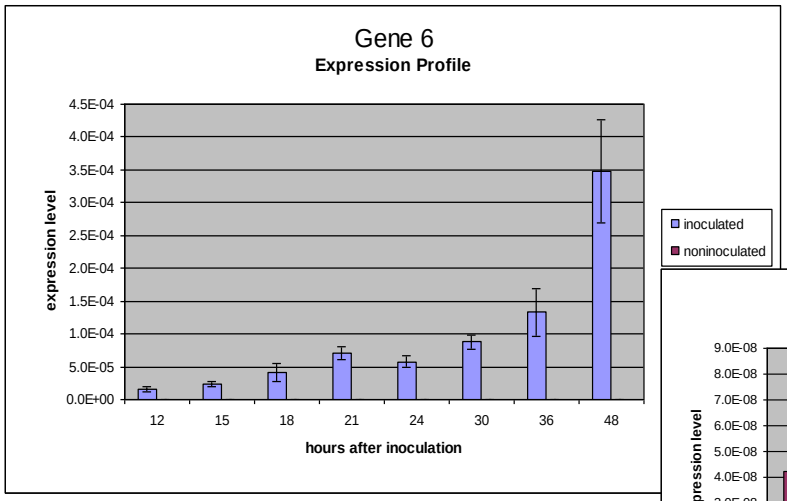
9. Appendix

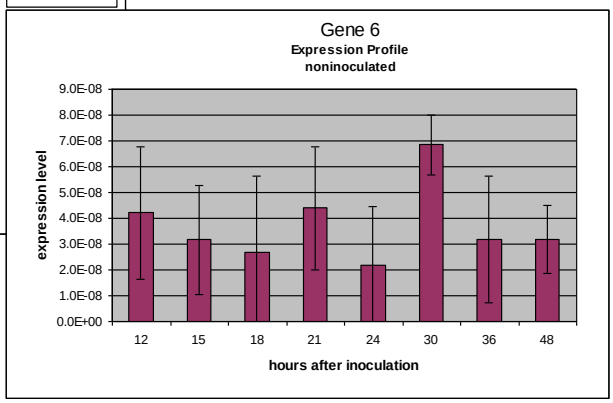

Gene 3
Expression Profile
1.8E-04
1.6E-04
1.4E-04
1.2E-04
1.0E-04
8.0E-05
6.0E-05
4.0E-05
2.0E-05
0.0E+00
expression level
12
15
18
21
24
30
36
48
hours after inoculation
inoculated
noninoculated

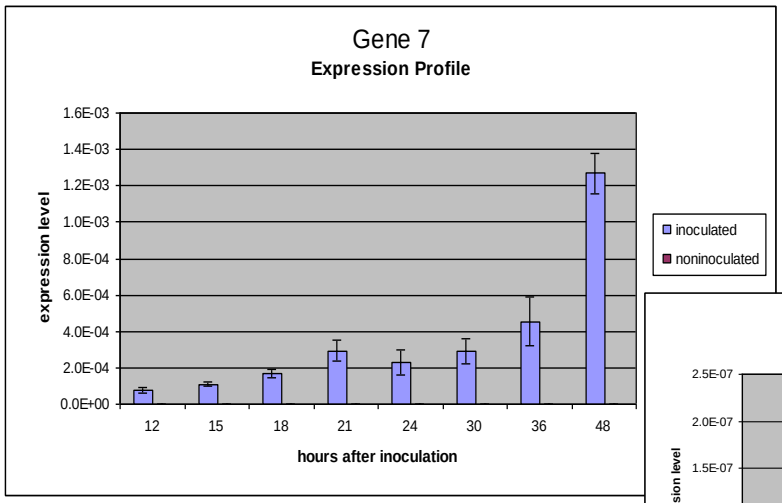

Gene 4
Expression Profile
1.4E-03
1.2E-03
1.0E-03
8.0E-04
6.0E-04
4.0E-04
2.0E-04
0.0E+00
expression level
12
15
18
21
24
30
36
48
hours after inoculation
inoculated
noninoculated

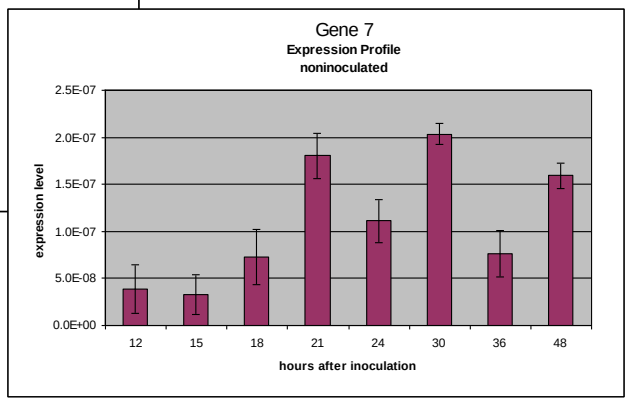

Gene 5
Expression Profile
2.0E-03
1.8E-03
1.6E-03
1.4E-03
1.2E-03
1.0E-03
8.0E-04
6.0E-04
4.0E-04
2.0E-04
0.0E+00
expression level
12
15
18
21
24
30
36
48
hours after inoculation
inoculated
noninoculated

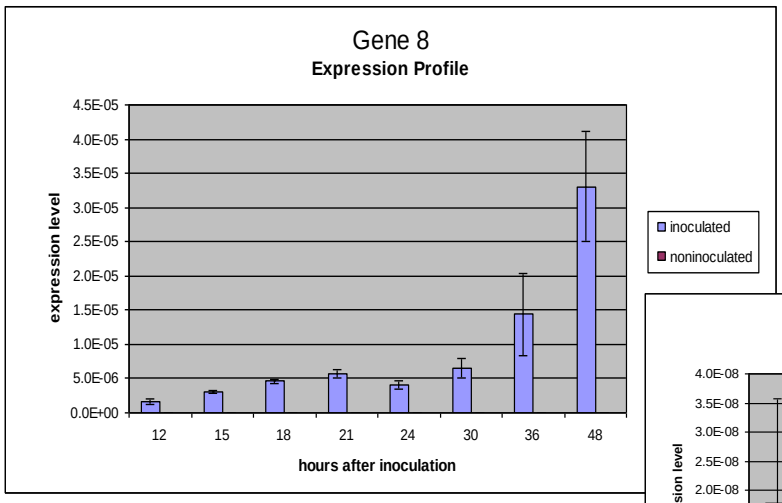

Gene 5
Expression Profile
noninoculated
7.0E-07
6.0E-07
5.0E-07
4.0E-07
3.0E-07
2.0E-07
1.0E-07
0.0E+00
expression level
12
15
18
21
24
30
36
48
hours after inoculation

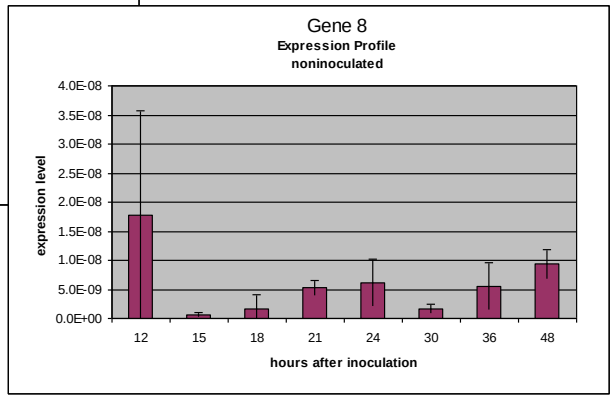

Gene 6
Expression Profile
4.5E-04
4.0E-04
3.5E-04
3.0E-04
2.5E-04
2.0E-04
1.5E-04
1.0E-04
5.0E-05
0.0E+00
expression level
12 15 18 21 24 30 36 48
hours after inoculation
inoculated
noninoculated


Gene 6
Expression Profile
noninoculated
9.0E-08
8.0E-08
7.0E-08
6.0E-08
5.0E-08
4.0E-08
3.0E-08
2.0E-08
1.0E-08
0.0E+00
expression level
12 15 18 21 24 30 36 48
hours after inoculation


Gene 7
Expression Profile
1.6E-03
1.4E-03
1.2E-03
1.0E-03
8.0E-04
6.0E-04
4.0E-04
2.0E-04
0.0E+00
expression level
12 15 18 21 24 30 36 48
hours after inoculation
inoculated
noninoculated


Gene 7
Expression Profile
noninoculated
2.5E-07
2.0E-07
1.5E-07
1.0E-07
5.0E-08
0.0E+00
expression level
12 15 18 21 24 30 36 48
hours after inoculation


Gene 8
Expression Profile
4.5E-05
4.0E-05
3.5E-05
3.0E-05
2.5E-05
2.0E-05
1.5E-05
1.0E-05
5.0E-06
0.0E+00
expression level
12 15 18 21 24 30 36 48
hours after inoculation
inoculated
noninoculated


Gene 8
Expression Profile
noninoculated
4.0E-08
3.5E-08
3.0E-08
2.5E-08
2.0E-08
1.5E-08
1.0E-08
5.0E-09
0.0E+00
expression level
12 15 18 21 24 30 36 48
hours after inoculation