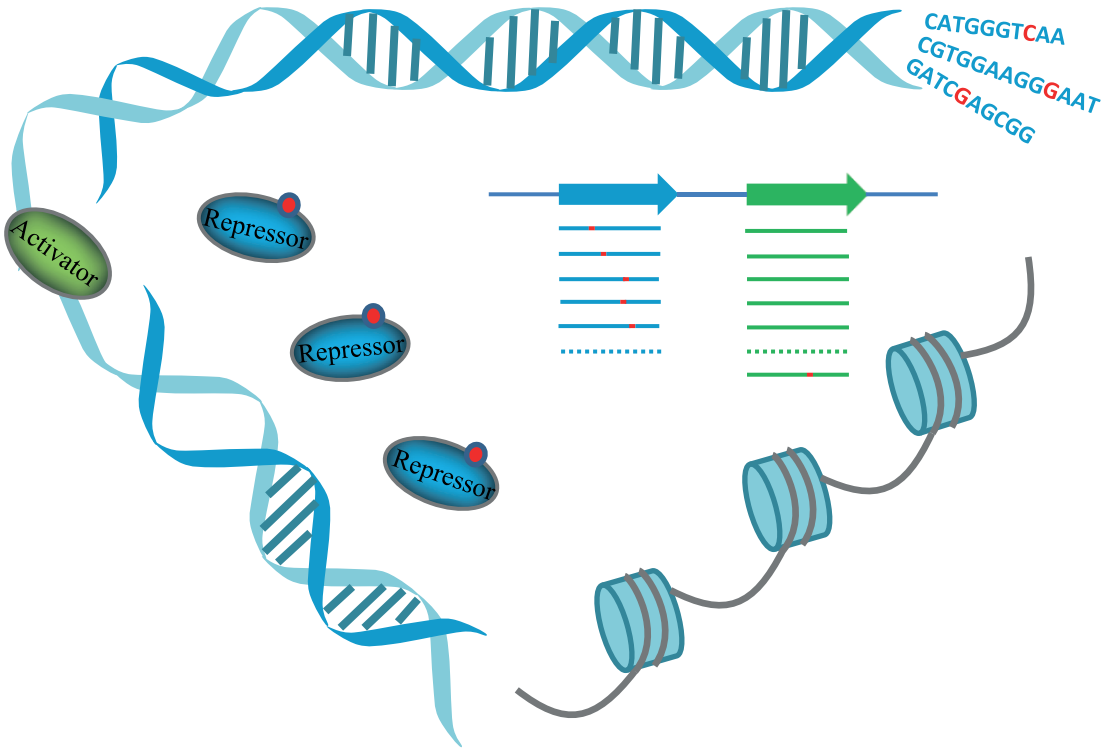
Transcriptional control of pectin degrading enzymes in Aspergillus niger

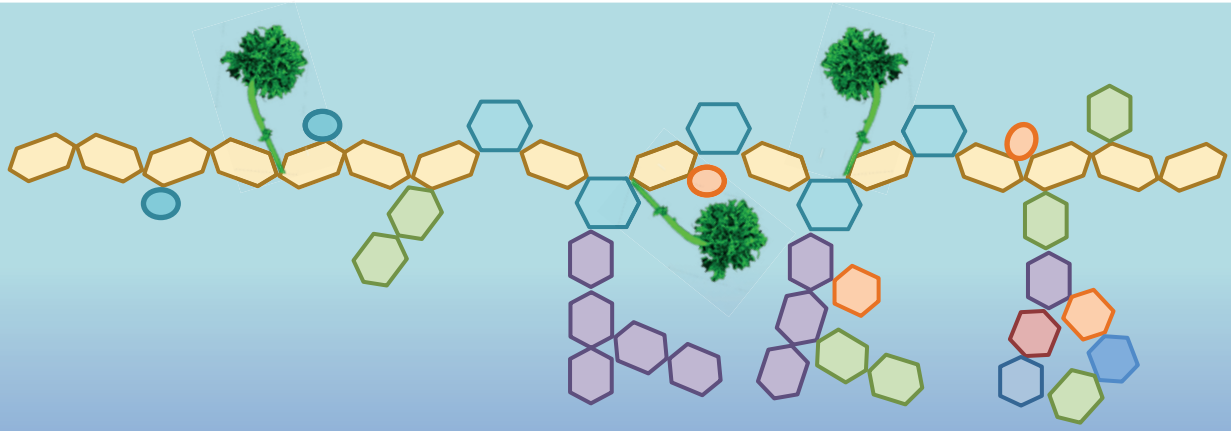
Niu, J.; Niu J.

Transcriptional control of pectin degrading enzymes in *Aspergillus niger*



Jing Niu

牛　静



Transcriptional control of pectin degrading enzymes in *Aspergillus niger*

Jing Niu

Cover: Designed by Jing Niu

Layout: Jing Niu

Printed by: Ridderprint BV, Ridderkerk, The Netherlands

ISBN: 978-94-6299-740-0

Transcriptional control of pectin degrading enzymes in *Aspergillus niger*

Proefschrift

ter verkrijging van

De graad van Doctor aan de Universiteit Leiden,

op gezag van Rector Magnificus Prof.mr. C.J.J.M. Stolker,

volgens besluit van het College voor Promoties

te verdedigen op woensdag 18 oktober 2017

klokke 10.00 uur

door

Jing Niu

Geboren te Qujing -Yunnan, China

in 1986

Supervisor

Promotor: Prof. Dr. P.J. Punt

Co-promotor: Dr. A.F.J. Ram

Promotiecommissie

Prof. Dr. C.A.M.J.J. van den Hondel (Leiden University)

Prof. Dr. P.J.J. Hooykaas (Leiden University)

Prof. Dr. V. Meyer (Technical University of Berlin)

Prof. Dr. H.A.B. Wösten (Utrecht University)

Dr. J.A.L van Kan (Wageningen University)

Contents **pages**

Chapter 1

General introduction

1. Filamentous fungi and *Aspergillus niger*

1.1 Filamentous fungi

Filamentous fungi refer to organisms that produce multi-cellular filaments called hyphae which are, with the exception of the Zygomycetes, regularly septated. About 60,000 filamentous fungi species (excluding yeasts) had been known by 1990s, and the number in nature was estimated to exceed 250,000 (Hawksworth and Kirsop 1988). Current estimates (T. Boekhout, personal communication) reach to several million species altogether. Most of the biochemical properties of these fungi have not been exploited at all. With more and more of these properties coming to light, an increasing number of filamentous fungi have been shown to be useful in industrial biotechnology (Meyer *et al.* 2011; Pessoa *et al.* 2017).

Filamentous fungi have been used for a long time in various applications. The most well-known filamentous fungi are probably those connected with food for humans. It is estimated that over 500 species are editable, but only less than 20 have been exploited commercially, including *Agaricus bisporus* (common mushroom), *Lentinula edodes* (shii-take), *Tricholoma matsutake* (matsu-take), *Volvariella volvacea* (paddy straw mushroom) (Chang and Hayes 1978). Another category of filamentous fungi is traditionally used to produce fermented food and drinks. The most prominent fungi in this category include species of *Aspergillus*, *Monascus*, *Mucor*, *Rhizopus* (Hesseltine 1965; Batra and Millner 1974; Steinkraus 1983) as well as certain *Penicillium* species (e.g. *P. camembertii* and *P. roqueforti*), which have long been used in cheese production (Pitt 1980). The third category of filamentous fungi constitutes of fungi known to produce interesting metabolites including antibiotics and organic acids (Turner 1971; Turner and Aldridge 1983). The species in this category include those that produce antibiotics, e.g. *Penicillium rubens*, which produces penicillin (Houbraken *et al.* 2011), and *Penicillium griseofulvum*, which produces anti-fungal griseofulvin (Macmillan 1954); those that produce ergot alkaloids, e.g. *Claviceps purpurea* (Amici *et al.* 1969); and those that produce growth hormones, e.g. *Fusarium moniliforme* (Meleigy and Khalaf 2009) and *Fusarium graminearum* (Mirocha and Devay 1971). The fourth category of filamentous fungi is a source for production of non-protein compounds with that serve as food additive such as vitamins (*Eremothecium ashbyii*: (Goodwin and Pendlington 1954), or polysaccharides (*Aureobasidium pullulan*s: (Heald and Kristiansen 1985). The fifth category of filamentous fungi includes important producers of industrial enzymes such as *A. niger* (amylases and pectinases) (Pandey *et al.* 1999), *Aspergillus oryzae* (proteases) (Chutmanop *et*

al. 2008), *Trichoderma reesei* (cellulase) (Montenecourt and Eveleigh 1977) and *Penicillium roquefortii* (lipase) (Eitenmiller *et al.* 1970). In the remaining part of this introductionairy chapter, I focus on the filamentous fungus *A. niger* in relation to its enzyme producing characteristics.

1.2 *Aspergillus niger*

A. niger, commonly known as black mold, is an asexual reproducing filamentous fungus that is ubiquitous in the environment. It is a common species of the fungal genus *Aspergillus*. *A. niger* can grow in a wide range of temperature conditions, with an optimal temperature for growth between 35–37°C (Schuster *et al.* 2002). A typical *A. niger* colony consists of the youngest, actively extending hyphae at the edge of the colony and the oldest, non-extending, sporulating mycelium at the center. During *A. niger* conidial development, the vegetative mycelium of the air-exposed colonies forms aerial hyphae, whose tips may swell to form a vesicle. Buds are formed on the vesicle that develops into metulae. Then on top of the metulae formed phialides, which give rise to chains of conidia (Krijgsheld *et al.* 2013). These conidia are the asexual reproductive structures.

A. niger has no known sexual life cycle. The asexual life cycle of *A. niger* goes through different stages and starts with spore swelling, followed by germtube outgrowth, germ-tube elongation and branch formation. On plates, but also in liquid cultures (Jorgensen *et al.* 2010), carbon starvation leads to the induction of the process of asexual development and the formation of conidiospores. In *A. niger*, conidia are black because of the present of melanin. *A. niger* is a biotechnologically important filamentous fungus and is used as an industrial cell factory for the production of organic acids and enzymes (Pel *et al.* 2007; Andersen *et al.* 2011). The oldest and most well-known application of *A. niger* is the production of citric acid. The bioprocess of citric acid production by *A. niger* is highly efficient, and over one million metric tons of citric acid being produced each year (Karaffa *et al.* 2001). As a soil saprobe, *A. niger* is also important for global carbon recycling. Like many other saprophytic fungi, *A. niger* can secrete a variety of hydrolytic and oxidative enzymes to degrade plant biomass. Moreover, *A. niger* is also an important model organism for studying some basic processes in life science, e.g. eukaryotic protein secretion (Baker 2006). Comparing with other microorganisms, *A. niger* has some attractive properties to be used for industrial fermentation. It can secrete enzymes efficiently and in large scale, and is generally regarded as a safe (GRAS) production organism. Due to long-standing experience with fermentation, many *A.*

niger strains with improved production have been generated and used by companies (Das and Roy 1978; Schuster *et al.* 2002).

2. Functional Genomics to understand gene function and regulation

In any organism, it is the regulated pattern of gene expression that determines the phenotype. Gene regulation is the means by which cells orchestrate gene activities to ensure that the right genes are expressed at the right time. The proper control of gene expression is important for cells to adapt to changing conditions such as nutrient availability, temperature and environmental stress. The aim of studying the control of gene expression is to understand how different regulatory networks exert their function. In this chapter, I will describe functional genomics approaches to study fungal gene function and gene regulation in relation to carbon source availability in more detail. These approaches include the rapid developments in (genome) sequencing, construction of gene editing and gene-knockout methods, which in combination allow new and efficient transcriptome analysis and forward genomics approaches required for studying gene regulation.

2.1 Genome sequencing technologies

Sequencing DNA molecules contributes greatly to research progress in biology and medicine. During the last 10 years, considerable progress has been made in genome sequencing technologies, allowing individual researchers to sequence fungal genomes within a few weeks and allow transcriptome analysis to study gene regulation without the need to generate microarrays.

DNA sequencing techniques have been going through three generations. Sanger sequencing is the most important first-generation sequencing technique. It has been the most widely used sequencing technique before being replaced by the next generation sequencing, such as Roche 454, Illumina, ABI/SOLiD which allows sequencing DNA samples in high-throughput. Most recently, the third-generation sequencing techniques was developed for single molecule sequencing. Table 1 shows the characteristics of different DNA sequencing methods.

Table 1: Characteristics of different DNA sequencing methods used for whole genome sequencing.

Technology	**Generation**	**Year**	**Approach**	**Reads per run**	**Average Read length (bp)**	**% Accuracy**	**References**
Sanger ABI 3730xI	1st	2002	Synthesis with dye terminators	~100	800	>99	(Kerstens 2010; Rhoads and Au 2015)
454/Roche FLX	2nd	2005	Sequencing by synthesis	>200.000	100	>95	(Kerstens 2010) Website*
IIIumina/Solexa	2nd	2006	Sequencing by synthesis	30.000.000	25-50	>95	(Kerstens 2010) Website*
PacBio	3rd	2011	Sequencing by synthesis	400.000	1300	80~90	Website*
Oxford Nanopore MinION	3rd	2015	Direct Sequencing	4.000.000	10.000	60~70	(Mikheyev and Tin 2014; Laver *et al.* 2015) Website*

*Website: https://flxlexblog.wordpress.com/2016/07/08/developments-in-high-throughput-sequencing-july-2016-edition/

The first-generation DNA sequencing technologies include Maxam-Gilbert sequencing and Sanger sequencing. Maxam and Gilbert developed chemical cleavage method for DNA sequencing in 1977 (Maxam and Gilbert 1977). The basic principle is using hydrazine to selectively remove bases from pyrimidines (cytosine and thymine), while hydrazine can only remove bases from cytosine in the presence of high salt concentrations; using acid to remove bases from purines (adenine and guanine) and to attack guanine with dimethyl sulfate. The resulting backbone is then cleaved at the abasic sites by piperidine, yielding fragment of different length, which can be visualized via polyacrylamide gel electrophoresis, and the order of nucleotides can be deduced by reading up the gel (Maxam and Gilbert 1977). This technique is considered the start of "first-generation" sequencing. A major breakthrough in DNA sequencing was made in late 1977, when Sanger and colleagues developed the "Sanger chain termination" method. Sanger chain termination is also called dideoxy chain termination (Sanger *et al.* 1977). The principle is adding four types of radiolabeled dideoxynucleotides (ddATP, ddTTP, ddGTP, ddCTP) instead of chemicals to four polymerase reactions respectively. Due to lack of the 3'OH group, which is required for DNA polymerase-mediated strand elongation, random incorporation of ddNTPs in a PCR reaction can terminate DNA extension at different positions, resulting in the generation of 3' truncated sequence fragments of different sizes in each of the four reactions. The fragments can be visualized by polyacrylamide gel electrophoresis and the order of nucleotides can be inferred after autoradiograph (Sanger *et al.* 1977) . This technique has improved with time, such as using fluorescent labelled dNTP which can read DNA sequence without gel electrophoresis. Whole genome sequencing became possible with this technique (e.g. *E. coli* genome, yeast genome), although still very expensive and time consuming. Towards this purpose, the invention of new sequencing techniques was therefore required to lead the way.

The second-generation DNA sequencing (or next generation sequencing) represents the first high-throughput DNA sequencing approach. The beginning of the second-generation sequencing techniques was marked by pyrosequencing, which was first introduced in 1996 (Ronaghi *et al.* 1996). This approach uses a luminescent method to measure pyrophosphate release on nucleotide incorporation. The basic principle is an enzymatic cascade reaction catalyzed by four enzymes (DNA polymerase, ATP sulfurylase, luciferase and bisphosphatase). During the reaction, if the added dNTP can be paired with the template DNA at a given position, the dNTP will be incorporated under the help of DNA polymerase, releasing equal amount of pyrophosphate. ATP sulfurylase then converts pyrophosphate to

ATP, which, in turn, drives the conversion of luciferin to oxyluciferin mediated by luciferase. Oxyluciferin produces light signals proportional to the number of pyrophosphates (Nyren and Lundin 1985). This finding evolved into the first "next generation sequencing" technology, and was used to sequence the *Mycoplasma genitalium* genome (Margulies *et al.* 2005). Pyrosequencing was initially licensed to 454 Life Science and later purchased by Roche. It allows parallel production of sequence reads from a large number of wells, and is therefore called high-throughput DNA sequencing. Several parallel sequencing techniques sprung up following the application of Roche-454. Among them, the most important one is Solexa/Illumina sequencing. Solexa/Illumina sequencing uses fluorescent reversible-terminator, and sequencing results are visualized by detection of fluorescent signals. These second generation approaches have some significant advantages over the first generation approaches, e.g. they can sequence DNA in high-throughput and can be observed in real time. Albeit these advantages, the short read length in NGS makes it difficult to assemble large genomes or those with lots of repeats.

The third-generation DNA sequencing is featured by single molecule sequencing (SMS), real time sequencing and long-read sequencing. Currently, the most widely used third-generation approach is the single molecule real time (SMRT) platform from Pacific Biosciences (PacBio). SMRT sequencing takes place in nano photonic visualization chambers called zero-mode waveguides (ZMWs). During SMRT sequencing, a DNA template is replicated by a DNA polymerase immobilized at the bottom of the ZMW. As the DNA template passes through the polymerase during synthesis, nucleotides that are phospho-labelled with four different fluorescents are incorporated into the newly synthesized strand one by one and are illuminated from below by an excitation beam, emitting four different bright light pulses. This process occurs in parallel in thousands of ZMWs that make up the SMRT cell (Goodwin *et al.* 2016; Heather and Chain 2016).

Another promising third-generation DNA sequencing technique is the Oxford nanopore sequencing (MinIon). This technique can directly detect the nucleotide composition of a ssDNA, negating the need for incorporation or hybridization of nucleotides guided by template DNA strand. The principle behind nanopore sequencing is that ssDNA can be passed through a protein nanopore by electrophoresis. The nanopore protein is embedded in a membrane. During sequencing, a constant voltage is applied to each side of the membrane, which drives the translocation of DNA through the pore. As the DNA passes through the pore, the nucleotides will change the ionic conductivity of the nanopore, resulting in nucleotide-

specific shifts in the magnitude and duration of the ion current. These shifts can then be recorded by a sensitive ammeter and serve as a read out to determine the nucleotide sequence (Goodwin *et al.* 2016; Heather and Chain 2016).

Starting from the second-generation approaches, these different sequencing technologies provide a powerful tool for research on functional genomics of fungi. One important application of these technologies is whole genome sequencing. There has been substantial investment in sequencing of filamentous fungi genomes, with a clear focus on sequence analysis of a very important class of fungi, the Aspergilli. The first sequenced filamentous fungus was *N. crassa*, a well-established filamentous fungus for basic fundamental research (Galagan *et al.* 2003; Mannhaupt *et al.* 2003). The first sequenced *Aspergillus* genomes include the model organism *Aspergillus nidulans*, as well as *Aspergillus fumigatus* and *A. oryzae* in 2005 (Galagan *et al.* 2005; Machida *et al.* 2005; Margulies *et al.* 2005). Subsequently, genomes of *Aspergillus flavus*, *Aspegillus fischeri*, *Aspergillus clavatus* and *Aspergillus terreus* were also sequenced (Payne and Loomis 2006; Fedorova *et al.* 2008; Arnaud *et al.* 2012). Two *A. niger* strains CBS513.88 and ATCC1015 were sequenced at 2007 and 2011, respectively (Pel *et al.* 2007; Andersen *et al.* 2011). In a recent study contributed by global consortium, ten more *Aspergillus* strains were sequenced and annotated. They are *Aspergillus luchuensis*, *Aspergillus tubingensis*, *Aspergillus brasiliensis*, *Aspergillus carbonarius*, *Aspergillus aculeatus*, *Aspergillus versicolor*, *Aspergillus sydowii*, *Aspergillus glaucus*, *Aspergillus wentii* and *Aspergillus zonatus* (De Vries *et al.* 2017). In the meantime, the complete genus *Aspergillus* has been sampled for genome sequencing, leading to the sequencing of almost 46 genomes (http://genome.jgi.doe.gov/eurotiomycetes/eurotiomycetes.info.html) with even more *Aspergillus* genomes underway (M. Andersen, personal information). These genome sequence data provide a resource-rich platform for evolutionary and functional genomics studies and provide reference genomes for transcriptomic studies via RNA sequencing. In the following chapters, the use of these platforms will be described in relation to the genetic characterization of mutants and to study gene functions.

2.2 Methods to genetically characterize mutant genes from forward genetic screens

During the last decades, forward genetic screens have identified many new genes in various species and contributed greatly to our understanding of gene functions. The essence of a forward genetic screens is to make random mutations to create mutants with specific

phenotypes and to identify the genetic basis of the mutations responsible for these phenotypes. In forward genetic screens, chemicals (e.g. ethyl methanesulfonate, EMS) or radiation-based mutagens (e.g. UV) are commonly used to generate random mutants, which are then screened for interesting mutant phenotypes. For the genetic identification of mutations from a forward screen, different approaches can be used. In the pre-genomic era, genetic linkage analyses in combination with chromosome walking and complementation approaches with cosmid libraries were generally used.

2.2.1 Genetic linkage analysis based methods

Genetic linkage analysis utilizes genetic markers to map the mutation of interest. Depending on the genetic background of the species, a variety of crossing schemes can be used to map the mutation that causes the phenotype of interest to a specific region of the genome (mapping interval). During the crossing process, markers which are closely linked to the causal mutation will be co-segregated with the causal mutation due to infrequent recombination between them. Therefore, there is a distinct allele distribution of the mutation and the closely linked markers in the progeny from a cross. Once the mutation region is mapped, a targeted search e,g, via chromosome walking can be conducted to find the actual causal mutation within that region by sequence analysis (see for review (Schneeberger 2014). Genetic mapping by this method is largely dependent on the density of the polymorphic markers genotyped. Moreover, as this method can only locate the genomic region that contains the causal mutation, further sequencing within this region is required (Schneeberger 2014).

2.2.2 Complementation analysis based methods

Confirmation of the mutated candidate gene responsible for the phenotype can also be achieved by complementation analysis. In this approach, a cosmid library is constructed by ligating genomic DNA fragments into the cosmid vector. Introduction of the cosmid library into the mutant strain allows selection of transformants functionally complementing the causal mutation. If the cosmid clone contains a wild-type allele of the mutated gene, it can rescue the phenotype by complementing the endogenous disrupted allele. Further analysis of complementing cosmid clones will reveal the gene contained in the complementing sequences (Damveld *et al.* 2008; Punt *et al.* 2008; Meyer *et al.* 2009). This method has been successfully used in *A. niger*, for example for the identification of PrtT, a unique regulator of extracellular protease encoding genes (Punt *et al.* 2008). However, the complementation method is time

and labor intensive and has some limitations, such as that the gene might be lacking in the library, and that certain mutant phenotypes are difficult to screen for complementation among thousands of transformants. The next generation sequencing approaches provides a promising alternative method for identifying specific gene mutations.

2.2.3 Next-generation sequencing based methods

With the advent of next generation sequencing (NGS) techniques it is possible to directly sequence individual mutant genomes to identify causal mutations (Srivatsan *et al.* 2008). However, multiple mutations might be found in mutants and requiring a lot of research to identify the mutation responsible for the phenotype. Therefore, several approaches have been developed to facilitate identification of the mutation related to the phenotype of the mutant. Recently, the combination of the classical bulk segregant analysis (BSA) (Michelmore *et al.* 1991) with NGS has proven to greatly accelerate this process, leading to the development of an approach named mapping-by-sequencing (Schneeberger *et al.* 2009; Niu *et al.* 2016).

BSA is traditionally used to identify makers linked to gene(s) of interest (Michelmore *et al.* 1991). It involves comparing the pooled DNA sample of mutant segregants with that of wild-type segregants. Both segregants are resulted from a single cross of the parental strains. The individuals in each pool have the same version of the target gene (either wild-type or mutated), but are arbitrary in all the other genes. By genome sequencing, single-nucleotide polymorphisms (SNPs) are analyzed between the two parental strains and serve as markers. Markers that are homozygously polymorphic between the two segregant pools are within physical proximity of the mutation and thus genetically linked to the locus of the target mutation (Michelmore *et al.* 1991; Lister *et al.* 2009). This approach, combining BSA with NGS, allows simultaneous mapping and identification of the target mutation. In our lab, we used bulk segregant analysis in combination with high-throughput genome sequencing to identify the mutation gene *laeA*, which is responsible for the non-acidifying phenotype in *A. niger* (Niu *et al.* 2015). In case of a very specific mutant selection approach, spontaneous mutants with the same mutant phenotype can be directly used to identify the causal mutations without bulk segregant analysis, as there are less non-targeted mutations and the selection scheme only results in one phenotype in multiple mutants. In our lab, directly sequencing of several individual *A. niger* mutants revealed a transcriptional repressor which control expression of genes for D-galacturonic acid utilization (Niu et al., 2017).

2.3 *Aspergillus niger* functional genomics

The availability of high quality genome sequence of *A. niger* in combination with improved annotations of the genome has resulted in the identification of 11800 potential genes (http://genome.fungalgenomics.ca/new_gene_model_pages/species_search_page.php?predname=Aspni_NRRL3), of which most still await further functional analysis. In this thesis we focus on functional analysis of regulatory genes in particular. There are two common ways to study the function of a (regulatory) gene *in vivo*: deletion analysis or overexpression analysis. In this chapter, I will focus on two more recent and highly efficient technologies for making gene deletion mutants: the split marker approach and the CRISPR-Cas9 system.

2.3.1 NHEJ mutants combined with the split marker approach

Although in unicellular fungi random integration of DNA via the non-homologous end joining (NHEJ) pathway does occur, targeted integration is already much more efficient than in filamentous fungi (Kooistra *et al.* 2004). In particular in filamentous fungi, new approaches were required to make targeted integration more efficient. An important breakthrough in fungal genetics was the discovery of gene *YKu70* encoding protein responsible for NHEJ in yeast *Saccharomyces cerevisiae* (Van Attikum *et al.* 2001). Making NHEJ-deficient mutants was first performed in yeast *S. cerevisiae*, and thereafter in *N. crassa* via deletion of the *ku70* gene, and deletion of the *ku70* gene was shown to lead high frequencies of HR and consequently high efficiencies of obtaining targeted mutants (Van Attikum *et al.* 2001; Van Attikum and Hooykaas 2003; Ninomiya *et al.* 2004). In addition to *ku70*, deleting other components of the NHEJ machinery, such as *ku80* and *lig4* resulted in fungal NHEJ-deficient recipient strain for gene targeted deletion (for reviews see (Meyer and Bailis 2008; Kuck and Hoff 2010) and references therein).

Gene targeted deletion is normally performed by constructing a linear DNA fragment that contains the 5'and 3'flanks of the gene of interest (GOI) and a selection marker between them. The easiest way to generate these fragments is by fusion PCR in which the three fragments (5'flank, selection marker and 3'flank) are fused together by primer overlap extension. Although these methods work in general well, the full length PCR fragments are quite large in size (4-5 kb, depending on the size of the flanking sequence and selection marker used) with sometimes leads to PCR problems and low yields. To circumvent amplification of these large fragments the split marker approach was developed. In split marker approach, the gene

deletion cassette consists of two fragments. The first fragment contains the 5'flank of the GOI fused with a 3' truncated version of the selection marker. The second fragment contains a 5' truncated version of the selection marker that still overlaps with the first one and is fused with 3'flank of the GOI (Fairhead *et al.* 1996; Nielsen *et al.* 2006; Goswami 2012). Using this approach, PCR fragments are smaller in size (3 kb) thereby increasing success rate and yield of the PCR reaction. Both fragments are transformed simultaneously to the strain of choice. Strategies using the split marker approach lead to more efficient gene deletion in strains with an intact NHEJ-machinery (Nielsen *et al.* 2006). We have used the spilt marker approach in combination with NHEJ mutants for generating in an even more efficient way gene deletion mutants. In chapter 2, efficient generation of *A. niger* knockout strains by combining NHEJ mutants and a split marker approach is described in detail. Although these two methods can be used separately, the NHEJ mutants help to significantly increase the frequency of homologous recombination when using the split marker approach.

2.3.2. CRISPR-Cas9 approaches

Genome editing technologies that allow us to delete, insert, and modify DNA sequences have greatly accelerated our understanding of the functional organization of the genome. Currently, the most rapidly developing genome editing technique is the CRISPR-Cas9 system, a RNA-guided DNA editing technique that originates from type II CRISPR-Cas systems. In bacteria, CRISPRs (clustered regularly interspaced short palindromic repeats) provide acquired immunity against viruses and plasmids (Horvath and Barrangou 2010; Wiedenheft *et al.* 2012). Typical CRISPR loci consist of a CRISPR array of repeated sequences separated by variable sequences called spacers, which match the sequences within the invading foreign DNA (protospacer), and are often adjacent to CRISPR associated (Cas) genes that encode RNA-guided DNA nucleases (Hsu *et al.* 2014). During adaptive immunity, certain Cas enzymes incorporate segments of the invading DNA into the CRISPR array as spacers. In type II CRISPR-Cas systems, the CRISPR array is firstly transcribed into pre-CRISPR RNA (pre-crRNA). A trans-activating crRNA (tracrRNA) then hybridizes with pre-crRNA to form a RNA duplex, which can be cleaved and processed by RNAase III to produce mature tracrRNA:crRNA hybrids. In the hybrid, the small crRNA contains a repeat portion that hybridizes with tracrRNA and a spacer portion that can recognize the target DNA sequence by base pairing. The tracrRNA:crRNA duplex then pairs with the target DNA sequence and directs the Cas protein to introduce a site-specific double strand break (DSB) in the DNA (Doudna and Charpentier 2014; Hsu *et al.* 2014; Sander and Joung 2014).

In the CRISPR-Cas9 genome editing system, the tracrRNA:crRNA duplex is engineered as a single guide RNA (sgRNA). By redesigning crRNA, the CRISPR-Cas9 system can target any region of interest in the genome as long as it is adjacent to a protospacer adjacent motif (PAM). Due to ease of use and efficiency of this technique, it holds great promise to help us understand gene function. The CRISPR-Cas9 system has been tested in several *Aspergillus* species and has been shown to be effective in targeting genes. For example, CRISPR-Cas9 can efficiently introduce directed mutations into the *yA* gene in *A. nidulans*, the *albA* and *pyrG* gene in *A. aculeatus*, and *albA* homologs in five Aspergilli (*A. brasiliensis*, *A. carbonarius*, *A. luchuensis*, *A. niger* and *A. tubingensis*) (Nodvig *et al.* 2015). Moreover, it has been reported that a strain generated by CRISPR-Cas9 and containing a *pyrG* marker is capable for iterative gene targeting (Nodvig *et al.* 2015). Combining CRISPR-Cas9 gene targeting with transformation with "repair DNA" allows not only disrupting a gene but also specific gene editing. Together, it is clear that CRISPR-Cas9 is a promising technique to employ genetic engineering in these fungi and holds great potential in helping us understand their biology.

2.3.3. Overexpression analysis

Yet another way to study gene function is by overexpressing the GOI and study the phenotypic effects of overexpression. The most common strategy of overexpressing a gene is to put the gene under control of a strong constitutive promoter (Zhang 2003) or using an inducible promoter system such as the Tet-on system (Vogt *et al.* 2005; Meyer *et al.* 2011). The Tet system is involving the repressor protein TetR from *Escherichia coli*, which binds to the operator sequence (*tetO*) of the Tn*10* in the absence of tetracyclines and prevents the transcription of the operon. In the presence of tetracycline, TetR dissociates from *tetO*, initiating the transcription of the operon (Beck *et al.* 1982). This system was modified to generate a hybrid transactivator tTA by combining the TetR with the minimal transcriptional activation domain derived from the herpes simplex virus protein 16 (VP16) for application in eukaryotic systems. In this system (Tet-off system), tTA stimulates gene expression in the absence of tetracycline. Alternatively a Tet-on system has been developed. In the Tet-on system, the reverse hybrid transactivator rtTA was generated by introduction of mutations to TetR, which lead to induction of gene expression in the presence of tetracycline instead of repression. The Tet-on system can be used for maximum expression levels by placing several copies of the *tetO* sequence upstream of a minimal promoter. Both the Tet-ON system and the Tet-OFF system has been adapted to be functional in *A. niger* (Meyer *et al.* 2011; Wanka *et al.*

2016). Overexpression using these strategies often gives rise to an exaggerated phenotype due to overexpression of the targeted genes of the regulatory network, which directly imply the function of the gene.

2.4 Transcriptomics and related technologies

As described above, to dissect the role of regulatory proteins such as transcriptional activators in the regulatory networks, overexpression and deletion strategies are frequently used to study their effects on the expression of their target genes. The set of genes that are regulated as a unit or controlled by the same regulatory gene comprise a regulon (Anderson 2010). The regulon includes genes whose expression is collectively controlled and likely to be involved in a specific functional program. The approaches described below allow us to study gene regulation on a large scale to identify these regulons. In this thesis, these technologies have been studied to understand the role of a transcription al activator (GaaR) and repressor protein (GaaX) in relation to polygalacturonic acid utilization in *A. niger*.

2.4.1 First generation genome wide transcriptome analysis: Microarrays

Traditional approaches of detecting gene expression include northern blot, *in situ* hybridization and quantitative PCR (Q-PCR). While they are useful for studying single or a few genes, it is not possible to systematically survey genome wide gene expression using these traditional methods. The invention of DNA microarrays has greatly transformed the traditional way of studying gene expression, and allowed to detect and quantify tens of thousands of genes simultaneously (Kurella *et al.* 2001). The basic principle behind DNA microarrays is to immobilize a large number of known DNA sequences (probes) on a solid surface (e.g. nylon membrane, microscope slides and silica gel etc.) in an ordered array. These probes are used to specifically hybridize complementary DNA (sometimes termed the 'target') that is present in a sample (Schena *et al.* 1995; Huang *et al.* 2007). In case of transcriptional profiling the fluorescently labelled target DNA is generated from RNA samples by RT-PCR based approaches. After hybridization, the fluorescently labelled target sequences that bind to a probe generate a signal, which can be detected by laser-scanning and fluorescence detection devices such as CCD cameras. The hybridization patterns generated on the microarray can be read and the results can be quantitatively analyzed. In a previous study, researchers made use of gene expression profiling of *A. niger* grown on various carbon sources using Affymetrix DNA microarrays, and found several genes that were specifically induced by galacturonic

acid (Martens-Uzunova and Schaap 2008). The results of these studies allowed us to search for regulation factors involved in the co-regulation of these genes (Chapter 4).

Since its conception in 1995, DNA microarrays have developed into a powerful tool for surveying gene expression efficiently and comprehensively on a genomic scale. At the start of my thesis, microarrays were a state of the art technology that was available for *A. niger* and was therefore used most often. Although alternative gene expression techniques, among which is the next generation sequencing (NGS), are available today, these technologies were expensive at that time and as microarrays had been designed, production costs of arrays are relatively low. An increasing amount of knowledge about gene regulation in *Aspergillus* has been gained from studies using microarrays. A trispecies *Aspergillus* microarray was also developed for transcriptome analysis of *A. nidulans*, *A. niger*, and *A. oryzae* (Andersen *et al.* 2008). Despite their wide spread use, DNA microarrays continue to have some limitations, which include: inflexible probe design and strain variations which may influence hybridization signals of genes containing multiple DNA polymorphisms. Moreover, DNA microarrays can only be developed for species whose genome sequence has been determined. With the advent of even cheaper high-throughput DNA sequencing technology, DNA microarray is rapidly replaced by RNA sequencing (RNA-Seq).

2.4.2. RNA-Seq

RNA-Seq is an emerging technology that uses next generation sequencing to map and quantify transcriptomes. It provides a powerful tool to reveal many different properties of the transcriptome and to accurately measure all transcripts of an organism, including messenger RNAs, microRNAs, small interfering RNAs, and long noncoding RNAs (Wang *et al.* 2009). The typical protocol for RNA-Seq is to extract RNA, convert it into a library of cDNA fragments and attach them to sequencing adaptors, and sequence the cDNA library using high-throughput sequencing technology. After sequencing, the resulting reads, including exonic reads, junction reads, and poly(A) end-reads can be mapped to a reference genome or de novo assembled if the genome is unknown. This generates a base-resolution expression profile for each gene in the genome. Comparing with microarrays, RNA-seq has several key advantages (Wang *et al.* 2009). First, it can be used for species whose genomic sequences have not yet been determined and does not require an optimal genome annotation to predict open reading frames. Second, it can actually reveal transcript structure to a single-base resolution. Many properties of transcript structure (e.g. the precise location of transcription

boundaries, the connectivity of exons etc.) can therefore be accurately determined, making it useful for studying complex transcriptomes. Third, it has much lower background signals for sequence mapping and a higher dynamic range for measurement of transcriptional levels than microarrays. Fourth, RNA-seq is highly accurate for measurement of expression levels and the results have high levels of reproducibility. Finally, RNA-seq needs less RNA sample due to that no cloning steps are required. Due to these advantages, most researchers, including our own group, have completely shifted form (Affymetrix) microarray analysis to perform transcriptomic studies using RNA-seq based technology. In Chapter 4 and 5, RNA-seq has been used to perform transcriptomic studies to identify the genes controlled by the transcriptional activator and repressor module that control the expression of galacturonic acid induced gene expression in *A. niger*.

2.4.3 CHIP-seq analysis

Essential components of any gene regulatory network are DNA-binding proteins, such as transcription factors. Transcription factors can be activators, that is activation of the transcription factor stimulate gene expression, but transcription factor can also bind to DNA and act as repressor. DNA binding transcription factors (either acting as an activator or repressor) often bind to specific transcription factor binding sites in the promoter of target genes, thereby controlling their expression. Therefore, DNA-protein interactions play a fundamental role in the regulation of gene expression.

Historically, DNA-protein interactions can be identified by chromatin immunoprecipitation (CHIP) experiments. In CHIP studies, proteins are binding to DNA, followed by immunoprecipitation of the protein of interest with a protein-specific antibody, the precipitated protein-DNA complexes are then purified and the bound DNA is characterized. In early stages, the bound DNA was characterized by dot blot or Southern blot analysis. Further development of CHIP combines this technique with genome wide microarrays, leading to the invention of CHIP-chip method which allows hybridizing fluorescently labeled bound-DNA to an appropriate microarray at a relatively high-throughput (Ren *et al.* 2000). With recent advances in the next-generation sequencing, CHIP sequencing (CHIP-seq) was developed to sequence the released bound-DNA with short reads at a higher throughput. The short reads delivered in CHIP-seq allow identification of interaction sites with more precision. CHIP-seq has first been applied to identify the binding sites of STAT1 and NRSF at the genome-wide scale (Johnson *et al.* 2007; Robertson *et al.* 2007) and has

been used in several studies to define direct binding sites. The DNA binding targets of the Clr-1, Clr-2 and Xlr-1 transcription factors in *N. crass*a that are involved in the regulation of genes involved in plant cell wall biomass deconstruction have been identified via Chip-seq (Craig *et al.* 2015). The CHIP-seq technology represents a powerful tool to verify direct binding of a transcription factor to a promoter element.

2.5 Functional Genomics to understand gene regulation

With the rapid development of DNA and RNA sequencing technologies, more and more genomic sequences and transcriptional data of fungi are available. Bioinformaticians assemble whole genomic sequence for each species and create websites to store the sequences information and related protein information and other information as soon as they are available. Several websites are accessible such as JGI (http://genome.jgi.doe.gov/), fungal special database FungiDB (http://fungidb.org/fungidb/) or *Aspergillus* genome database AspGD (http://www.aspgd.org/) to view or download information. The study of gene regulation is clearly focused on studying the role of pathway specific and wide domain regulatory proteins, which in the majority of cases are DNA binding proteins governing transcription. These so-called transcription factors consist of two or more domains. One is a DNA binding domain (DBD), which attaches to a specific DNA sequence that is present upstream to the translational start site of a regulated gene. The second is a transactivation domain (TAD), to which other proteins (co-regulatory proteins) bind. DBD domains are commonly classified into different type including Zinc finger, helix-turn-helix, leucine zipper and helix-loop-helix based on the secondary structure. Zinc finger are categorized into three main classes Cys_2His_2 (C2H2), Cys_4 (C4) and Cys_6 (C6) (Macpherson *et al.* 2006; Shelest 2017). Proteins with a Zn(II)2Cys6 domain are found exclusively in fungi and yeasts. Chang et al., conducted genome-wide analysis of the Zn(II)2Cys6 zinc cluster-encoding gene family in *Aspergillus flavus* resulting in 199 genes encoding proteins with a Cys6 domain (Chang and Ehrlich 2013). Detailed genome mining in *A. niger* revealed the presence of 694 putative DNA-binding transcription factor of which 453 belong to the Zn(II)2Cys6 zinc cluster family (A. Ram, personal communication).

3. *A. niger* as an industrial important enzyme producer

A. niger is an industrial important enzyme producer, it can produce a wide range of enzymes involved in modification and degradation of plant polysaccharides, such as starch, inulin, cellulose, hemicellulose (mainly xylan and arabinan), galactomannan and pectin (De Vries

and Visser 2001). In this paragraph, I will focus on the structure and regulation of gene expression in relation to starch, xylan and pectin utilization by *A. niger*.

3.1 Major polysaccharides starch, xylan and polygalacturonic acid

Polysaccharides are polymeric carbohydrates, composed of 10 to up to several thousand monosaccharides linked together by glycosidic linkages. The most common monosaccharides that appear as parts of polysaccharides are glucose, fructose, xylose, arabinose, galactose, rhamnose and mannose. In addition, galacturonic acid (GA) is the most important sugar acid in plant cell wall and present as the main component of pectin. Plant cell wall polysaccharides can be classified into storage components (starch and inulin) and structure components such as cellulose, hemicellulose, and pectin.

3.1.1. Starch

Starch or glycogen are multi-branched polysaccharides consisting of α-1,4 and α-1,6-linked glucose residues that serves as a form of energy storage in humans, animals, and fungi. It is one of the most abundant storage polysaccharides in nature, functioning as a short- and long-term reserve carbohydrate. Starch is produced in the plastids of higher plants and accumulated as granules in chloroplasts of source organs such as leaves (transitory starch) or in amyloplasts of sink organs such as seeds, tubers and roots (storage starch). Some plants with high starch content include corn, potato, rice, sorghum, wheat, and cassava. Starch is made up of two substructures including amylose and amylopectin. Amylose is a linear chain composed of 100-10000 glucose units with α-1,4-glucosidic bonds. Amylopectin consists not only the linear backbone of glucose units with α -1,4-glucosidic bonds but also branches composed of α -1,6- glucosidic linkages. The number of branches and the length of the side chains vary among different sources of starch. A complete amylopectin molecule contains on average about 2,000,000 glucose units (Myers *et al.* 2000). Starch granules consist of tightly packed glucan chains resulting in a semicrystalline, water–insoluble structure, which is suitable for long-term storage. In general, the more the chains are branched, the more the starch is soluble. Glycogen is very similar in structure in relation to starch but containing an average a higher portion of α-1, 6- glucosidic linkages and is therefore more branched. Glycogen is found in fungi and as well in mammalian cells and also functions as a storage carbohydrate (Gilbert 2000).

3.1.2. Xylan

Xylan is a major constituent of plant cell wall hemicellulose. The content of xylan comprises 10 - 35 % of the hemicellulose in hardwoods and 10 - 15 % of the hemicellulose in softwoods. After cellulose, xylans are the second most abundant structural polysaccharides in plants. The deposition of xylan in the secondary cell wall contributes to the construction of a strong and flexible plants cell wall and hence helps to defend against herbivores and pathogens. Therefore, xylan is important for normal plant growth and development. The structure of xylan is characterized by a backbone composed of a linear polymer of β-1,4-glycoside-linked xylose residues. Xylans of all higher plants possess this backbone, which is usually substituted with acetyl, glucuronic acid, 4-O-methylglucuronic acid, and arabinose residues. Despite the common features, variations in xylan structures also among different species and even among different tissues in the same species (Rennie and Scheller 2014).

3.1.3 Pectin

Pectin is the main constituent of the middle lamella of plant cell wall. The middle lamella is found as the outermost layer of the plant cell wall which consists of up to three layers. The layer formed between the middle lamella and plasma membrane is called primary cell wall. The primary cell wall is mainly composed of cellulose microfibrils contained within a gel-like matrix of hemicellulose fibers and pectin polysaccharides. The third layer is called secondary cell wall, and is formed between the primary cell wall and plasma membrane in some plants. In addition to polysaccharides, plant cell walls also contain lignin as structural component and many proteins with enzymatic functions and (hydroxyproline-rich) glycoproteins with structural functions (Rose and Lee 2010).

Galacturonic acid (GA) is the most abundant component of pectin. Pectin is a collective name for GA-rich structures and four substructures have been defined which include: *i)* homogalacturonan (HGA) or polygalacturonic acids (PGA), *ii)* xylogalacturonan (XGA), *iii)* rhamnogalacturonan I (RG-I) and *iv)* rhamnogalacturonan II (RG-II) (reviewed in (Mohnen 2008). PGA is a linear polymer, consisting of α-1,4-linked D-galacturonic acid residues. The backbones of XGA and RG-II are made up of α-1,4-linked D-galacturonic acid residues. In XGA, β-D-xylose residues are β-1,3-linked to GA residues of the PGA backbone. The backbone of RG-I is made up of alternating GA and L-rhamnose residues (reviewed in (Mohnen 2008; Leijdekkers *et al.* 2015). The side chains of RG-I are mainly arabinan and arabinogalactan comprising of L-arabinose and D-galactose residues (Mohnen 2008). RG-II

is the most complex structure and side chains of RG-II are composed of up to twelve different types of monosaccharides in >20 different linkages (Mohnen 2008). The abundance of each substructure varies with plant species, but typically homogalacturonan is the most abundant polysaccharide in pectin (65%) followed by rhamnogalacturonan I (25-30%). Xylogalacturonan and rhamnogalacturonan II comprise less than 10% of the total pectin (Mohnen 2008).

3. 2 Degradation of polysaccharides by *A. niger*

A. niger is a typical saprophytic fungus feeding on plant litter. Saprophytic fungi convert the plant polysaccharides into mainly monosaccharides before uptake of the monosaccharides into the cell for further catabolism. They degrade plant litter by secreting substrates-specific enzymes (mainly hydrolytic enzymes). The expression and consequent secretion of these enzymes is tightly controlled and dependent on which carbon source is available. Like many other filamentous fungi, *A. niger* has a rich arsenal of different enzymes able to plant polysaccharides. In the introduction I will focus on the enzymes and their regulation on relation to starch, xylan and pectin.

The storage polysaccharide starch/glycogen is the principle carbon reserve in many plants as well as for microorganisms (both bacteria and fungi), and higher eukaryotes, including humans. A variety of enzymes participate in the hydrolysis of starch (Steup 1988). Bacteria and fungi are also specialized in the extracellular degradation of plant-derived starch by secreting starch-degrading enzymes. These enzymes are categorized into three major glycoside hydrolase (GH) families (Coutinho and Henrissat 1999): α-amylases belonging to the GH13 family of endo-amylases, glucoamylase type enzymes (exo-acting enzymes) of family GH15 for releasing α-(1,4)- and α-(1,6)-glucose, and additionally α-(1,4)-glucosidases of family GH31 for releasing α-glucose from the non-reducing end of starch. A detailed annotation of starch degrading enzymes in *A. niger* was performed previously (Yuan *et al.* 2008b).

Most fungi are also capable of degrading xylan and metabolizing the resulting xylose. Since *S. cerevisiae* cannot naturally convert xylose into ethanol, The uptake and intracellular metabolism of xylose has received a lot of attention to construct recombinant *S. cerevisiae* strains, which containing the intracellular enzymes involved in xylose metabolism. As a result the xylose part in the plant cell wall could not be metabolized by natural *S. cerevisiae* strains to produce ethanol. Xylose fermenting yeasts have been developed to also utilize the

xylose form plant biomass (Azhar *et al.* 2017). To degrade xylans, several hydrolytic enzymes need to work synergistically. Among these enzymes, the most important one is endo-1,4-β-xylanase, which cleaves the xylosyl backbone and releases short xylooligosaccharides, and xylan 1,4-β-xylosidase, which hydrolyzes xylooligosaccharides into xylose units (Shallom and Shoham 2003). Depending on the type of xylan, various auxiliary enzymes, such as α-arabinofuranosidases, β-galactosidases, α-galactosidases, α-glucuronidases and feruloyl esterases, are also required for efficient deconstruction of xylan sidechains. Many microorganisms, including bacteria, yeasts and filamentous fungi, are important producers of these xylanolytic enzymes (Biely *et al.* 2016). Filamentous fungi like *Aspergillus* and *Penicillium* species are particularly important xylanase producers because they secrete the enzyme into media at higher levels than other microorganisms (De Vries and Visser 2001; Chavez *et al.* 2006).

Pectin degrading enzymes are mainly produced in nature by saprophytes and many bacterial and fungal pathogens of plants for degradation of plant cell wall. Commercial pectinase preparations are primarily derived from *A. niger* (Voragen and Pilnik 1989). Genome mining has revealed a large array of extracellular pectinolytic enzymes in *A. niger* (Coutinho *et al.* 2009; Martens-Uzunova and Schaap 2009). Pectin degrading enzymes can be grouped in two major classes "pectinases" and "accessory enzymes" according to the complex structure of pectin. The "pectinases" attack the backbone of pectin, and "accessory enzymes" degrade the side chains of pectin. Homogalacturonan (HGA) is most abundant component in pectin (Harholt *et al.* 2010). During HGA degradation, pectin methylesterases hydrolyze methoxy groups in pectin to yield pectate and methanol. Endo-polygalacturonases and exo-polygalacturonases are hydrolytic enzymes that hydrolyze pectate, producing oligogalacturonic acid and GA respectively. Pectate lyases are endo-acting enzymes that catalyze pectate to unsaturated oligogalacturonides with an eliminative cleavage mechanism. Pectin lyases are endo-acting enzymes with an eliminative cleavage mechanism on naturally methylated pectin (Hsiao *et al.* 2008). The backbone XGA can be degraded by endo-xylogalacturonan and exo-polygalacturonan hydrolases, whereas RGI requires the additional activity of rhamnogalacturonan hydrolases and rhamnogalacturonan lyases.

3.3 Galacturonic acid metabolism in *A. niger*.

To utilize GA as a carbon source, GA have to be taken up into the cell by specific sugar transporters (Sloothaak *et al.* 2014). GA can be metabolized both by bacteria and in

eukaryotes using different enzymatic pathways. In bacteria, GA is metabolized in a five-step pathway via D-tagaturonate, D-altronate, 2-keto-3-deoxy-gluconate and 2-keto-3-deoxy-6-phospho-gluconate resulting in the formation of pyruvate and glyceraldehyde-3-phosphate (Ashwell *et al.* 1960; Huisjes *et al.* 2012). In eukaryotes, the metabolism of GA takes a different metabolic route. Metabolism of GA in fungi is well studied and involves four enzymatic reactions to convert GA into glycerol and pyruvate. The genes encoding these enzymes (*gaaA*, *gaaB*, *gaaC*, and *gaaD*) have been identified and the biochemical properties of the enzymes have been determined (Kuorelahti *et al.* 2005; Kuorelahti *et al.* 2006; Liepins *et al.* 2006; Hilditch *et al.* 2007; Mojzita *et al.* 2010; Wiebe *et al.* 2010; Zhang *et al.* 2011; Kuivanen *et al.* 2012). Specific sugar transporters that are able to transport GA over the plasma membrane have recently been identified and characterized in *N. crassa* (Benz *et al.* 2014) as well as in *A. niger* (Sloothaak *et al.* 2014) and *Botrytis cinerea* (Zhang *et al.* 2014).

4.1 Transcriptional regulation of genes encoding polysaccharides degrading enzymes

As described above, *A. niger* can secret wide range of enzymes to synergistically degrade plant cell wall polysaccharides. The expression of these enzymes is tightly regulated in filamentous fungi including *A. niger*. In many cases, the expression is under the control of substrate specific transcriptional activators, which belong to the fungal specific transcription factors with a Zn(II)2Cys6 DNA binding motif (Todd and Andrianopoulos 1997). Their expression of the genes encoding the extracellular enzymes, sugar transporters, intracellular metabolic enzymes, and in some cases also the transcriptional activator is also controlled by wide-domain regulators, such as carbon catabolite repressor CreA and the ambient pH regulator PacC which are both members of the C2H2 family of transcription factors. Table 2 shows the main pathway specific transcription factors from *Aspergilli* involved in the degradation of various plants derived polysaccharides known to date (see review (Benocci *et al.* 2017)). Here, I will focus on the transcription factors AmyR, XlnR and the carbon catabolite repressor CreA as an introduction to our study on the search for regulatory factors involved in controlling the expression of enzymes related to PGA utilization.

4.2 AmyR

Aspergillus spp can produce different types of amylolytic enzymes to degrade starch. The expression of genes encoding these synergistically acting amylolytic enzymes is regulated by a transcription activator AmyR. AmyR was first cloned and characterized in *A. oryzae* (Petersen *et al.* 1999), and encodes a 604 amino acids protein containing a zinc binuclear

cluster motif (Zn(II)2Cys6) (Petersen *et al.* 1999; Gomi *et al.* 2000). AmyR was also cloned and sequenced in *A. nidulans* and *A. niger* by heterologous hybridization of the *A. oryzae amyR* gene (Tani *et al.* 2001). The *A. nidulans* AmyR protein comprises 662 amino acids and shares 72% overall amino acid identity to the AmyR of *A. oryzae* (Tani *et al.* 2001). The *A. niger* AmyR protein is 610 amino acids long and shows 66.18% identity to the AmyR of *A. oryzae*.

The regulatory mechanism of expression of the amylolytic genes mediated by AmyR is relatively well studied in *A. nidulans* by subcellular localization studies using a green fluorescent protein (GFP)-labeled AmyR fusion protein (GFP-AmyR). Unfortunately, these localization studies were performed by placing the GFP-AmyR fusion protein under control of the inducible *alcR* promoter which could affect the conclusions related to AmyR localization under inducing and non-inducing conditions (Makita *et al.* 2009). AmyR possess five conserved domains with *S. cerevisiae* Mal regulators (Tani *et al.* 2001) (Figure 1) which include the Zn(II)2Cys6 domain (Zn) and four domains, named MH1-4. The N-terminal Zn(II)2Cys6 domain is the DNA- binding domain, which binds to the proposed AmyR binding site ($CGGN_8CGG$) which is present in the promoter regions of the various amylolytic genes (Petersen *et al.* 1999; Tani *et al.* 2001; Ito *et al.* 2004). An alternative binding site ($CGGN_8AGG$) in the Taka-amylase A (taaG2) promoter in *A. oryzae* has been shown to be functional (Ito *et al.* 2004). The N-terminus also contains the nuclear localization signal (NLS) sequences which is responsible for the nuclear localization of AmyR (Makita *et al.* 2009). The nuclear localization of AmyR which is required to activate amylolytic gene expression, is inducer-dependent. Interestingly, deletion of the MH4 domain results in inducer independent localization of GFP-AmyR and constitutive expression of amylolytic genes (Makita *et al.* 2009). The physiological role of the MH1 domain is currently unknown. The MH2 domain is required for the transcriptional activation, as truncation of the C-terminal half from the MH2 domain onwards leads to a defect in transactivation of *taaG2* expression, while it does not affect nuclear localization. Simultaneous deletion of MH3 and the MH4 domains also leads to reduced transactivation activity (Makita *et al.* 2009).

AmyR localizes to the nucleus in response to various inducers which include isomaltose, maltose, kojibiose and pentose in both *A. oryzae* and *A. nidulans* (Kato *et al.* 2002). These inducers are α-linked glucobioses and glucotrioses. In *A. nidulans* isomaltose has the strongest inducing activity and is proposed to be the physiological inducer (Kato *et al.* 2002). Studies in *A. oryzae* and *A. niger* have shown that D-glucose also acts an inducer of α-amylase

production in a AmyR-dependent manner (Carlsen and Nielsen 2001; Murakoshi *et al.* 2012; Vankuyk *et al.* 2012). Whether glucose is directly inducing or whether the glucose needs to be converted by intracellular (trans)glycosylation or glycosyltransferase reactions is currently not known.

Table 2. Functionally characterized transcription factors related to plant biomass degradation in *Aspergilli.*.

	Substrate specificity of TF		Functionally characterized TFs in three *Aspergillus* species			Three-species synteny
TF	**polymer**	**di/monomer**	***A. niger***	***A. oryzae***	***A. nidulans***	
AmyR	starch	glucose	(Yuan *et al.* 2008b; Vankuyk *et al.* 2012)	(Petersen *et al.* 1999; Gomi *et al.* 2000)	(Tani *et al.* 2001; Makita *et al.* 2009; Kojima *et al.* 2010)	alpha-glucosidase & alpha-amylase gene
MalR	starch	maltose	not present (Xiong *et al.* 2017)	(Hasegawa *et al.* 2010; Suzuki *et al.* 2015)	not present (Xiong *et al.* 2017)	NR
InuR	inulin/sucrose	glucose/fructose	(Yuan *et al.* 2006; Yuan *et al.* 2008a)	not analyzed	not analyzed	sugar transporter & invertase gene
XlnR	xylan	xylose	(Van Peij *et al.* 1998b; Klaubauf *et al.* 2014)	(Noguchi *et al.* 2009)	(Klaubauf *et al.* 2014; Kowalczyk *et al.* 2015)	xylan/xylose unrelated genes
ClrB/ManR	cellulose/xylan mannan	cellobiose mannose	(Raulo *et al.* 2016) not analyzed	(Ogawa *et al.* 2013) (Ogawa *et al.* 2012; Ogawa *et al.* 2013)	(Coradetti *et al.* 2012) not analyzed	mannosidase & mannosyl-transferase gene
ClrA	cellulose/xylan	cellobiose	(Raulo *et al.* 2016)	not analyzed	(Coradetti *et al.* 2012)	cellulose/xylan unrelated genes
GalX	galactan	galactose	(Gruben *et al.* 2012)	not analyzed	(Christensen *et al.* 2011)	No
GalR	galactan	galactose	not present (Christensen *et al.* 2011)	not present (Christensen *et al.* 2011)	(Christensen *et al.* 2011; Kowalczyk *et al.* 2015)	NR
AraR	arabinan	arabinose	(Battaglia *et al.* 2011; Battaglia *et al.* 2014)	not analyzed	(Kowalczyk *et al.* 2015)	No
RhaR	rhamnan	rhamnose	(Gruben *et al.* 2014)	not analyzed	(Pardo and Orejas 2014)	rhamnose catabolic pathway genes
GaaR	polygalacturonic acid	galacturonic acid	(Alazi *et al.* 2016; Niu *et al.* 2017)	not analyzed	not analyzed	GA repressor *gaaX*
GaaX	polygalacturonic acid	galacturonic acid	(Niu *et al.* 2017)	not analyzed	not analyzed	GA activator *gaaR*

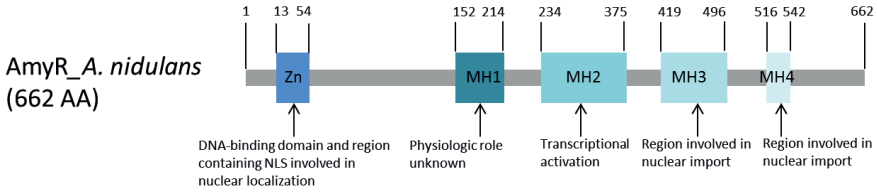


Figure1. Schematic diagram of the AmyR protein domains in *A. nidulans*. This figure was adapted from (Suzuki *et al.* 2015)

In *A. oryzae*, two Zn(II)2Cys6 transcription factors, AmyR and MalR are involved in the regulation of amylolytic enzymes. MalR is the ortholog of yeast maltose utilizing (MAL) activator (Hasegawa *et al.* 2010). Similar to the yeast MAL activator, the *malR* gene in *A. oryzae* is part of a small cluster together with genes encoding putative maltose permease (MalP) and maltase (MalT). MalR controls the expression of maltose-utilizing (MAL) cluster genes independent of AmyR (Hasegawa *et al.* 2010). The activation of AmyR and MalR is regulated in a different manner and illustrates well that even related transcription factor can be activated by different mechanisms. AmyR translocates from cytoplasm to nucleus under the induction of glucose, maltose or isomaltose, and subsequently triggers the expression of amylase genes (Suzuki *et al.* 2015). Different to AmyR, MalR is constitutively localized in nucleus and the expression of MAL cluster genes was induced by maltose, but not by glucose or isomaltose. Amino acid sequences analysis shows that MalR is homologous to AmyR, but MalR seems to lack the MH4 domain which could explain the constitutive nuclear localization of MalR (Suzuki *et al.* 2015). Deletion of *malR* indicates that MalR is essential for maltose utilization, and evidence was provided that MalR is essential for the activation of AmyR (Suzuki *et al.* 2015). As a working model, these authors suggest that the transport of maltose into the cell is mediated via MalR controlled expression of *malT*. In the cell, the maltose is converted into isomaltose via transglycosylation mediated by intracellular alpha-glucosidases. The isomaltose subsequently triggers the activation and translocation of AmyR into the nucleus. This model explains why the preceding activation of MalR is essential for the utilization of maltose as an inducer for AmyR activation (Suzuki *et al.* 2015). *A. niger* and *A. nidulans* do not have orthologs of MalR (Niu, unpublished).

Regulation of amylolytic gene expression is not only regulated via the AmyR mediated activation mechanism, the expression of these genes are also controlled by carbon catabolite repression mechanisms. The carbon catabolite repressor protein CreA has been shown to bind to the promoter sequence of amylolytic genes in the presence of glucose in order to repress transcription (Tsukagoshi *et al.* 2001). Binding affinities experiments of the recombinant CreA protein produced in *E. coli* suggested two CreA binding sites at around -145 to -150 and -90 to -95 within the promoter region of the Taka-amylase A gene (*taaG2*) to be involved in glucose repression (Kato *et al.* 1996).

In *A. niger*, AmyR has been shown to regulate the expression of genes encoding α-amylases, α-glucosidases and glucoamylases (Yuan *et al.* 2008b). By using deletion and multicopy strains, a later study in *A. niger* showed that AmyR also regulates the expression of additional genes encoding α- and β-glucosidases, and α- and β- galactosidases (Vankuyk *et al.* 2012). When grown on D-glucose, lactose, maltose and starch, the activities of α- and β-glucosidases and α- and β- galactosidases were lower in the *amyR* deletion strain and higher in the *amyR* multicopy strain. Consistent with these enzyme activity assays, gene expression analysis suggests that AmyR controls a small subset of genes encoding two β-glucosidases, two α-amylases, two α-glucosidases, two glucoamylases, two α-galactosidases, and one β-galactosidases. These results were further supported by growth profiling, which showed reduced growth on starch, maltose, melibiose, melezitose, raffinose, sucrose) and β-linked D-glucose (cellobiose) as well as α- (melibiose, raffinose, carrageenan) and β-linked D-galactose (lactose, carrageenan) for the *amyR* deletion strain, while improved growth on several of these substrates for the *amyR* multicopy strain (Vankuyk *et al.* 2012). Together, these results indicate that AmyR has a broader physiological role not only in starch degradation but also in regulation of the production of enzymes not directly related to starch. AmyR is commonly suggested to be induced by maltose. However, *A. niger* secretes high levels of glucoamylase when exposed to maltose or starch (Barton *et al.* 1972; Schrickx *et al.* 1995; Gouka *et al.* 1997b; Gouka *et al.* 1997a; Pedersen *et al.* 2000), leading to high glucose levels in the medium. It is suggested that all maltose is hydrolyzed extracellularly to D-glucose, which still is a condition to activate AmyR. Therefore, D-glucose or a metabolic product rather than maltose may be the inducer of the AmyR system in *A. niger*. This explanation is supported by the induction of AmyR regulated genes during growth on low levels of D-glucose (Vankuyk *et al.* 2012).

4.3 XlnR

The expression of genes encoding xylan degrading enzymes is under the control of a transcriptional activator XlnR. The *xlnR* gene was first isolated and characterized by complementation of an *A. niger* mutant lacking xylanolytic activity (Van Peij *et al.* 1998b). The XlnR protein is 875 amino acids long and regulates both xylanolytic and endo-glucanases gene expression in *A. niger* (Van Peij *et al.* 1998a; Van Peij *et al.* 1998b). Later, researchers found that pentose catabolic genes and are also under regulation of XlnR in *A. niger* (Battaglia *et al.* 2014). XlnR is highly conserved among Aspergilli. In *A. oryzae*, the homolog of *A. niger* XlnR is 971 amino acids long and shows 77.5% identity to the *A. niger* XlnR (Marui *et al.* 2002b). The *A. oryzae* XlnR was also found to control expression of xylanolytic and cellulolytic genes (Marui *et al.* 2002a; Marui *et al.* 2002b; Noguchi *et al.* 2009). *A. nidulans* XlnR is a 875 amino acids long protein and shows 76% identity to the *A. niger* XlnR, 73% identity to the *A. oryzae* XlnR (Tamayo *et al.* 2008).

The XlnR contains a Zn(II)2Cys6 domain in the N-terminal region responsible for DNA-binding (Van Peij *et al.* 1998a). The XlnR-binding site (5'-GGCTAAA-3') was identified by electrophoretic mobility shift assays (EMSA) in *A. niger* (Van Peij *et al.* 1998a). The 5'-GGCTAG-3' sequence was found to be functional in the α-glucuronidase gene (*aguA*) promoter (De Vries *et al.* 2002). In *A. oryzae*, two XlnR-binding sites (5'-GGCTAA-3' and 5'-GGCTGA-3') were shown to be functional (Marui *et al.* 2002a). Prediction using *in silico* analysis suggested a putative coiled-coil domain directly C-terminal to the DNA-binding domain, and a second coiled-coil domain at the C-terminal part of XlnR (Hasper *et al.* 2004). The function of each part of XlnR was studied by cellular localization studies using C-terminal GFP-tagged XlnR and xylanase activity of various truncated versions of XlnR (Hasper *et al.* 2004). When GFP was fused with full-length wild-type XlnR, the XlnR-GFP was translocated to the nucleus after the strain was grown on 10 mM D-xylose for 24h (Hasper *et al.* 2004). A stop codon mutation between the DNA-binding sites (amino acid numbers 51 to 86) and the putative basic cluster nuclear localization signal (amino acid numbers 87 to 90) at position 83 (Tyr83stop) of XlnR abolished xylanase activity and caused cytoplasmic localization of XlnR. A stop codon mutation located upstream of the predicated C-terminal coiled-coil region (Asp635stop) within XlnR also resulted in cytosolic localization of XlnR-GFP and a low GFP signal in the nucleus. These results indicate that also the C-terminal coiled-coil domain is involved in the nuclear import of XlnR (Hasper *et al.* 2004). A stop codon mutation located downstream of the predicated C-terminal coiled-coil region

(Leu668stop) within XlnR resulted in nuclear localization of XlnR as in the wild-type and strongly increased xylanase activity compared to the wild-type under non-inducing conditions, indicating that this region is important to support XlnR activity when no inducer is present. Deletion of the putative coiled-coil region (Δ636-666) of XlnR resulted in a total loss of xylanase activity, Moreover, the Δ636-666 mutant and two other mutants containing mutation in the coiled-coil region (Leu650Pro and Tyr664stop) showed no fluorescence signal, suggesting that the XlnR protein may be rapidly degraded (Hasper *et al.* 2004). These results indicate the C-terminal coiled-coil domain is important as an activation domain and important for protein stability.

Finally, in the same study by (Hasper *et al.* 2004), it was found a mutation at position 756 (Val756Phe) in XlnR, as well as a missense mutation at position 668 (Leu668stop), resulted in increased xylanase activity under non-inducing conditions. Deletion of the last 78 amino acids from the C-terminus by introducing a stop codon at position Gly797 resulted in increased xylanase activity com=pared to the wild-type under inducing conditions (Figure 2). Because mutations downstream of the C-terminal coiled-coil region increased the expression of xylanases, it was suggested that this region inhibits XlnR activity under non-inducing condition and that certain mutations or deletion in this region leads to constitutive activation of XlnR. However, two other mutations in this region (Leu823Ser and Tyr864Asp) and deletion of amino acids 802-836 within the last 60 amino acids of XlnR resulted in complete loss of xylanase activity on D-xylose as in the Δ*xlnR* mutant, indicating this region also contains an activation domain (Figure 2).

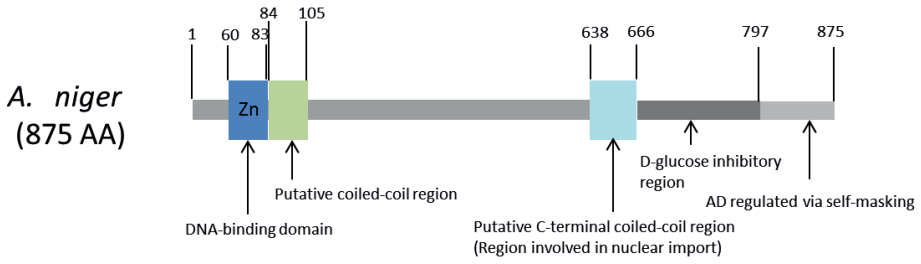


Figure 2. Schematic diagram of the XlnR protein domains in *A. niger*. This figure was adapted from Hasper et al., 2004. Zn domain was analyzed manually. The two coiled-coil regions were predicted on line by website: http://www.bioinformatics.nl/cgi-bin/emboss/pepcoil.

The exact mechanism by which XlnR is activated is currently unknown. Overexpression of XlnR in *A. nidulans* by fusing *xlnR* to the strong promoter *gpdA* from *A. nidulans* did not result in expression of XlnR target genes (*xlnA*, *xlnB* and *xlnD)* under non-inducing condition (Tamayo *et al.* 2008). These results indicate that the presence of xylose is required for XlnR and additional post-transcriptional modifications of XlnR are required for activation. Regulation of XlnR via phosphorylation has been postulated as an important mechanism to control XlnR activity in *A. oryzae* (Noguchi *et al.* 2011). The *A. oryzae* XlnR was found to be present as a mixture of variously phosphorylated forms in the absence of D-xylose, and D-xylose triggered additional phosphorylation (Noguchi *et al.* 2011). However, it is still unclear where these phosphorylation sites are located within XlnR and which proteins mediate the hyperphosphorylation of XlnR.

The expression of xylanolytic genes is not only controlled by induction via XlnR, but also by repression via the carbon catabolite repressor CreA on xylose (De Vries *et al.* 1999; Mach-Aigner *et al.* 2012). CreA controls not only the expression of xylanase genes but also the expression of *xlnR* itself in the presence of glucose (Tamayo et al., 2008). CreA indirectly represses *xlnA* and *xlnB* genes via repression of *xlnR* as well as exerting direct repression on *xlnA and xlnD* expression (Tamayo *et al.* 2008). The repression mechanism of the *xlnA* and *xlnD* genes is different; whereas glucose repression of *xlnA* is mostly repressed indirectly, repression of *xlnD* is mediated via direct repression of CreA by binding to the *xlnD* promoter region (Tamayo *et al.* 2008). This could be explained by the observation that there are three more CreA binding sites within the promoter region of *xlnD* than that of *xlnA*.

Hemicellulose are heterogeneous polysaccharides including xylans, xyloglucans, galactoglucomannan, and arabinogalactan. Full and efficient degradation of hemicellulose requires coordinated action of several transcription factors. In *A. nidulans*, XlnR acts together with AraR and GalR to in regulating genes expression involved in efficient degradation of complex hemicelluloses (Kowalczyk *et al.* 2015).

4.1.3 CreA

In their natural environment, microorganisms select the most energetically favorable carbon source and simultaneously repress the use of less favorable carbon sources. This process is known as carbon catabolite repression (CCR), which supports rapid growth and development required for colonizing diverse habitats (for reviews on carbon repression in fungi see (Kelly 1994; Scazzocchio *et al.* 1995; Ruijter and Visser 1997). CCR is mediated by the transcription

factor CreA, a C2H2 zinc finger DNA-binding protein which was first identified and characterized in *A. nidulans* (Dowzer and Kelly 1989; Dowzer and Kelly 1991). Thereafter CreA was also identified in *A. niger* by screening of mutants relieved of carbon repression (Ruijter *et al.* 1997; Ruijter and Visser 1997). Subsequent studies showed that CreA acts as a repressor to control the expression of genes encoding enzymes required for degradation of different carbon sources including starch, xylan and pectin (De Vries *et al.* 2002; Tamayo *et al.* 2008; Ichinose *et al.* 2014).

CreA has been studied extensively in *A. nidulans*. The consensus binding motif of CreA is 5' SYGGRG (Kulmburg *et al.* 1993; Cubero and Scazzocchio 1994; Cubero *et al.* 2000). By analysing the conserved domains and special features of the amino acid sequence, five different regions in CreA have been defined (Figure 3) (Roy *et al.* 2008). The two C2H2 zinc fingers DNA-binding domains (region I) are followed by a conserved region containing seven alanine residues (region II). Region III contains an acidic acid–rich region that is located adjacent to a highly conserved region (region IV). Region IV is a highly conserved between *A. niger* and *T. reesei* (Roy *et al.* 2008). This conserved region is followed by region V which has been shown to be important for repression (Roy *et al.* 2008). Except for the two C2H2 zinc fingers region, the function of each region was studied by individual deletion of CreA regions and studying the functionality of CreA and by analyzing the localization of CreA using CreA-GFP fusion proteins (Roy *et al.* 2008; Ries *et al.* 2016). Roy et al., found that the C_2H_2 domain and the C-terminal repression domain (region V) of CreA are required for repressing function of CreA (Roy *et al.* 2008).

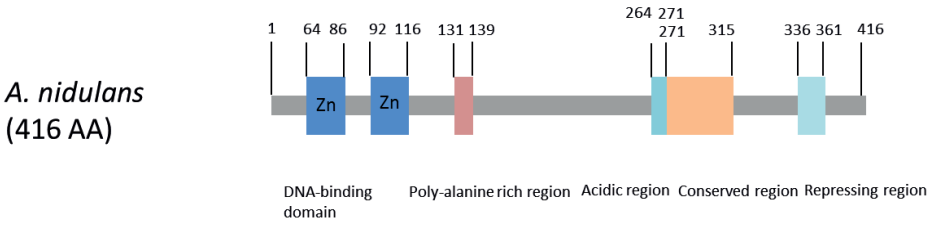


Figure 3. Schematic diagram of different regions in CreA (Roy *et al.* 2008).

Western blot analysis of HA:CreA:GFP showed that the expression levels of CreA were similar in mycelia grown in repressing or depressing conditions, indicating that there was little or no transcriptional/translation control on CreA. Similarly, constitutive over-expression of CreA did not affect normal repression or depression, indicating carbon catabolite

repression was independent on the transcription level of *creA* (Roy *et al.* 2008). Analysis of the cellular localization of CreA under repressing and non-repressing conditions revealed that there was not strictly correlation between CreA localization and it is activity. In a range of glucose concentrations (from 1% to 0.01%), GFP-CreA was mainly localized in the nucleus. However, fluorescence was not exclusively in the nucleus as GFP was also present in the cytoplasm (Roy *et al.* 2008).

Ubiquitination/deubiquitination processes are important components of gene post-translational regulation mechanisms in eukaryotes. Once a protein is identified for degradation, it will be marked for degradation by the attachment of a ubiquitin moiety, a small regulatory protein that has been found in almost all eukaryotes (Lecker *et al.* 2006). The ubiquitinated protein will be delivered to proteasome for degradation. There is evidence that deubiquitination (DUB) carried out by the CreB-CreC DUB complex plays a role in CCR, and it was suggested that CreA was deubiquitinylated by the CreB-CreC complex (Hynes and Kelly 1977; Lockington and Kelly 2002).

CreB is the deubiquitinating enzyme (Lockington and Kelly 2002) and CreC is the scaffold protein containing WD-40 repeats (Todd *et al.* 2000). The genetic evidence for this is that mutants in *creB* or *creC* are, like *creA* mutants, unable to perform CCR. Thus, it is expected that in *creB* and *creC* mutants, CreA protein is always ubiquinated (no deubiquitination) resulting in low levels of CreA. In addition, the interaction of CreD with a ubiquitin protein ligase HulA may also be involved in the ubiquitination process (Boase and Kelly 2004). Kelly proposed a model that in derepressing conditions, CreA or a protein that acts in a complex with CreA might be ubiquitinated by CreD/HulA complex. This action would target CreA to the proteasome and thus prevent repression. In the presence of a repressing carbon source however, the repressing activity of CreA or a CreA-complex might be restored by deubiquitination that is mediated by CreB and CreC (Kelly 2004). Results of co-immunoprecipitation using CreA and CreB antibodies, indicated that CreA and CreB are not present in the same complex (Alam *et al.* 2016). Studies with antibodies against phosphorylated proteins (Phos-tag system) as well as antibodies against unbiquitin, showed that CreA is a phosphorylated protein, but not an ubiquinated protein. This conclusion was further supported by mass spectrometry and indicates that CreA is not a direct target of CreB (Alam *et al.* 2016). Alam and Kelly (Alam and Kelly 2016) further conducted experiments to identify possible proteins that may be of a CreA/CreB complex and thereby form a bridge between CreA and CreB under repressing and derepressing conditions, and they found Hir3 to

be present in both repressing and derepressing conditions for CreB, suggesting that Hir3, or proteins interacting with Hir3, could be a possible target of CreB. Further research and the possible role of CreC as a scaffold protein is needed to illuminate the regulatory mechanism involved in CCR.

4.4 State of the art of understanding research on the regulation of pectinolytic genes at the start of this thesis

At the start of my thesis, it was well established that the expression of polygalacturonic or pectin degrading enzymes was highly regulated. Early studies in *A. nidulans* performed by Dean and Timberlake showed that the *pelA* mRNA (encoding a pectate lyase) was detectable on polygalacturonic acid as carbon source and undetectable on glucose or acetate as carbon source (Dean and Timberlake 1989). Similarly, it was shown in *A. niger* that several pectinases are specifically induced on GA or pectin (De Vries *et al.* 2002). In addition, it was found that the expression of polygalacturonase and pectate lyase in *A. nidulans* is completely repressed by glucose due to carbon catabolite repression (Dean and Timberlake 1989; De Vries *et al.* 2002), suggesting that the expression of enzymes involved in pectin/polygalacturonic acid degradation is also under the carbon catabolite repression. With the possibility of performing genome-wide expression studies using Affymetrix gene arrays, it was shown that at least 11 genes were specifically induced on galacturonic acid (Martens-Uzunova and Schaap 2008). Analysis of the promoter region of these 11 genes identified a conserved promoter element "YCCNCCAAT" (Martens-Uzunova and Schaap 2008) which was suggested to play an important role in the regulation and co-expression of these genes. These results indicate that the expression of genes encoding enzymes involved in pectin degradation is specifically induced by polygalacturonic acid or galacturonic acid. Taken together, genes encoding enzymes involved in pectin/polygalacturonic acid degradation are specifically induced on polygalacturonic acid or galacturonic acid and are under carbon catabolite repression control.

5. Aim and outline of the thesis

A. niger is an important industrial enzyme producer. These enzymes find their way in a broad spectrum of industrial applications in food and non-food products or processes. Highly efficient production of enzymes mediated by modulating transcriptional regulation is meaningful. In this thesis I focus on the complex regulation of the expression of pectinolytic genes in *A. niger*.

The co-regulation of pectin degrading enzymes and the conserved promoter element in the co-regulated genes (Martens-Uzunova and Schaap 2008) strongly suggested the existence of a transcriptional activator coordinating the activation of gene expression of these GA-induced genes in response to GA or pectin. Whereas over the last few years several new transcription factors involved in plant cell wall degradation have been identified (Kowalczyk *et al.* 2014), a possible transcription factor involved in the regulation of pectinases was not identified.

With the start of the project, I first developed two important new tools for functional genomics in *A. niger*. The first tool is an efficient system for making gene deletion mutants by combining non-homologous end joining (NHEJ) mutants and a split marker approach (Chapter 2). This system was used to make (in collaboration with DSM) a library of 240 *A. niger* transcription factor mutants (Arentshorst, Arendsen, van Peij, Pel, Ram, unpublished data). Unfortunately, screening of this collection did not yield a mutant with a specific growth defect on GA indicating that the specific transcription factor mutant was not present in the collection.

The second important new tool was to construct auxotrophic mutants in which multiple gene deletions could be made easily as we anticipated that maybe multiple transcription factor encoding genes should be deleted to obtain a pectin-non-utilizing mutant. With the aim to combine whole genome sequencing of mutants in combination with parasexual crosses to facility mutant identification via next generation sequencing techniques (Ram 2013), isogenic, auxotrophic colour mutants were constructed via targeted deletion approaches. Genome sequencing of two auxotrophic colour mutants showed a high level of isogenicity between them, which could facilitate the selection of diploids and the isolation of haploid segregants from the diploid using the parasexual cycle (Chapter 3).

Since the targeted approach by constructing gene deletion mutants in selected transcription factors was not successful, a non-targeted approach was designed for a forward genetic screen to isolate mutants with constitutive expression of pectinases. First of all, I selected promoter region of five genes (*pgaX*, *pgxB*, *pgxC*, *gatA* and *gaaB*) that were specifically induced by GA based on available genome-wide expression profiles from literature to construct promoter-*amdS* reporter strains. These reporter strains were used to analyze gene expression *in vivo* by assaying the growth of these strains on acetamide. The rationale of the screen is that high expression of the *amdS* gene allows the fungal strain to grow on acetamide as the nitrogen source and as such the ability to grow on acetamide is a direct measurement of promoter

activity. Growth analysis of the reporter strains indicated the promoter regions of first four genes (*pgaX*, *pgxB*, *pgxC*, *gatA)* were specifically induced on GA even under the carbon catabolite derepression condition, allowing forward genetic screens for inducer-independent mutants (Chapter 4 and Chapter 6).

In Chapter 5, I identified the transcriptional activator required for GA-induced gene expression (GaaR) by homology to BcGaaR (Zhang *et al.* 2016). Growth phenotype and genome-wide expression analysis of the *A. niger* Δ*gaaR* strain showed that GaaR is required for the expression of genes involved in releasing GA from PGA and more complex pectins, in transporting GA into the cell, and in inducing the GA catabolic pathway.

Subsequently, the reporter strain containing the P*pgaX-amdS* reporter construct (Chapter 4) was selected to screen for inducer-independent mutants which constitutively expressed pectinases. Whole genome sequencing of five constitutive mutants revealed the gene NRRL3_08194, named *gaaX*, that was responsible for the constitutive expression of pectinases when deleted (Chapter 6). *GaaX* is located next to *gaaR* (NRRL3_08195) in the genome. In this Chapter, we provide the first evidence that *gaaX* is likely to encode a repressor protein that controls the activity of GaaR and keeps GaaR inactive under non-inducing conditions. In Chapter 7, I summarize that main findings obtained in the thesis, discuss the current working model, and propose some future experiments to further understand the molecular details on how the repressor (GaaX) and the activator (GaaR) and the inducer molecule (2-keto-3-deoxy galactonate) (Alazi *et al*. 2017) might interact to control GA-specific gene expression.

References

Alam, M. A., N. Kamlangdee and J. M. Kelly, 2016 The CreB deubiquitinating enzyme does not directly target the CreA repressor protein in *Aspergillus nidulans*. Curr Genet.

Alam, M. A., and J. M. Kelly, 2016 Proteins interacting with CreA and CreB in the carbon catabolite repression network in *Aspergillus nidulans*. Curr Genet.

Alazi, E., J. Niu, J. E. Kowalczyk, M. Peng, M. V. Aguilar Pontes *et al.*, 2016 The transcriptional activator GaaR of *Aspergillus niger* is required for release and utilization of d-galacturonic acid from pectin. FEBS Lett 590**:** 1804-1815.

Amici, A. M., A. Minghetti, T. Scotti, C. Spalla and L. Tognoli, 1969 Production of peptide ergot alkaloids in submerged culture by three isolates of *Claviceps purpurea*. Appl Microbiol 18**:** 464-468.

Andersen, M. R., M. P. Salazar, P. J. Schaap, P. J. van de Vondervoort, D. Culley *et al.*, 2011 Comparative genomics of citric-acid-producing *Aspergillus niger* ATCC 1015 versus enzyme-producing CBS 513.88. Genome Res 21**:** 885-897.

Andersen, M. R., W. Vongsangnak, G. Panagiotou, M. P. Salazar, L. Lehmann *et al.*, 2008 A trispecies *Aspergillus* microarray: comparative transcriptomics of three Aspergillus species. Proc Natl Acad Sci U S A 105**:** 4387-4392.

Anderson, P., 2010 Post-transcriptional regulons coordinate the initiation and resolution of inflammation. Nat Rev Immunol 10**:** 24-35.

Arnaud, M. B., G. C. Cerqueira, D. O. Inglis, M. S. Skrzypek, J. Binkley *et al.*, 2012 The *Aspergillus* Genome Database (AspGD): recent developments in comprehensive multispecies curation, comparative genomics and community resources. Nucleic Acids Res 40**:** D653-659.

Ashwell, G., A. J. Wahba and J. Hickman, 1960 Uronic acid metabolism in bacteria. I. Purification and properties of uronic acid isomerase in *Escherichia coli*. J Biol Chem 235**:** 1559-1565.

Azhar, S. H. M., R. Abdulla, S. A. Jambo, H. Marbawi, J. A. Gansau *et al.*, 2017 Yeasts in sustainable bioethanol production: A review. Biochemistry and Biophysics Reports 10**:** 9.

Baker, S. E., 2006 *Aspergillus niger* genomics: past, present and into the future. Med Mycol 44 Suppl 1**:** S17-21.

Barton, L. L., C. E. Georgi and D. R. Lineback, 1972 Effect of maltose on glucoamylase formation by *Aspergillus niger*. J Bacteriol 111**:** 771-777.

Batra, L. R., and P. D. Millner, 1974 Some Asian Fermented Foods and Beverages, and Associated Fungi. Mycologia 66**:** 942-950.

Battaglia, E., S. F. Hansen, A. Leendertse, S. Madrid, H. Mulder *et al.*, 2011 Regulation of pentose utilisation by AraR, but not XlnR, differs in *Aspergillus nidulans* and *Aspergillus niger*. Appl Microbiol Biotechnol 91**:** 387-397.

Battaglia, E., M. Zhou and R. P. de Vries, 2014 The transcriptional activators AraR and XlnR from *Aspergillus niger* regulate expression of pentose catabolic and pentose phosphate pathway genes. Res Microbiol 165**:** 531-540.

Beck, C. F., R. Mutzel, J. Barbe and W. Muller, 1982 A multifunctional gene (tetR) controls Tn10-encoded tetracycline resistance. J Bacteriol 150**:** 633-642.

Benocci, T., M. V. Aguilar-Pontes, M. Zhou, B. Seiboth and R. P. de Vries, 2017 Regulators of plant biomass degradation in ascomycetous fungi. Biotechnol Biofuels 10**:** 152.

Benz, J. P., R. J. Protzko, J. M. S. Andrich, S. Bauer, J. E. Dueber *et al.*, 2014 Identification and characterization of a galacturonic acid transporter from *Neurospora crassa* and its application for Saccharomyces cerevisiae fermentation processes. Biotechnology for Biofuels 7.

Biely, P., S. Singh and V. Puchart, 2016 Towards enzymatic breakdown of complex plant xylan structures: State of the art. Biotechnol Adv 34**:** 1260-1274.

Boase, N. A., and J. M. Kelly, 2004 A role for creD, a carbon catabolite repression gene from *Aspergillus nidulans*, in ubiquitination. Mol Microbiol 53**:** 929-940.

Carlsen, M., and J. Nielsen, 2001 Influence of carbon source on alpha-amylase production by *Aspergillus oryzae*. Appl Microbiol Biotechnol 57**:** 346-349.

Chang, P. K., and K. C. Ehrlich, 2013 Genome-wide analysis of the Zn(II)(2)Cys(6) zinc cluster-encoding gene family in *Aspergillus flavus*. Appl Microbiol Biotechnol 97**:** 4289-4300.

Chang, S. T., and W. A. Hayes, 1978 *The Biology and Cultivation of Edible Mushroom*. London & New York: Academic Press.

Chavez, R., P. Bull and J. Eyzaguirre, 2006 The xylanolytic enzyme system from the genus Penicillium. J Biotechnol 123**:** 413-433.

Christensen, U., B. S. Gruben, S. Madrid, H. Mulder, I. Nikolaev *et al.*, 2011 Unique regulatory mechanism for D-galactose utilization in *Aspergillus nidulans*. Appl Environ Microbiol 77**:** 7084-7087.

Chutmanop, J., S. Chuichulcherm, Y. Chisti and P. Sirinophakun, 2008 Protease production by *Aspergillus oryzae* in solid-state fermentation using agroindustrial substrates. Journal of Chemical Technology and Biotechnology 83**:** 1012-1018.

Coradetti, S. T., J. P. Craig, Y. Xiong, T. Shock, C. Tian *et al.*, 2012 Conserved and essential transcription factors for cellulase gene expression in ascomycete fungi. Proc Natl Acad Sci U S A 109**:** 7397-7402.

Coutinho, P. M., M. R. Andersen, K. Kolenova, P. A. vanKuyk, I. Benoit *et al.*, 2009 Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. Fungal Genetics and Biology 46**:** S161-S169.

Coutinho, P. M., and B. Henrissat, 1999 Carbohydrate-active enzymes: an integrated database approach, pp. 3-12 in *Recent advances in Carbohydrate Bioengineering*, edited by H. J. Gilbert, G. D'avies, B. Henrissat and B. Svensson. Cambridge: Royal Society of Chemistry.

Craig, J. P., S. T. Coradetti, T. L. Starr and N. L. Glass, 2015 Direct target network of the *Neurospora crassa* plant cell wall deconstruction regulators CLR-1, CLR-2, and XLR-1. MBio 6**:** e01452-01415.

Cubero, B., D. Gomez and C. Scazzocchio, 2000 Metabolite repression and inducer exclusion in the proline utilization gene cluster of *Aspergillus nidulans*. J Bacteriol 182**:** 233-235.

Cubero, B., and C. Scazzocchio, 1994 Two different, adjacent and divergent zinc finger binding sites are necessary for CREA-mediated carbon catabolite repression in the proline gene cluster of *Aspergillus nidulans*. EMBO J 13**:** 407-415.

Damveld, R. A., A. Franken, M. Arentshorst, P. J. Punt, F. M. Klis *et al.*, 2008 A novel screening method for cell wall mutants in *Aspergillus niger* identifies UDP-galactopyranose mutase as an important protein in fungal cell wall biosynthesis. Genetics 178**:** 873-881.

Das, A., and P. Roy, 1978 Improved production of citric acid by a diploid strain of *Aspergillus niger*. Can J Microbiol 24**:** 622-625.

de Vries, D. J., M. A. King, E. J. Soares, B. M. Tsui and C. E. Metz, 1999 Effects of scatter substraction on detection and quantitation in hepatic SPECT. J Nucl Med 40**:** 1011-1023.

de Vries, R. P., R. Riley, A. Wiebenga, G. Aguilar-Osorio, S. Amillis *et al.*, 2017 Comparative genomics reveals high biological diversity and specific adaptations in the industrially and medically important fungal genus *Aspergillus*. Genome Biol 18**:** 28.

de Vries, R. P., P. J. van de Vondervoort, L. Hendriks, M. van de Belt and J. Visser, 2002 Regulation of the alpha-glucuronidase-encoding gene (aguA) from *Aspergillus niger*. Mol Genet Genomics 268**:** 96-102.

de Vries, R. P., and J. Visser, 2001 Aspergillus enzymes involved in degradation of plant cell wall polysaccharides. Microbiol Mol Biol Rev 65**:** 497-522, table of contents.

Dean, R. A., and W. E. Timberlake, 1989 Production of cell wall-degrading enzymes by *Aspergillus nidulans*: a model system for fungal pathogenesis of plants. Plant Cell 1**:** 265-273.

Doudna, J. A., and E. Charpentier, 2014 Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science 346**:** 1258096.

Dowzer, C. E., and J. M. Kelly, 1989 Cloning of the creA gene from *Aspergillus nidulans*: a gene involved in carbon catabolite repression. Curr Genet 15**:** 457-459.

Dowzer, C. E., and J. M. Kelly, 1991 Analysis of the creA gene, a regulator of carbon catabolite repression in *Aspergillus nidulans*. Mol Cell Biol 11**:** 5701-5709.

Eitenmiller, R. R., J. R. Vakil and K. M. Shahani, 1970 Production and Properties of *Pencillium Roqueforti* Lipase. Journal of Food Science 35**:** 130-+.

Fairhead, C., B. Llorente, F. Denis, M. Soler and B. Dujon, 1996 New vectors for combinatorial deletions in yeast chromosomes and for gap-repair cloning using 'split-marker' recombination. Yeast 12**:** 1439-1457.

Fedorova, N. D., N. Khaldi, V. S. Joardar, R. Maiti, P. Amedeo *et al.*, 2008 Genomic islands in the pathogenic filamentous fungus *Aspergillus fumigatus*. PLoS Genet 4**:** e1000046.

Galagan, J. E., S. E. Calvo, K. A. Borkovich, E. U. Selker, N. D. Read *et al.*, 2003 The genome sequence of the filamentous fungus *Neurospora crassa*. Nature 422**:** 859-868.

Galagan, J. E., S. E. Calvo, C. Cuomo, L. J. Ma, J. R. Wortman *et al.*, 2005 Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature 438**:** 1105-1115.

Gilbert, G. F., 2000 *Basic Concepts in Biochemistry*. McGraw-Hill health Professions Division, New York.

Gomi, K., T. Akeno, T. Minetoki, K. Ozeki, C. Kumagai *et al.*, 2000 Molecular cloning and characterization of a transcriptional activator gene, *amyR*, involved in the amylolytic gene expression in *Aspergillus oryzae*. Biosci Biotechnol Biochem 64**:** 816-827.

Goodwin, S., J. D. McPherson and W. R. McCombie, 2016 Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 17**:** 333-351.

Goodwin, T. W., and S. Pendlington, 1954 Studies on the biosynthesis of riboflavin; nitrogen metabolism and flavinogenesis in *Eremothecium ashbyii*. Biochem J 57**:** 631-641.

Goswami, R. S., 2012 Targeted gene replacement in fungi using a split-marker approach. Methods Mol Biol 835**:** 255-269.

Gouka, R. J., P. J. Punt and C. A. van den Hondel, 1997a Efficient production of secreted proteins by *Aspergillus*: progress, limitations and prospects. Appl Microbiol Biotechnol 47**:** 1-11.

Gouka, R. J., P. J. Punt and C. A. van den Hondel, 1997b Glucoamylase gene fusions alleviate limitations for protein production in *Aspergillus awamori* at the transcriptional and (post) translational levels. Appl Environ Microbiol 63**:** 488-497.

Gruben, B. S., M. Zhou and R. P. de Vries, 2012 GalX regulates the D-galactose oxido-reductive pathway in *Aspergillus niger*. FEBS Lett 586**:** 3980-3985.

Gruben, B. S., M. Zhou, A. Wiebenga, J. Ballering, K. M. Overkamp *et al.*, 2014 *Aspergillus niger* RhaR, a regulator involved in L-rhamnose release and catabolism. Appl Microbiol Biotechnol 98**:** 5531-5540.

Harholt, J., A. Suttangkakul and H. Vibe Scheller, 2010 Biosynthesis of pectin. Plant Physiol 153**:** 384-395.

Hasegawa, S., M. Takizawa, H. Suyama, T. Shintani and K. Gomi, 2010 Characterization and expression analysis of a maltose-utilizing (MAL) cluster in *Aspergillus oryzae*. Fungal Genet Biol 47**:** 1-9.

Hasper, A. A., L. M. Trindade, D. van der Veen, A. J. van Ooyen and L. H. de Graaff, 2004 Functional analysis of the transcriptional activator XlnR from *Aspergillus niger*. Microbiology 150**:** 1367-1375.

Hawksworth, D. L., and B. E. Kirsop, 1988 *Filamentous Fungi. Living Resources for Biotechnology*. Cambridge University Press, Cambridge.

Heald, P. J., and B. Kristiansen, 1985 Synthesis of polysaccharide by yeast-like forms of Aureobasidium pullulans. Biotechnol Bioeng 27**:** 1516-1519.

Heather, J. M., and B. Chain, 2016 The sequence of sequencers: The history of sequencing DNA. Genomics 107**:** 1-8.

Hesseltine, C. W., 1965 A millenium of fungi, food and fermentation. Mycologia 57**:** 149-197.

Hilditch, S., S. Berghall, N. Kalkkinen, M. Penttila and P. Richard, 2007 The missing link in the fungal D-galacturonate pathway - Identification of the L-threo-3-deoxy-hexulosonate aldolase. Journal of Biological Chemistry 282**:** 26195-26201.

Horvath, P., and R. Barrangou, 2010 CRISPR/Cas, the immune system of bacteria and archaea. Science 327**:** 167-170.

Houbraken, J., J. C. Frisvad and R. A. Samson, 2011 Fleming's penicillin producing strain is not *Penicillium chrysogenum* but *P. rubens*. IMA Fungus 2**:** 87-95.

Hsiao, Y. M., M. H. Zheng, R. M. Hu, T. C. Yang and Y. H. Tseng, 2008 Regulation of the *pehA* gene encoding the major polygalacturonase of Xanthomonas campestris by Clp and RpfF. Microbiology 154**:** 705-713.

Hsu, P. D., E. S. Lander and F. Zhang, 2014 Development and applications of CRISPR-Cas9 for genome engineering. Cell 157**:** 1262-1278.

Huang, H. L., C. C. Lee and S. Y. Ho, 2007 Selecting a minimal number of relevant genes from microarray data to design accurate tissue classifiers. Biosystems 90**:** 78-86.

Huisjes, E. H., M. A. Luttik, M. J. Almering, M. M. Bisschops, D. H. Dang *et al.*, 2012 Toward pectin fermentation by *Saccharomyces cerevisiae*: expression of the first two steps of a bacterial pathway for D-galacturonate metabolism. J Biotechnol 162**:** 303-310.

Hynes, M. J., and J. M. Kelly, 1977 Pleiotropic mutants of *Aspergillus nidulans* altered in carbon metabolism. Mol Gen Genet 150**:** 193-204.

Ichinose, S., M. Tanaka, T. Shintani and K. Gomi, 2014 Improved alpha-amylase production by *Aspergillus oryzae* after a double deletion of genes involved in carbon catabolite repression. Appl Microbiol Biotechnol 98**:** 335-343.

Ito, T., S. Tani, T. Itoh, N. Tsukagoshi, M. Kato *et al.*, 2004 Mode of AmyR binding to the CGGN8AGG sequence in the *Aspergillus oryzae* taaG2 promoter. Biosci Biotechnol Biochem 68**:** 1906-1911.

Johnson, D. S., A. Mortazavi, R. M. Myers and B. Wold, 2007 Genome-wide mapping of in vivo protein-DNA interactions. Science 316**:** 1497-1502.

Jorgensen, T. R., B. M. Nitsche, G. E. Lamers, M. Arentshorst, C. A. van den Hondel *et al.*, 2010 Transcriptomic insights into the physiology of *Aspergillus niger* approaching a specific growth rate of zero. Appl Environ Microbiol 76**:** 5344-5355.

Karaffa, L., E. Sandor, E. Fekete and A. Szentirmai, 2001 The biochemistry of citric acid accumulation by *Aspergillus niger*. Acta Microbiol Immunol Hung 48**:** 429-440.

Kato, M., K. Sekine and N. Tsukagoshi, 1996 Sequence-specific binding sites in the Taka-amylase A G2 promoter for the CreA repressor mediating carbon catabolite repression. Biosci Biotechnol Biochem 60**:** 1776-1779.

Kato, N., Y. Murakoshi, M. Kato, T. Kobayashi and N. Tsukagoshi, 2002 Isomaltose formed by alpha-glucosidases triggers amylase induction in *Aspergillus nidulans*. Curr Genet 42**:** 43-50.

Kelly, J. M., 1994 Carbon catabolite repression. Prog Ind Microbiol 29**:** 355-367.

Kelly, J. M., 2004 The regulation of carbon metabolism in filamentous fungi in *Mycota lii: Biochemistry and Molecular Biology*.

Kerstens, H. H. D., 2010 Bioinformatics approaches to detect genetic variation in whole genome sequencing data, pp. Wageningen University.

Klaubauf, S., H. M. Narang, H. Post, M. Zhou, K. Brunner *et al.*, 2014 Similar is not the same: differences in the function of the (hemi-)cellulolytic regulator XlnR (Xlr1/Xyr1) in filamentous fungi. Fungal Genet Biol 72**:** 73-81.

Kojima, T., Y. Hashimoto, M. Kato, T. Kobayashi and H. Nakano, 2010 High-throughput screening of DNA binding sites for transcription factor AmyR from *Aspergillus nidulans* using DNA beads display system. J Biosci Bioeng 109**:** 519-525.

Kooistra, R., P. J. J. Hooykaas and H. Y. Steensma, 2004 Efficient gene targeting in *Kluyveromyces lactis*. Yeast 21**:** 781-792.

Kowalczyk, J. E., I. Benoit and R. P. de Vries, 2014 Regulation of plant biomass utilization in *Aspergillus*. Adv Appl Microbiol 88**:** 31-56.

Kowalczyk, J. E., B. S. Gruben, E. Battaglia, A. Wiebenga, E. Majoor *et al.*, 2015 Genetic Interaction of *Aspergillus nidulans* galR, xlnR and araR in Regulating D-Galactose and L-Arabinose Release and Catabolism Gene Expression. PLoS One 10**:** e0143200.

Krijgsheld, P., B. M. Nitsche, H. Post, A. M. Levin, W. H. Muller *et al.*, 2013 Deletion of flbA results in increased secretome complexity and reduced secretion heterogeneity in colonies of *Aspergillus niger*. J Proteome Res 12**:** 1808-1819.

Kuck, U., and B. Hoff, 2010 New tools for the genetic manipulation of filamentous fungi. Appl Microbiol Biotechnol 86**:** 51-62.

Kuivanen, J., D. Mojzita, Y. M. Wang, S. Hilditch, M. Penttila *et al.*, 2012 Engineering Filamentous Fungi for Conversion of D-Galacturonic Acid to L-Galactonic Acid. Applied and Environmental Microbiology 78**:** 8676-8683.

Kulmburg, P., M. Mathieu, C. Dowzer, J. Kelly and B. Felenbok, 1993 Specific binding sites in the alcR and alcA promoters of the ethanol regulon for the CREA repressor mediating carbon catabolite repression in *Aspergillus nidulans*. Mol Microbiol 7**:** 847-857.

Kuorelahti, S., P. Jouhten, H. Maaheimo, M. Penttila and P. Richard, 2006 L-galactonate dehydratase is part of the fungal path for D-galacturonic acid catabolism. Molecular Microbiology 61**:** 1060-1068.

Kuorelahti, S., N. Kalkkinen, M. Penttila, J. Londesborough and P. Richard, 2005 Identification in the mold *Hypocrea jecorina* of the first fungal D-galacturonic acid reductase. Biochemistry 44**:** 11234-11240.

Kurella, M., L. L. Hsiao, T. Yoshida, J. D. Randall, G. Chow *et al.*, 2001 DNA microarray analysis of complex biologic processes. J Am Soc Nephrol 12**:** 1072-1078.

Laver, T., J. Harrison, P. A. O'Neill, K. Moore, A. Farbos *et al.*, 2015 Assessing the performance of the Oxford Nanopore Technologies MinION. Biomol Detect Quantif 3**:** 1-8.

Lecker, S. H., A. L. Goldberg and W. E. Mitch, 2006 Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. J Am Soc Nephrol 17**:** 1807-1819.

Leijdekkers, A. G., J. H. Huang, E. J. Bakx, H. Gruppen and H. A. Schols, 2015 Identification of novel isomeric pectic oligosaccharides using hydrophilic interaction chromatography coupled to traveling-wave ion mobility mass spectrometry. Carbohydr Res 404**:** 1-8.

Liepins, J., S. Kuorelahti, M. Penttila and P. Richard, 2006 Enzymes for the NADPH-dependent reduction of dihydroxyacetone and D-glyceraldehyde and L-glyceraldehyde in the mould *Hypocrea jecorina*. Febs Journal 273**:** 4229-4235.

Lister, R., B. D. Gregory and J. R. Ecker, 2009 Next is now: new technologies for sequencing of genomes, transcriptomes, and beyond. Curr Opin Plant Biol 12**:** 107-118.

Lockington, R. A., and J. M. Kelly, 2002 The WD40-repeat protein CreC interacts with and stabilizes the deubiquitinating enzyme CreB in vivo in *Aspergillus nidulans*. Mol Microbiol 43**:** 1173-1182.

Mach-Aigner, A. R., J. Omony, B. Jovanovic, A. J. van Boxtel and L. H. de Graaff, 2012 d-Xylose concentration-dependent hydrolase expression profiles and the function of CreA and XlnR in *Aspergillus niger*. Appl Environ Microbiol 78**:** 3145-3155.

Machida, M., K. Asai, M. Sano, T. Tanaka, T. Kumagai *et al.*, 2005 Genome sequencing and analysis of *Aspergillus oryzae*. Nature 438**:** 1157-1161.

Macmillan, J., 1954 Griseofulvin. Part 9. Isolation of the bromoanalogue from *Penicillium griseofulvum* and *Penicillium nigricans*. J. Chem. Soc.**:** 2585-2587.

MacPherson, S., M. Larochelle and B. Turcotte, 2006 A fungal family of transcriptional regulators: the zinc cluster proteins. Microbiol Mol Biol Rev 70**:** 583-604.

Makita, T., Y. Katsuyama, S. Tani, H. Suzuki, N. Kato *et al.*, 2009 Inducer-dependent nuclear localization of a Zn(II)(2)Cys(6) transcriptional activator, AmyR, in *Aspergillus nidulans*. Biosci Biotechnol Biochem 73**:** 391-399.

Mannhaupt, G., C. Montrone, D. Haase, H. W. Mewes, V. Aign *et al.*, 2003 What's in the genome of a filamentous fungus? Analysis of the Neurospora genome sequence. Nucleic Acids Res 31**:** 1944-1954.

Margulies, M., M. Egholm, W. E. Altman, S. Attiya, J. S. Bader *et al.*, 2005 Genome sequencing in microfabricated high-density picolitre reactors. Nature 437**:** 376-380.

Martens-Uzunova, E. S., and P. J. Schaap, 2008 An evolutionary conserved d-galacturonic acid metabolic pathway operates across filamentous fungi capable of pectin degradation. Fungal Genet Biol 45**:** 1449-1457.

Martens-Uzunova, E. S., and P. J. Schaap, 2009 Assessment of the pectin degrading enzyme network of *Aspergillus niger* by functional genomics. Fungal Genet Biol 46 Suppl 1**:** S170-S179.

Marui, J., N. Kitamoto, M. Kato, T. Kobayashi and N. Tsukagoshi, 2002a Transcriptional activator, AoXlnR, mediates cellulose-inductive expression of the xylanolytic and cellulolytic genes in *Aspergillus oryzae*. FEBS Lett 528**:** 279-282.

Marui, J., A. Tanaka, S. Mimura, L. H. de Graaff, J. Visser *et al.*, 2002b A transcriptional activator, AoXlnR, controls the expression of genes encoding xylanolytic enzymes in *Aspergillus oryzae*. Fungal Genet Biol 35**:** 157-169.

Maxam, A. M., and W. Gilbert, 1977 A new method for sequencing DNA. Proc Natl Acad Sci U S A 74**:** 560-564.

Meleigy, S. A., and M. A. Khalaf, 2009 Biosynthesis of gibberellic acid from milk permeate in repeated batch operation by a mutant Fusarium moniliforme cells immobilized on loofa sponge. Bioresour Technol 100**:** 374-379.

Meyer, D. H., and A. M. Bailis, 2008 Mating type influences chromosome loss and replicative senescence in telomerase-deficient budding yeast by Dnl4-dependent telomere fusion. Mol Microbiol 69**:** 1246-1254.

Meyer, V., M. Arentshorst, S. J. Flitter, B. M. Nitsche, M. J. Kwon *et al.*, 2009 Reconstruction of signaling networks regulating fungal morphogenesis by transcriptomics. Eukaryot Cell 8**:** 1677-1691.

Meyer, V., F. Wanka, J. van Gent, M. Arentshorst, C. A. van den Hondel *et al.*, 2011 Fungal gene expression on demand: an inducible, tunable, and metabolism-independent expression system for *Aspergillus niger*. Appl Environ Microbiol 77**:** 2975-2983.

Michelmore, R. W., I. Paran and R. V. Kesseli, 1991 Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci U S A 88**:** 9828-9832.

Mikheyev, A. S., and M. M. Tin, 2014 A first look at the Oxford Nanopore MinION sequencer. Mol Ecol Resour 14**:** 1097-1102.

Mirocha, C. J., and J. E. DeVay, 1971 Growth of fungi on an inorganic medium. Can J Microbiol 17**:** 1373-1378.

Mohnen, D., 2008 Pectin structure and biosynthesis. Curr Opin Plant Biol 11**:** 266-277.

Mojzita, D., M. Wiebe, S. Hilditch, H. Boer, M. Penttila *et al.*, 2010 Metabolic Engineering of Fungal Strains for Conversion of D-Galacturonate to meso-Galactarate. Applied and Environmental Microbiology 76**:** 169-175.

Montenecourt, B. S., and D. E. Eveleigh, 1977 Preparation of mutants of Trichoderma reesei with enhanced cellulase production. Appl Environ Microbiol 34**:** 777-782.

Murakoshi, Y., T. Makita, M. Kato and T. Kobayashi, 2012 Comparison and characterization of alpha-amylase inducers in *Aspergillus nidulans* based on nuclear localization of AmyR. Appl Microbiol Biotechnol 94**:** 1629-1635.

Myers, A. M., M. K. Morell, M. G. James and S. G. Ball, 2000 Recent progress toward understanding biosynthesis of the amylopectin crystal. Plant Physiol 122**:** 989-997.

Nielsen, M. L., L. Albertsen, G. Lettier, J. B. Nielsen and U. H. Mortensen, 2006 Efficient PCR-based gene targeting with a recyclable marker for *Aspergillus nidulans*. Fungal Genet Biol 43**:** 54-64.

Ninomiya, Y., K. Suzuki, C. Ishii and H. Inoue, 2004 Highly efficient gene replacements in *Neurospora* strains deficient for nonhomologous end-joining. Proc Natl Acad Sci U S A 101**:** 12248-12253.

Niu, J., E. Alazi, I. D. Reid, M. Arentshorst, P. J. Punt *et al.*, 2017 An Evolutionarily Conserved Transcriptional Activator-Repressor Module Controls Expression of Genes for D-Galacturonic Acid Utilization in *Aspergillus niger*. Genetics 205**:** 169-183.

Niu, J., M. Arentshorst, P. D. Nair, Z. Dai, S. E. Baker *et al.*, 2015 Identification of a Classical Mutant in the Industrial Host *Aspergillus niger* by Systems Genetics: LaeA Is Required for Citric Acid Production and Regulates the Formation of Some Secondary Metabolites. G3 (Bethesda) 6**:** 193-204.

Niu, J., M. Arentshorst, F. Seelinger, A. F. Ram and J. P. Ouedraogo, 2016 A set of isogenic auxotrophic strains for constructing multiple gene deletion mutants and parasexual crossings in *Aspergillus niger*. Arch Microbiol 198**:** 861-868.

Nodvig, C. S., J. B. Nielsen, M. E. Kogle and U. H. Mortensen, 2015 A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi. PLoS One 10**:** e0133085.

Noguchi, Y., M. Sano, K. Kanamaru, T. Ko, M. Takeuchi *et al.*, 2009 Genes regulated by AoXlnR, the xylanolytic and cellulolytic transcriptional regulator, in *Aspergillus oryzae*. Appl Microbiol Biotechnol 85**:** 141-154.

Noguchi, Y., H. Tanaka, K. Kanamaru, M. Kato and T. Kobayashi, 2011 Xylose triggers reversible phosphorylation of XlnR, the fungal transcriptional activator of xylanolytic and cellulolytic genes in *Aspergillus oryzae*. Biosci Biotechnol Biochem 75**:** 953-959.

Nyren, P., and A. Lundin, 1985 Enzymatic method for continuous monitoring of inorganic pyrophosphate synthesis. Anal Biochem 151**:** 504-509.

Ogawa, M., T. Kobayashi and Y. Koyama, 2012 ManR, a novel Zn(II)2Cys6 transcriptional activator, controls the beta-mannan utilization system in *Aspergillus oryzae*. Fungal Genet Biol 49**:** 987-995.

Ogawa, M., T. Kobayashi and Y. Koyama, 2013 ManR, a transcriptional regulator of the beta-mannan utilization system, controls the cellulose utilization system in *Aspergillus oryzae*. Biosci Biotechnol Biochem 77**:** 426-429.

Pandey, A., P. Selvakumar, C. R. Soccol and P. Nigam, 1999 Solid state fermentation for the production of industrial enzymes. Current Science 77**:** 149-162.

Pardo, E., and M. Orejas, 2014 The *Aspergillus nidulans* Zn(II)2Cys6 transcription factor AN5673/RhaR mediates L-rhamnose utilization and the production of alpha-L-rhamnosidases. Microb Cell Fact 13**:** 161.

Payne, S. H., and W. F. Loomis, 2006 Retention and loss of amino acid biosynthetic pathways based on analysis of whole-genome sequences. Eukaryot Cell 5**:** 272-276.

Pedersen, H., M. Beyer and J. Nielsen, 2000 Glucoamylase production in batch, chemostat and fed-batch cultivations by an industrial strain of *Aspergillus niger*. Appl Microbiol Biotechnol 53**:** 272-277.

Pel, H. J., J. H. de Winde, D. B. Archer, P. S. Dyer, G. Hofmann *et al.*, 2007 Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. Nat Biotechnol 25**:** 221-231.

Pessoa, M. G., B. N. Paulino, M. C. R. Mano, I. A. Neri-Numa, G. Molina *et al.*, 2017 Fusarium species-a promising tool box for industrial biotechnology. Appl Microbiol Biotechnol 101**:** 3493-3511.

Petersen, K. L., J. Lehmbeck and T. Christensen, 1999 A new transcriptional activator for amylase genes in *Aspergillus*. Mol Gen Genet 262**:** 668-676.

Pitt, J. I., 1980 *The Genus Penicillium and its Teleomorphic States Eupenicillium and Talaromyces*. London: Academic Press.

Punt, P. J., F. H. Schuren, J. Lehmbeck, T. Christensen, C. Hjort *et al.*, 2008 Characterization of the *Aspergillus niger* prtT, a unique regulator of extracellular protease encoding genes. Fungal Genet Biol 45**:** 1591-1599.

Ram, A. F., 2013 Modern Technologies Boost Classical Genetics: Whole Genome Sequencing Revives Forward Genetic Mutant Screens in Filamentous Fungi. Fungal Genomics & Biology 2**:** e105.

Raulo, R., M. Kokolski and D. B. Archer, 2016 The roles of the zinc finger transcription factors XlnR, ClrA and ClrB in the breakdown of lignocellulose by *Aspergillus niger*. AMB Express 6**:** 5.

Ren, B., F. Robert, J. J. Wyrick, O. Aparicio, E. G. Jennings *et al.*, 2000 Genome-wide location and function of DNA binding proteins. Science 290**:** 2306-2309.

Rennie, E. A., and H. V. Scheller, 2014 Xylan biosynthesis. Curr Opin Biotechnol 26**:** 100-107.

Rhoads, A., and K. F. Au, 2015 PacBio Sequencing and Its Applications. Genomics Proteomics Bioinformatics 13**:** 278-289.

Ries, L. N., S. R. Beattie, E. A. Espeso, R. A. Cramer and G. H. Goldman, 2016 Diverse Regulation of the CreA Carbon Catabolite Repressor in *Aspergillus nidulans*. Genetics 203**:** 335-352.

Robertson, G., M. Hirst, M. Bainbridge, M. Bilenky, Y. Zhao *et al.*, 2007 Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. Nat Methods 4**:** 651-657.

Ronaghi, M., S. Karamohamed, B. Pettersson, M. Uhlen and P. Nyren, 1996 Real-time DNA sequencing using detection of pyrophosphate release. Anal Biochem 242**:** 84-89.

Rose, J. K., and S. J. Lee, 2010 Straying off the highway: trafficking of secreted plant proteins and complexity in the plant cell wall proteome. Plant Physiol 153**:** 433-436.

Roy, P., R. A. Lockington and J. M. Kelly, 2008 CreA-mediated repression in *Aspergillus nidulans* does not require transcriptional auto-regulation, regulated intracellular localisation or degradation of CreA. Fungal Genet Biol 45**:** 657-670.

Ruijter, G. J., S. A. Vanhanen, M. M. Gielkens, P. J. van de Vondervoort and J. Visser, 1997 Isolation of *Aspergillus niger* creA mutants and effects of the mutations on expression of arabinases and L-arabinose catabolic enzymes. Microbiology 143 (Pt 9)**:** 2991-2998.

Ruijter, G. J., and J. Visser, 1997 Carbon repression in *Aspergilli*. FEMS Microbiol Lett 151**:** 103-114.

Sander, J. D., and J. K. Joung, 2014 CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol 32**:** 347-355.

Sanger, F., S. Nicklen and A. R. Coulson, 1977 DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74**:** 5463-5467.

Scazzocchio, C., V. Gavrias, B. Cubero, C. Panozzo, M. Mathieu *et al.*, 1995 Carbon Catabolite Repression in *Aspergillus nidulans*-a Review. Canadian Journal of Botany-Revue Canadienne De Botanique 73**:** S160-S166.

Schena, M., D. Shalon, R. W. Davis and P. O. Brown, 1995 Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270**:** 467-470.

Schneeberger, K., 2014 Using next-generation sequencing to isolate mutant genes from forward genetic screens. Nat Rev Genet 15**:** 662-676.

Schneeberger, K., S. Ossowski, C. Lanz, T. Juul, A. H. Petersen *et al.*, 2009 SHOREmap: simultaneous mapping and mutation identification by deep sequencing. Nat Methods 6**:** 550-551.

Schrickx, J. M., A. H. Stouthamer and H. W. van Verseveld, 1995 Growth behaviour and glucoamylase production by *Aspergillus niger* N402 and a glucoamylase overproducing transformant in recycling culture without a nitrogen source. Appl Microbiol Biotechnol 43**:** 109-116.

Schuster, E., N. Dunn-Coleman, J. C. Frisvad and P. W. Van Dijck, 2002 On the safety of *Aspergillus niger*-a review. Appl Microbiol Biotechnol 59**:** 426-435.

Shallom, D., and Y. Shoham, 2003 Microbial hemicellulases. Curr Opin Microbiol 6**:** 219-228.

Shelest, E., 2017 Transcription Factors in Fungi: TFome Dynamics, Three Major Families, and Dual-Specificity TFs. Front Genet 8**:** 53.

Sloothaak, J., M. Schilders, P. J. Schaap and L. H. de Graaff, 2014 Overexpression of the *Aspergillus niger* GatA transporter leads to preferential use of D-galacturonic acid over D-xylose. AMB Express 4**:** 66.

Srivatsan, A., Y. Han, J. Peng, A. K. Tehranchi, R. Gibbs *et al.*, 2008 High-precision, whole-genome sequencing of laboratory strains facilitates genetic studies. PLoS Genet 4**:** e1000139.

Steinkraus, K. K., 1983 *Handbook of Indigenous Fermented Foods*. New York & Basel: Marcel Dekker.

Steup, M., 1988 Starch degradation, pp. 255-296 in *The Biochemistry of Plants*, edited by J. Preiss. Academic Press, San Diego, CA, USA.

Suzuki, K., M. Tanaka, Y. Konno, T. Ichikawa, S. Ichinose *et al.*, 2015 Distinct mechanism of activation of two transcription factors, AmyR and MalR, involved in amylolytic enzyme production in *Aspergillus oryzae*. Appl Microbiol Biotechnol 99**:** 1805-1815.

Tamayo, E. N., A. Villanueva, A. A. Hasper, L. H. de Graaff, D. Ramon *et al.*, 2008 CreA mediates repression of the regulatory gene xlnR which controls the production of xylanolytic enzymes in *Aspergillus nidulans*. Fungal Genet Biol 45**:** 984-993.

Tani, S., Y. Katsuyama, T. Hayashi, H. Suzuki, M. Kato *et al.*, 2001 Characterization of the amyR gene encoding a transcriptional activator for the amylase genes in *Aspergillus nidulans*. Curr Genet 39**:** 10-15.

Todd, R. B., and A. Andrianopoulos, 1997 Evolution of a fungal regulatory gene family: the Zn(II)2Cys6 binuclear cluster DNA binding motif. Fungal Genet Biol 21**:** 388-405.

Todd, R. B., R. A. Lockington and J. M. Kelly, 2000 The *Aspergillus nidulans* creC gene involved in carbon catabolite repression encodes a WD40 repeat protein. Mol Gen Genet 263**:** 561-570.

Tsukagoshi, N., T. Kobayashi and M. Kato, 2001 Regulation of the amylolytic and (hemi-)cellulolytic genes in *aspergilli*. J Gen Appl Microbiol 47**:** 1-19.

Turner, W. B., 1971 *Fungal Metabolites*. London & New York: Academic Press.

Turner, W. B., and D. C. Aldridge, 1983 *Fungal Metabolite II*. London: Academic Press.

van Attikum, H., P. Bundock and P. J. Hooykaas, 2001 Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. EMBO J 20**:** 6550-6558.

van Attikum, H., and P. J. J. Hooykaas, 2003 Genetic requirements for the targeted integration of *Agrobacterium* T-DNA in Saccharomyces cerevisiae. Nucleic Acids Research 31**:** 826-832.

van Peij, N. N., M. M. Gielkens, R. P. de Vries, J. Visser and L. H. de Graaff, 1998a The transcriptional activator XlnR regulates both xylanolytic and endoglucanase gene expression in *Aspergillus niger*. Appl Environ Microbiol 64**:** 3615-3619.

van Peij, N. N., J. Visser and L. H. de Graaff, 1998b Isolation and analysis of xlnR, encoding a transcriptional activator co-ordinating xylanolytic expression in *Aspergillus niger*. Mol Microbiol 27**:** 131-142.

vanKuyk, P. A., J. A. Benen, H. A. Wosten, J. Visser and R. P. de Vries, 2012 A broader role for AmyR in *Aspergillus niger*: regulation of the utilisation of D-glucose or D-galactose containing oligo- and polysaccharides. Appl Microbiol Biotechnol 93**:** 285-293.

Vogt, K., R. Bhabhra, J. C. Rhodes and D. S. Askew, 2005 Doxycycline-regulated gene expression in the opportunistic fungal pathogen *Aspergillus fumigatus*. BMC Microbiol 5**:** 1.

Voragen, A. G. J., and W. Pilnik, 1989 Pectin-degrading enzymes in fruit and vegetable processiong, pp. 93-155 in *Biocatalysis in Agricultural Biotechnology*, edited by J. R. Whitaker and P. E. Sonnet. American Chemical Society, Washington, D.C.

Wang, Z., M. Gerstein and M. Snyder, 2009 RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10**:** 57-63.

Wanka, F., T. Cairns, S. Boecker, C. Berens, A. Happel *et al.*, 2016 Tet-on, or Tet-off, that is the question: Advanced conditional gene expression in *Aspergillus*. Fungal Genet Biol 89**:** 72-83.

Wiebe, M. G., D. Mojzita, S. Hilditch, L. Ruohonen and M. Penttila, 2010 Bioconversion of D-galacturonate to keto-deoxy-L-galactonate (3-deoxy-L-threo-hex-2-ulosonate) using filamentous fungi. Bmc Biotechnology 10.

Wiedenheft, B., S. H. Sternberg and J. A. Doudna, 2012 RNA-guided genetic silencing systems in bacteria and archaea. Nature 482**:** 331-338.

Xiong, Y., V. W. Wu, A. Lubbe, L. Qin, S. Deng *et al.*, 2017 A fungal transcription factor essential for starch degradation affects integration of carbon and nitrogen metabolism. PLoS Genet 13**:** e1006737.

Yuan, X. L., C. Goosen, H. Kools, M. J. van der Maarel, C. A. van den Hondel *et al.*, 2006 Database mining and transcriptional analysis of genes encoding inulin-modifying enzymes of *Aspergillus niger*. Microbiology 152**:** 3061-3073.

Yuan, X. L., J. A. Roubos, C. A. van den Hondel and A. F. Ram, 2008a Identification of InuR, a new Zn(II)2Cys6 transcriptional activator involved in the regulation of inulinolytic genes in *Aspergillus niger*. Mol Genet Genomics 279**:** 11-26.

Yuan, X. L., R. M. van der Kaaij, C. A. van den Hondel, P. J. Punt, M. J. van der Maarel *et al.*, 2008b *Aspergillus niger* genome-wide analysis reveals a large number of novel alpha-glucan acting enzymes with unexpected expression profiles. Mol Genet Genomics 279**:** 545-561.

Zhang, J. Z., 2003 Overexpression analysis of plant transcription factors. Curr Opin Plant Biol 6**:** 430-440.

Zhang, L., C. Hua, J. H. Stassen, S. Chatterjee, M. Cornelissen *et al.*, 2014 Genome-wide analysis of pectate-induced gene expression in *Botrytis cinerea*: identification and functional analysis of putative d-galacturonate transporters. Fungal Genet Biol 72**:** 182-191.

Zhang, L., R. J. Lubbers, A. Simon, J. H. Stassen, P. R. Vargas Ribera *et al.*, 2016 A novel Zn2 Cys6 transcription factor BcGaaR regulates D-galacturonic acid utilization in *Botrytis cinerea*. Mol Microbiol 100**:** 247-262.

Zhang, L. S., H. Thiewes and J. A. L. van Kan, 2011 The D-galacturonic acid catabolic pathway in *Botrytis cinerea*. Fungal Genetics and Biology 48**:** 990-997.

Chapter 2

Efficient Generation of *Aspergillus niger* Knock Out Strains by Combining NHEJ Mutants and a Split Marker Approach

Mark Arentshorst, Jing Niu and Arthur F.J. Ram

Published in *Genetic Transformation Systems in Fungi*, edited by M. A. van Berg and K. Maruthachalam. Volume 1, Fungal Biology

Abstract

To generate gene deletion mutants in *Aspergillus niger*, we combined the use of Non-Homologous End Joining (NHEJ) mutants (*ku70* mutant) and the split marker approach. The combination of both tools resulted in efficient PCR amplification because of the reduced length of the PCR fragments and efficient homologous recombination frequencies. A set of five selection markers, two dominant selection markers (*hph*; hygromycin B resistance and BLE; phleomycin resistance) and three auxotrophic markers (*pyrG*, *argB* and *nicB*) were successfully used in a split marker approach to obtain *amyR* knock outs with high efficiency. *AmyR* encodes a transcription factor that is required for the expression of starch degrading enzymes and disruption of *amyR* results in the inability to grow on starch. The strategy to generate the gene deletion constructs is such that with one set of four gene specific primers, a gene deletion mutant can be generated with either one of the five selection markers. The strategy is based on fusion PCR and omits the necessity for cloning the disruption cassettes. This accelerates the process of generating gene deletion cassettes which can now be accomplished within eight hours. The split marker approach can also be used to make gene deletions in a wild-type background instead of a *Δku70* background. In this chapter, we present protocols and considerations that we used to generate gene knock out constructs by fusion PCR and to obtain and verify gene knock outs with any of the five marker genes using the split marker approach. The method is easily transferable to other filamentous fungi.

Keywords

Aspergillus niger, ku70, split marker approach, five available selection markers.

1. Introduction

Targeted deletion of a Gene of Interest (GOI) is a powerful method to address gene functions and requires a double crossover homologous recombination event to exchange the GOI with a selection marker. In filamentous fungi, DNA integrates preferably via the Non-Homologous End Joining (NHEJ) pathway, which results in low frequencies of homologous recombination and consequently, in low efficiencies in obtaining gene deletion mutants. A successful approach to obtain gene deletion mutants with high efficiency has been the construction of mutants in the NHEJ-pathway, first described for *Neurospora crassa* (Ninomiya *et al.* 2004), and followed up by numerous other filamentous fungi including *Aspergillus niger* (Meyer *et al.* 2007; Carvalho *et al.* 2010; Arentshorst *et al.* 2012). Most often the fungal gene homologous to the gene encoding the Ku70 is used to generate a NHEJ mutant, but also Ku80 and Lig4 homologs have been disrupted to obtain NHEJ-deficient mutants (for reviews see (Meyer 2008), (Kuck and Hoff 2010) and references therein). The use of NHEJ mutants has greatly reduced time and effort to generate gene deletion mutants. The construction of a gene deletion cassette is also an important and time consuming factor. In principle, a gene deletion construct consists of a selection marker, flanked by upstream (5') and downstream (3') sequences of the GOI. Several approaches to generate gene deletion cassettes include traditional restriction enzyme and ligation based cloning, GATEWAY cloning, fusion PCR, or *in vivo* assembly either in *E. coli* or *S. cerevisiae*.

An additional tool for improving gene targeting efficiencies is making use of the split marker technology. In this approach the gene deletion construct is split in two parts and each part contains the flanking region and a truncated form of the selection marker (Fairhead *et al.* 1996; Nielsen *et al.* 2006; Goswami 2012).

For the selection of transformants in *A. niger* (and also other filamentous fungi) the number of available markers is limited. Dominant selection markers for *A. niger* include markers giving resistance to hygromycin (pAN7.1) (Punt *et al.* 1987) or phleomycin (pAN8.1) (Punt and Van Den Hondel 1992), which are well established and commonly used. The uridine and arginine markers (*pyrG* (An12g03570) and *argB* (An14g03400), respectively), have been described earlier and are used in this study (Buxton *et al.* 1985; Van Hartingsveldt *et al.* 1987; Lenouvel *et al.* 2002). The *pyrG* gene encodes for the enzyme orotidine-5'-phosphate-decarboxylase and is required for uracil biosynthesis. The *argB* gene, encoding for an ornithine carbamoyltransferase, is essential for arginine biosynthesis. In addition, a new auxotrophic mutant which requires nicotinamide for growth based on the *nicB* gene (An11g10910) was made. The *A. niger nicB* gene encodes a nicotinate-nucleotide

pyrophosphorylase. Identification and the construction of a gene deletion cassette to disrupt *nicB* is based on a previous work by (Verdoes *et al.* 1994), and will be described elsewhere in detail (Niu et al. manuscript in preparation). The *ΔnicB* strain is auxotrophic for nicotinamide and needs supplementation of nicotinamide to be able to grow. In addition, we reconstructed an *argB* deletion mutant (Niu et al. manuscript in preparation) to have all auxotrophic strains in the same strain background (Table 1).

Table 1. Strains used in this study.

Strain	Genotype	Description	Reference
N402	*cspA1*	derivative of N400	(Bos *et al.* 1988)
AB4.1	*cspA1, pyrG378*	UV mutant of N402	(Van Hartingsveldt *et al.* 1987)
MA169.4	*cspA1, pyrG378, kusA::DR-amdS-DR*	*ku70* deletion in AB4.1	(Carvalho *et al.* 2010)
MA234.1	*cspA1, kusA::DR-amdS-DR*	*ku70* deletion in N402	Arentshorst (unpublished)
JN1.17	*cspA1, pyrG378, kusA::DR-amdS-DR argB::hph*	*argB* deletion in MA169.4	Niu *et al*. unpublished
JN4.2	*cspA1, pyrG378, kusA::DR-amdS-DR nicB::hph*	*nicB* deletion in MA169.4	Niu *et al*. unpublished

Growth of all three auxotrophic strains (*pyrG*$^-$, *argB*$^-$ and *nicB*$^-$) on minimal medium requires the addition of uridine, L-arginine or nicotinamide, respectively[1], and no growth is observed in the absence of the relevant supplements (data not shown). To minimize homologous recombination of the selection markers used in the disruption cassettes, the *argB* and *nicB* homologs from *A. nidulans* (ANID_04409.1 and ANID_03431.1 respectively) and the *pyrG* homolog from *A. oryzae* (AO090011000868) were PCR amplified. All genes are able to complement the auxotrophy of the relevant strain. The hygromycin and phleomycin cassettes also contain only non-homologous sequences as both resistance genes are flanked by the *A. nidulans gpdA* promoter (*PgpdA*) and *trpC* terminator *(TtrpC)* (Table 2).

Table 2. Plasmids to amplify selection markers.

Plasmid	Selection marker	Remark	Reference
pAN7.1	Hygromycin; *hph*	*Pgpd* and *TtrpC* from *A.nidulans*	(Punt *et al.* 1987)
pAN8.1	Phleomycin; BLE	*Pgpd* and *TtrpC* from *A.nidulans*	(Punt and Van Den Hondel 1992)
pAO4-13	*pyrG*	*pyrG* from *A.oryzae*	(De Ruiter-Jacobs *et al.* 1989)
pJN2.1	*argB*	*argB* from *A. nidulans*	Niu *et al*. unpublished
pJN4.1	*nicB*	*nicB* from *A. nidulans*	Niu *et al*. unpublished

[1] The *argB* and *nicB* auxotrofic mutants are also *pyrG*$^-$ and therefore the growth medium for these strains needs to be supplemented with uridine.

2. General Methods

2.1 General split marker approach

The split marker approach used for deleting the GOI is schematically depicted in Figure 1 and consists of two overlapping DNA fragments to disrupt the GOI. The first fragment contains the 5'flank of the GOI and a truncated version of the selection maker. The second DNA fragment contains an overlapping, but truncated version of the selection marker and the 3'flank of the GOI. Both fragments are generated by fusion PCR as described below and transformed simultaneously to the recipient *A. niger* strain. The truncation of the selection marker at either site of the construct results in a non-functional marker and as a consequence transformation of only a single split marker fragment does not result in any transformants (data not shown).

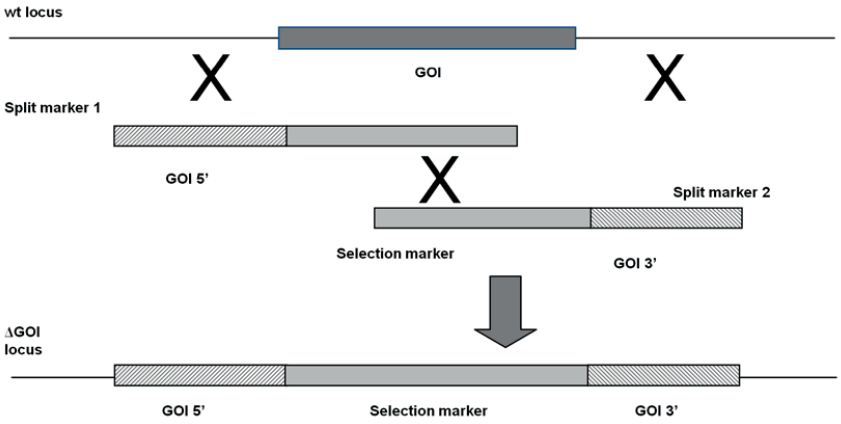


Figure 1. Schematic representation of the split marker gene deletion approach. 5' and 3'seqences flanking the GOI (5'and 3') are transformed simultaneously to the recipient strain. By recombination of the selection maker and homologous integration of the cassette in the genome, a successful gene deletion mutant can be obtained.

2.2 Generation of split marker fragments for *Aspergillus niger* transformation

In this section the experimental design for creating the split marker fragments is discussed. The split marker DNA fragments can be obtained in three steps (Figure 2). Each step is described in detail below.

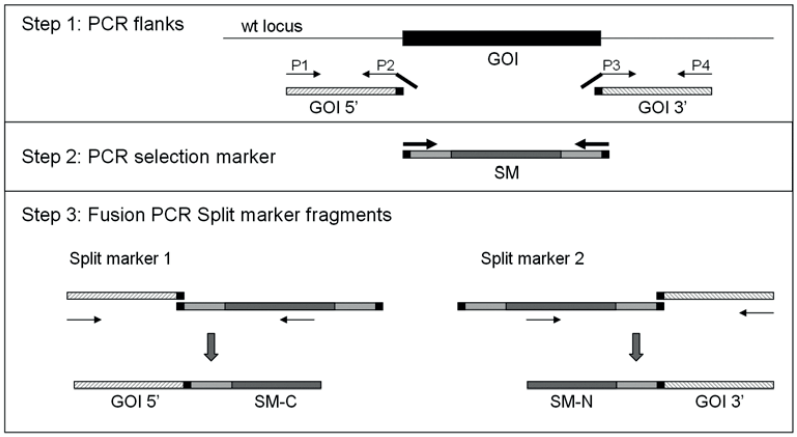


Figure 2. Experimental design for creating split marker fragments.

2.2.1 Experimental design for amplification of flanking regions of the GOI (Step 1)

Once the GOI has been identified, primers need to be designed for making gene deletion cassettes. First, two primers are required for the amplification of the 5' flank of the GOI. The first primer (P1) is chosen between 700 and 900 bases upstream of the start codon. The reverse primer (P2) is as close to the start codon as possible and contains a 5'-CAATTCCAGCAGCGGCTT-3' sequence, which is overlapping with all five selection markers and included for the subsequent fusion PCR. Also, two primers are required for the amplification of the 3' flank of the GOI (P3 and P4). Again, the aim is to generate a 700-900 base pair long flank. In this case, the forward primer (P3) needs a 5'-ACACGGCACAATTATCCATCG-3' sequence, which is also overlapping with all five selection markers for the subsequent fusion PCR (Step3).

2.2.2 Experimental design for amplification of suitable selection marker (Step 2)

For the amplification of the PCR fragments containing the appropriate selection marker the following plasmids can be used (see also Table 2):

The plasmid pAN7.1 (Punt *et al.* 1987) is used as template to amplify the hygromycin resistance cassette, containing the *hph* gene from *E. coli*, coding for hygromycin B phosphotransferase. Expression of the *hph* gene is driven by the *A. nidulans gpdA* promoter, and terminated by the *A. nidulans trpC* terminator. The plasmid pAN8.1 (Punt and Van Den Hondel 1992) is used as template to amplify the phleomycin resistance cassette, containing the BLE gene from *Streptoalloteichus hindustanus*, coding for a phleomycin binding protein.

Expression of the BLE gene is also driven by the *A. nidulans gpdA* promoter and terminated by the *A. nidulans trpC* terminator. The plasmid pAO4-13 (De Ruiter-Jacobs *et al.* 1989) is used as template to amplify the *A. oryzae pyrG* gene (AO090011000868), including promoter and terminator region. The *argB* gene (ANID_04409.1) and the *nicB* gene (ANID_03431.1) of *A. nidulans*, including promoter and terminator region, were amplified using primer pairs argBnidP5f and argBnidP6r or nicBnidP5f and nicBnidP6r, and genomic DNA of *A. nidulans* strain FGSC A234 (*yA2, pabaA1, veA1*), obtained from the Fungal Genetics Stock Center, as template. The resulting PCR products were ligated into PCR-cloning vector pJet1.2 (K1231, Thermo Fisher), to give plasmids pJN2.1 and pJN4.1 respectively (Table 2). Plasmid pJN2.1 and pJN4.1 can be used to amplify the *argB* gene or the *nicB* gene.

We developed a generic split marker approach in such a way that with a single set of four GOI primers, all five different selection markers can be used to generate the deletion cassette. Each primer, used to amplify a specific selection marker (Figure 3, Table 3), contains sequences which are overlapping with the GOI primer sequences (see section 2.2.1) to create gene deletion mutants with either one of the different selection markers.

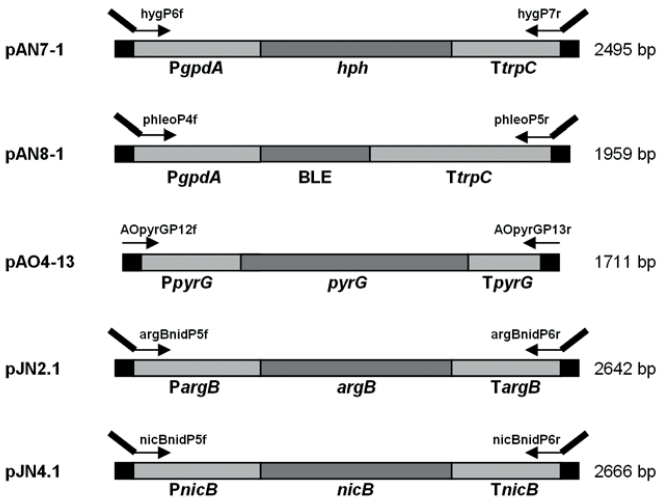


Figure 3. PCR products for all five selection markers. Overlapping sequences of the primers are indicated by bold lines. The size of the PCR products is indicated for each selection marker.

2.2.3 Experimental design for the generation of split marker fragments (Step 3)

Once both flanks of the GOI (Fig. 2, step 1) and the required selection marker (Fig. 2 step 2 and Fig. 3) have been amplified, the split marker fragments can be obtained by fusion PCR (Fig. 2, step 3). Exact details are described in section 3.2.3. After column purification, the resulting split marker fragments can directly be used to transform *A. niger*[2].

3. Detailed procedure description

As proof of principle, the *A. niger amyR* gene (An04g06910), encoding the amylase transcriptional regulator, has been used. The *ΔamyR* strain cannot grow on starch, allowing an easy screen for *ΔamyR* transformants (Petersen *et al.* 1999). This section contains a detailed description of the whole procedure of deleting *amyR*, using all five selection markers, illustrated with results of the experiments. Sequences of all primers used are listed in Tables 3, 4, and 5.

Table 3. Primers used to generate selection markers.

Primer name	Sequence (5'-3')	Remark	Template
hygP6for	**AAGCCGCTGCTGGAATTG-**GGCTCTGAGGTGCAGTGGAT	Amplification of hph marker	pAN7.1
hygP7rev	**CGATGGATAATTGTGCCGTGT-**TGGGTGTTACGGAGCATTCA	Amplification of hph marker	pAN7.1
phleoP4for	**AAGCCGCTGCTGGAATTG -**CTCTTTCTGGCATGCGGAG	Amplification of BLE marker	pAN8.1
phleoP5rev	**CGATGGATAATTGTGCCGTGT-**GGAGCATTCACTAGGCAACCA	Amplification of BLE marker	pAN8.1
AOpyrGP12f	**AAGCCGCTGCTGGAATTG**	Amplification of *pyrG* marker	pAO4-13
AOpyrGP13r	**CGATGGATAATTGTGCCGTGT**	Amplification of *pyrG* marker	pAO4-13
argBnidP5f	**AAGCCGCTGCTGGAATTG -**TTTCGACCTCTTTCCCAATCC	Amplification of *argB* marker	pJN2.1
argBnidP6r	**CGATGGATAATTGTGCCGTGT-**TCCTGTGGGTCTTTGTCCG	Amplification of *argB* marker	pJN2.1
nicBnidP5f	**AAGCCGCTGCTGGAATTG-**CGTTATGCACAGCTCCGTCTT	Amplification of *nicB* marker	pJN4.1
nicBnidP6r	**CGATGGATAATTGTGCCGTGT-**GCGCATACACAGAAGCATTGA	Amplification of *nicB* marker	pJN4.1

Note: overlapping sequences for fusion PCR are indicated in *bold*.

[2] A small sample of PCR fragments is routinely analyzed for purity and size. Optional is to confirm PCR product integrity by restriction analysis or sequencing.

Table 4. GOI (amyR) specific primers to amplify 5'and 3'flanks.

Primer name	Sequence (5'-3')	Remark
amyRP7f	ATCGTCAGCGAGCCTCAGA	Amplification of *amyR* 5'flank
amyRP8r	**CAATTCCAGCAGCGGCTT-**TTGTATGCGGAGACAAGTGTGAC	Amplification of *amyR* 5'flank
amyRP9f	**ACACGGCACAATTATCCATCG-**CCCTCATGAACAAGAAGCAGC	Amplification of *amyR* 3'flank
amyRP10r	GAGGACGCCATCATTGACG	Amplification of *amyR* 3'flank

Note: overlapping sequences for fusion PCR are indicated in *bold*.

Table 5. Generic primers used to amplify bipartite fragments.

Primer name	Sequence (5'-3')	Remark
hygP9r	GGCGTCGGTTTCCACTATC	reverse primer split marker fragment 1 *hph*
hygP8f	AAAGTTCGACAGCGTCTCC	forward primer split marker fragment 2 *hph*
phleoP7r	CACGAAGTGCACGCAGTTG	reverse primer split marker fragment 1 BLE
phleoP6f	AAGTTGACCAGTGCCGTTCC	forward primer split marker fragment 2 BLE
AOpyrGP15r	CCGGTAGCCAAAGATCCCTT	reverse primer split marker fragment 1 *pyrG*
AOpyrGP14f	ATTGACCTACAGCGCACGC	forward primer split marker fragment 2 *pyrG*
argBnidP8r	TGGTTTGCAGAAGCTTTCCT	reverse primer split marker fragment 1 *argB*
argBnidP7f	ACTCCTCGCAAACCATGCC	forward primer split marker fragment 2 *argB*
nicBnidP8r	GAACAGCCTTCGGGATTGC	reverse primer split marker fragment 1 *nicB*
nicBnidP7f	CGCCTTATATCCGATTGGCT	forward primer split marker fragment 2 *nicB*

3.1 Materials and Reagents

For the medium composition of minimal medium, the preparation of stock solutions for the medium and for a detailed protocol of genomic DNA isolation of *A. niger* we refer to the Materials and Reagents section in Arentshorst *et al*. 2012.

1. PCR enzyme (we routinely use Phire Hot start II DNA Polymerase (F-122L, Thermo Fisher).
2. dNTPs (1.25 mM): Add 0.25 mL of all 4 dNTPs (dNTP Set 100 mM Solutions (4 x 0.25 mL, R0181, Thermo Fisher)) to 19 mL of MQ, mix well, make aliquots of 0.5 mL and store at -20˚C.
3. PCR purification Kit (we routinely use Genejet Gel Extraction Kit (K0692, Thermo Fisher), also for PCR purifications).
4. Hygromycin (100 mg/mL): Dissolve 1 g of hygromycin (InvivoGen, ant-hg-10p) in 10 mL of MQ, sterilize by filtration, make aliquots of 500 µL and store at −20°C. The final concentration in the medium is 100 µg/mL, except for transformation plates, then use 200 µg/mL.
5. Phleomycin (40 mg/mL), for 10 mL: add 400 mg of phleomycin (InvivoGen, ant-ph-10p) to 8 mL of warm MQ (~60˚C) in a 15 mL tube. When phleomycin is dissolved, add MQ up to 10 mL and filter sterilize. Make aliquots and store at -20˚C.

6. Uridine (1 M), for 100 mL: add 22.4 g of uridine (Acros, 140775000) to 50 mL of warm MQ (~60˚C) in a 100 mL cylinder. When uridine is dissolved, add MQ up to 100 mL, sterilize by filtration and store at 4°C. Final concentration in medium is 10 mM.
7. Arginine (2%), for 100 mL: add 2 g of L-arginine monohydrochloride (Sigma, A5131) to 50 mL of warm MQ (~60˚C) in a 100 mL cylinder. When arginine is dissolved, add MQ up to 100 mL, sterilize by filtration and store at 4˚C.
8. Nicotinamide (0.5%), for 100 mL: add 0.5 g of nicotinamide (Sigma, N0636) to 50 mL of warm MQ (~60˚C) in a 100 mL cylinder. When nicotinamide is dissolved, add MQ up to 100 mL, sterilize by filtration and store at 4˚C.
9. Transformation media + phleomycin: Prepare MMS and Top agar according to Arentshorst et al. 2012. After autoclaving, and cooling down to 50ºC, add phleomycin to a final concentration of 50 µg/mL, to both the MMS and the Top agar.
10. MM + agar + L-arginine: Prepare 500 mL of MM + agar according to Arentshorst et al. 2012. Add 5 mL of 2% L-arginine after autoclaving (100 x dilution).
11. MM + agar + nicotinamide: Prepare 500 mL of MM + agar according to Arentshorst et al. 2012. Add 0.25 mL of 0.5% nicotinamide after autoclaving (2000 x dilution).
12. MM + agar + starch: For 500 mL: Dissolve 5 g of starch (soluble, extra pure, Merck, 1.01253) in 450 mL of warm MQ (~60˚C). Add 10 mL of 50 x ASP+N, 1 ml of 1M $MgSO_4$, 50 mL of trace element solution, 15 mg of yeast extract (YE)[3] (Roth, 2363.2) and 7.5 g agar bact. (Scharlau, 07-004-500), and autoclave.

3.2 Methods

*3.2.1 Amplification of the amyR 5'- and 3' flan*k

1. *AmyR* primers were designed (Fig. 2, Step 1 and Table 4), and subsequently used in PCR reactions to amplify both the *amyR* 5' flank and 3' flank.
2. The PCR mix, total volume of 50 µL, contained 1 µL genomic DNA of *A. niger* wt strain N402 (1µg/µL), 8 µL dNTP's (1.25 mM), 10 µL 5x Phire buffer, 1 µL Primer F (20 pmol/µL), 1 µL Primer R (20 pmol/µL), 0.5 µL Phire Hot start II DNA Polymerase and 28.5 µL of MQ.

[3] YE is added to a final concentration of 0.003% to stimulate germination of *A.niger*. On MM + starch without YE, the wt strain also does not grow very well.

3. PCR was performed under the following conditions: initial denaturation for 5 min at 98˚C, 30 cycles of 5 sec at 98˚C, 5 sec at 58˚C, and 15 sec per 1 kb of template at 72˚C, followed by final extension of 5 min at 72˚C.
4. PCR reactions were analyzed by loading 5 µL PCR reaction on a 1% agarose gel.
5. After column purification and elution with 30 µL of MQ, DNA concentration for both flanks was ~37 ng/µL.

3.2.2 Amplification of the selection markers

1. Primers for all five selection markers were designed (Fig 3, Table 3) and used for PCR. In these PCR reactions 1 ng of plasmid (pAO4-13, pAN7.1, pAN8.1, pJN2.1 and pJN4.1, respectively) was used as template. For PCR mix and PCR conditions see section 3.2.1.
2. After confirmation on agarose gel, selection marker PCR products were column purified, yielding DNA concentrations of ~50 ng/µL. The markers were stored at -20˚C and used repeatedly.

3.2.3 Amplification of the split marker fragments

1. Fusion PCR fragments were amplified according to Figure 2, step 3 (see also Table 5 and 6). Both *amyR* flanks and all selection markers (section 3.2.1 and 3.2.2) were diluted to 2 ng/µL.
2. For each PCR reaction, 2 ng of *amyR* flank and 2 ng of selection marker PCR were used as template (Table 6). For PCR mix and PCR conditions see section 3.2.1.
3. Two identical fusion PCR reactions were performed, in order to increase the yield of PCR product.
4. Fusion PCR products were analyzed on agarose gel, followed by column purification. The DNA concentration for all fragments varied between 120-160 ng/µL in a total volume of 20 µL (Table 6, column DNA Yield)[4].

[4] The split marker fragments are not purified from gel and template DNA (pyrG, hygB, Ble , argB , and nicB genes, respectively) used for amplification of the split marker might remain present in the next steps. We therefore include control transformations with both split markers separately. As no transformants are obtained in the transformation with only one flank (data not shown), the purification of the split marker fragment is not required, but is optional.

Table 6. Overview of templates and primers used in Fusion PCR reactions to obtain split marker fragments.

Split Marker 1					Split Marker 2					Transformed
Template		Primer		Product	Template		Primer		Product	Strain
1	2	Forward	Reverse	Yield (µg)	1	2	Forward	Reverse	Yield (µg)	
5'amyR	PCR hph	amyRP7f	hygP9r	2.5	3'amyR	PCR hph	hygP8f	amyRP10r	2.6	MA234.1
5'amyR	PCR BLE	amyRP7f	phleoP7r	2.6	3'amyR	PCR BLE	phleoP6f	amyRP10r	2.6	MA234.1
5'amyR	PCR pyrG	amyRP7f	AOpyrG15r	2.8	3'amyR	PCR pyrG	AOpyrGP14f	amyRP10r	2.8	MA169.4
5'amyR	PCR argB	amyRP7f	argBnidP8r	3.0	3'amyR	PCR argB	argBnidP7f	amyRP10r	3.2	JN1.17
5'amyR	PCR nicB	amyRP7f	nicBnidP8r	3.3	3'amyR	PCR nicB	nicBnidP7f	amyRP10r	2.5	JN4.2

3.2.4 Transformation of split marker fragments to Aspergillus niger Δku70 strains

1. Split marker fragments were combined and transformed to different *A. niger* strains (Table 6, column Transformed strain), according to Arentshorst *et al*. 2012. Results of these transformations are shown in Figure 4.
2. As a control, also separate split marker fragments were transformed. None of the separately transformed split marker fragments yielded any transformants (data not shown).
3. Four transformants were purified for each selection marker tested[5]. For purification protocol, see Arentshorst *et al*. 2012.
4. After the second purification, all purified transformants were tested for growth on MM + starch (Fig.4). All transformants analyzed showed a *ΔamyR* phenotype.
5. Purified transformants can be further analyzed by isolating genomic DNA, followed by both Southern blot analysis and diagnostic PCR (Arentshorst *et al*. 2012).

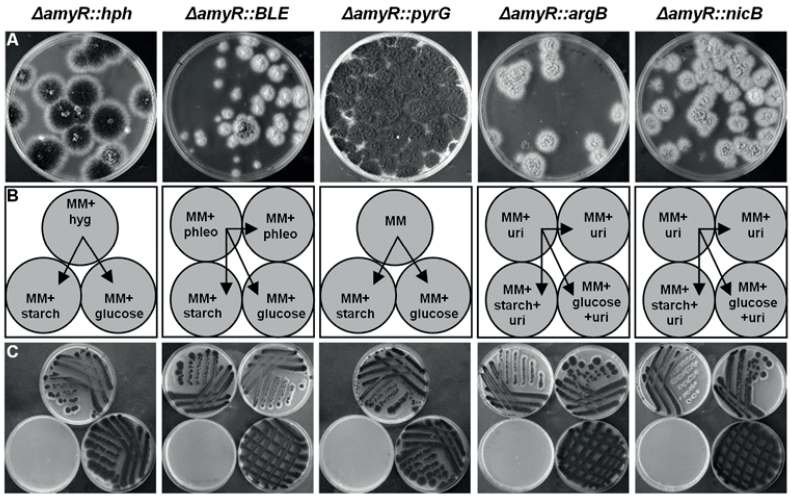


Figure 4. Phenotypic analysis of putative *amyR* disruptant strains using five different selection markers (*hph* = hygromycin resistance; BLE = phleomycin resistance; *pyrG* = uridine requiring; *argB* = arginine requiring; *nicB* = nicotinamide requiring). A) Transformation plates after transforming split marker fragment combinations for each of the five *amyR* deletion cassettes to the relevant recipient strain (Table 6). B, C) Purified transformants were analysed for their ability to grow on starch. The inability to grow on starch is indicative for the deletion of the *amyR* gene.

[5] Only the sporulating transformants on the phleomycin transformation plate (see Fig. 4) can grow on MM + phleomycin. The non-sporulating transformants do not grow, and are probably transient transformants, in which the split marker fragments have not integrated into the genome.

3.2.5 Transformation of split marker fragments to Aspergillus niger wt strains

For some experimental set-ups, it is preferred to analyze gene deletions in a *ku70* wild type strain. In order to show that the split marker approach also can be applied to a wild type (*ku70* plus) strain, both *A. niger* strains AB4.1 (Van Hartingsveldt *et al.* 1987) (*pyrG*$^{-}$) and MA169.4 (*Δku70, pyrG*$^{-}$) were transformed with *ΔamyR::pyrG* split marker fragments. After purification and screening on MM + starch, 25 out of 60 AB4.1-transformants (41%) showed a *ΔamyR* phenotype[6]. For MA169.4, 39 out of 40 transformants (98%) showed a *ΔamyR* phenotype. This result clearly shows that the split marker approach can also be used to make gene deletions in a wt background instead of a *Δku70* background.

[6] The percentages of Homologous Recombination (HR) for the *amyR* gene are very high (41% for wt, 98% for *Δku70*). Usually we find 5-10% HR for wt, and 80-100% for *Δku70Meyer, V., M. Arentshorst, A. El-Ghezal, A. C. Drews, R. Kooistra et al., 2007 Highly efficient gene targeting in the Aspergillus niger kusA mutant. J Biotechnol 128: 770-775, ibid. ibid.*(Meyer et al. 2007).

Acknowledgments

Jing Niu was supported by a grant from the China Scholarship Council. The research group of A.F.J. Ram is part of the Kluyver Centre for Genomics of Industrial Fermentation which is supported by the Netherlands Genomics Initiative.

References

Arentshorst, M., A. F. Ram and V. Meyer, 2012 Using non-homologous end-joining-deficient strains for functional gene analyses in filamentous fungi. Methods Mol Biol 835**:** 133-150.

Bos, C. J., A. J. Debets, K. Swart, A. Huybers, G. Kobus *et al.*, 1988 Genetic analysis and the construction of master strains for assignment of genes to six linkage groups in *Aspergillus niger*. Curr Genet 14**:** 437-443.

Buxton, F. P., D. I. Gwynne and R. W. Davies, 1985 Transformation of *Aspergillus niger* using the argB gene of *Aspergillus nidulans*. Gene 37**:** 207-214.

Carvalho, N. D., M. Arentshorst, M. Jin Kwon, V. Meyer and A. F. Ram, 2010 Expanding the ku70 toolbox for filamentous fungi: establishment of complementation vectors and recipient strains for advanced gene analyses. Appl Microbiol Biotechnol 87**:** 1463-1473.

de Ruiter-Jacobs, Y. M., M. Broekhuijsen, S. E. Unkles, E. I. Campbell, J. R. Kinghorn *et al.*, 1989 A gene transfer system based on the homologous *pyrG* gene and efficient expression of bacterial genes in *Aspergillus oryzae*. Curr Genet 16**:** 159-163.

Fairhead, C., B. Llorente, F. Denis, M. Soler and B. Dujon, 1996 New vectors for combinatorial deletions in yeast chromosomes and for gap-repair cloning using 'split-marker' recombination. Yeast 12**:** 1439-1457.

Goswami, R. S., 2012 Targeted gene replacement in fungi using a split-marker approach. Methods Mol Biol 835**:** 255-269.

Kuck, U., and B. Hoff, 2010 New tools for the genetic manipulation of filamentous fungi. Appl Microbiol Biotechnol 86**:** 51-62.

Lenouvel, F., P. J. van de Vondervoort and J. Visser, 2002 Disruption of the *Aspergillus niger* argB gene: a tool for transformation. Curr Genet 41**:** 425-432.

Meyer, V., 2008 Genetic engineering of filamentous fungi--progress, obstacles and future trends. Biotechnol Adv 26**:** 177-185.

Meyer, V., M. Arentshorst, A. El-Ghezal, A. C. Drews, R. Kooistra *et al.*, 2007 Highly efficient gene targeting in the *Aspergillus niger* kusA mutant. J Biotechnol 128**:** 770-775.

Nielsen, M. L., L. Albertsen, G. Lettier, J. B. Nielsen and U. H. Mortensen, 2006 Efficient PCR-based gene targeting with a recyclable marker for *Aspergillus nidulans*. Fungal Genet Biol 43**:** 54-64.

Ninomiya, Y., K. Suzuki, C. Ishii and H. Inoue, 2004 Highly efficient gene replacements in *Neurospora* strains deficient for nonhomologous end-joining. Proc Natl Acad Sci U S A 101**:** 12248-12253.

Petersen, K. L., J. Lehmbeck and T. Christensen, 1999 A new transcriptional activator for amylase genes in *Aspergillus*. Mol Gen Genet 262**:** 668-676.

Punt, P. J., R. P. Oliver, M. A. Dingemanse, P. H. Pouwels and C. A. van den Hondel, 1987 Transformation of Aspergillus based on the hygromycin B resistance marker from *Escherichia coli*. Gene 56**:** 117-124.

Punt, P. J., and C. A. van den Hondel, 1992 Transformation of filamentous fungi based on hygromycin B and phleomycin resistance markers. Methods Enzymol 216**:** 447-457.

van Hartingsveldt, W., I. E. Mattern, C. M. van Zeijl, P. H. Pouwels and C. A. van den Hondel, 1987 Development of a homologous transformation system for *Aspergillus niger* based on the *pyrG* gene. Mol Gen Genet 206**:** 71-75.

Verdoes, J. C., P. J. Punt, P. van der Berg, F. Debets, A. H. Stouthamer *et al.*, 1994 Characterization of an efficient gene cloning strategy for *Aspergillus niger* based on an autonomously replicating plasmid: cloning of the *nicB* gene of *A. niger*. Gene 146**:** 159-165.

Chapter 3

A set of isogenic auxotrophic strains for constructing multiple gene deletion mutants and parasexual crossings in *Aspergillus niger*

Jing Niu, Mark Arentshorst, Felix Seelinger, Arthur F.J. Ram, Jean Paul Ouedraogo

Published in Arch Microbiol. 2016 Nov; 198(9): 861-868

Abstract

To construct a set of isogenic auxotrophic strains in *Aspergillus niger* suited for creating multiple gene deletion mutants and executing parasexual crossings, we have combined mutations in genes involved in colour pigmentation (*fwnA*, and *olvA*) with well selectable auxotrophic markers (*pyrG*, *nicB*, *argB* and *adeA*). All markers, except for the *pyrG* marker, were introduced by targeted deletion, omitting UV mutagenesis of the strains. *Aspergillus oryzae* orthologous genes of the *argB*, *nicB* and *adeA* markers were used as heterologous selection markers and all markers were shown to complement to respective auxotrophic *A. niger* mutants. A quadruple auxotrophic marker was further constructed for suitable multiple gene deletions. Genome sequencing of two auxotrophic colour mutants JN3.2 (*olvA::pyrG*, *argB::hygB*) and JN6.2 (*olvA::pyrG*, *nicB::hygB*) revealed four SNPs between them in non-coding regions, indicating a high level of isogenicity between both strains. The availability of near isogenic complementary auxotrophic colour mutants facilitates the selection of diploids and the isolation of haploid segregants from the diploid using the parasexual cycle.

Keywords: isogenic strains, auxotrophy, multiple markers, parasexual crossing

Introduction

Aspergillus niger has attracted considerable interest as cell factories for the production of organic compounds (citric acid and secondary metabolites) or (recombinant) proteins (Pel *et al.* 2007; Ward 2012; Andersen *et al.* 2013; Meyer *et al.* 2015). *A. niger* is not only an important cell factory, it also has become an important model system for fungal development (Krijgsheld *et al.* 2013; Wosten *et al.* 2013). System biology-based approaches in combination with targeted metabolic engineering techniques are important tools to study and optimize production processes (Jacobs *et al.* 2009; Caspeta and Nielsen 2013). With relative ease gene knock outs can be made using the *ku70* mutants (Meyer *et al.* 2007; Carvalho *et al.* 2010) in combination with split marker approach (Nielsen *et al.* 2006; Goswami 2012; Arentshorst *et al.* 2015b). Together with tools for controlled overexpression of genes using the tetracycline promoter system (Meyer *et al.* 2011), metabolic engineering can be efficiently performed. A limiting factor for metabolic engineering in *A. niger* is the limited number of isogenic auxotrophic mutants with multiple auxotrophic markers, in which multiple gene deletion mutants can be made quickly without the need to recycle the selection markers. Selection markers such as the *pyrG* marker or the *amdS* marker are counter selectable, but when multiple deletions need to be made, these markers need to be recycled, which is time consuming. To overcome this limitation, we have selected the *nicB* gene (encoding nicotinate mononucleotide pyrophosphorylase) (Verdoes *et al.* 1994), the *argB* gene (encoding ornithine carbamoyltransferase) (Lenouvel *et al.* 2002) and the *adeA* gene (encoding phosphoribosylaminoimidazole-succinocarboxamidesynthase) (Ugolini and Bruschi 1996; Jin *et al.* 2004) of *A. niger* to construct near isogenic auxotrophic marker strains containing four auxotrophic markers (*pyrG, nicB, adeA* and *argB*). In combination with dominant selection markers such as hygromycin resistance (Punt and Van Den Hondel 1992), phleomycin resistance (Punt and Van Den Hondel 1992) and AmdS selection (Kelly and Hynes 1985), seven different markers are available for strain construction.

The lack of a sexual cycle in *A. niger* limits easy crossing of two strains to combine interesting properties or to construct double mutants. Despite the lack of a sexual cycle, the parasexual cycle can be used to combine genetic traits in *A. niger* (Pontecorvo *et al.* 1953; Swart *et al.* 2001a). The parasexual cycle includes the selection of a heterokaryon and subsequently the selection of a diploid strain. The frequency by which diploids are formed from a heterokaryotic mycelium in *A. niger* is very low and selection of diploids can be accomplished by crossing strains that have complementary auxotrophic and complementary

spore colour markers. Only when a diploid is formed, the resulting colony will produce solely black conidiospores which can be easily detected by eye. The genes encoding proteins involved in spore melanin production in *A. niger* have been identified (Jorgensen *et al.* 2011). Several studies, mainly conducted by Bos *et al*., have reported on the isolation of *A. niger* colour and auxotrophic mutants (see for review (Swart *et al.* 2001b)). However, most of these mutants were isolated by UV treatment. Although carried out with caution and relative high survival rates, unwanted random mutations are inevitable, leading to possible growth defects. By targeted deletion of spore colour genes and auxotrophies, we constructed a set of near-isogenic strains suitable for parasexual crossings. We performed genome sequencing of two auxotrophic colour mutants and confirmed the near isogenicity between these auxotrophic mutants.

Materials and methods

Strains and growth conditions

The *Aspergillus niger* strains used in this study are listed in Table 1. Auxotrophic strains are deposited at the Fungal Genetic Stock Centre. *A. niger* strains were grown on minimal medium (MM) (Bennet and Lasure 1991) or on complete medium (CM) consisting of minimal medium with the addition of 5 g l^{-1} yeast extract and 1 g l^{-1} casamino acids. When required, 10 mM uridine, 200 µg/ml L-arginine, 2.5 µg/ml nicotinamide, 100 µg/ml hygromycin or 40 µg/ml phleomycin was added. Adenine was directly added from the solid stock to the medium to a final concentration of 200 mg/L after autoclaving and dissolved by mixing. Fluoroacetamide (FAA) and 5-fluoro-orotic acid (5-FOA) counter selection was performed as described (Carvalho *et al.* 2010) to remove the *amdS* marker and the *pyrG* marker respectively.

Table 1. Strains used in this study

Name	Genotype/description	Reference/source
N402	*cspA1,* derivative of N400	Bos et al. 1988
A. oryzae	ATCC16868	-
MA169.4	*kusA::amdS, pyrG*$^-$	Carvalho et al. 2010
MA100.1	*cspA1, fwnA::hygB, kusA::amdS, pyrG*$^-$	Jørgensen et al., 2011
AW8.4	*cspA1, olvA::AopyrG* in MA169.4	Jørgensen et al. 2011
JN3.2	*argB*::*hygB*, *olvA*::*AopyrG* (derived from AW8.4)	This study
JN6.2	*nicB*::*hygB*, *olvA*::*AopyrG* (derived from AW8.4)	This study
JN1.17.1	*argB*::*hygB* in MA169.4	This study
OJP3.1	*nicB*::*phleo* in MA169.4	This study
OJP1.1	*adeA::pyrG* in MA169.4	This study
MA322.2	*ku70::amdS, nicB::AopyrG* in MA169.4	This study
MA323.1	*ku70::amdS, ΔnicB*$^-$*, pyrG*$^-$	This study
MA328.2	*ku70::amdS, ΔnicB*$^-$*, adeA::AopyrG*	This study
MA329.1	*ku70::amdS, ΔnicB*$^-$*, ΔadeA*$^-$*, pyrG*$^-$	This study
MA334.2	*ku70::amdS, ΔnicB*$^-$*, ΔadeA*$^-$*, argB::AopyrG*	This study
MA335.3	*ku70::amdS, ΔnicB*$^-$*, ΔadeA*$^-$*, ΔargB*$^-$*, pyrG*$^-$	This study

Molecular biological techniques

Transformation of *A. niger* and chromosomal DNA isolation of *A. niger* and *A. oryzae* was performed according to (Meyer *et al.* 2010). Southern blot analysis was performed according to (Sambrook and Russell 2001). α-^{32}P-dCTP-labelled probes were synthesized using the Rediprime II kit (Amersham, GE Healthcare), according to the instructions of the manufacturer. Restriction and ligation enzymes were obtained from Thermo Scientific and used according to instructions of the manufacturer. PCR was performed with Phire Hot Start II DNA polymerase or Phusion DNA polymerase (Thermo Scientific). Sequencing was performed by Macrogen.

Construction of plasmids and deletion cassettes

The deletion cassettes for the *argB*, *nicB* and *adeA* genes of *A. niger* were constructed with the *hygB*, *phleo* and *pyrG* selection marker respectively. The plasmid used to disrupt the *argB* gene (An14g03400) with the hygromycin selection marker was constructed as follows: ~ 0.8 kb DNA fragments flanking the *argB* ORF were amplified by PCR using N402 genomic DNA as template, with primers listed in Supplementary Table 1. The PCR products were cloned into pJet1.2 (Thermo Scientific). The 5'flank of *argB* was excised from pJet1.2 using *Kpn*I/*Hind*III and inserted into the same site of pBlueScript II sk(+) to obtain plasmid pJN3.3. Subsequently, pJN3.3 was digested with *Hind*III/*Not*I and used in a three way ligation with the 3'flank of *argB* excised from pJet1.2 using *Xho*I/*Not*I and the 3 kb *Hind*III/*Xho*I fragment containing the *hygB* gene, obtained from plasmid pΔ2380 (Damveld *et al.* 2008), resulting in the *argB* disruption plasmid pJN4.5. The *argB* gene deletion cassette was amplified by PCR using pJN4.5 DNA as template with primer argBKO1 and argBKO4 and the purified linear PCR fragment was used for subsequent transformation to *A. niger* strain MA169.4 (*ku70*$^{-}$, *pyrG*$^{-}$) to give JN1.19.1 (*ku70*$^{-}$*, pyrG*$^{-}$, *ΔargB::hygB*) or to *A. niger* strain AW8.4 (*ku70*$^{-}$, *ΔolvA::AOpyrG*), resulting in JN3.2 (*ku70*$^{-}$*, ΔolvA::AOpyrG*, *ΔargB::hygB*).

The same approach was used to construct the disruption cassettes of the *nicB* gene (An11g10910) of *A. niger* with either the phleomycin or hygromycin marker. The DNA fragments flanking the *nicB* ORF were amplified from N402 genomic DNA, with primers listed in Supplementary Table 1. After cloning in pJet1.2, the 5'flank of *nicB* was isolated as a *Kpn*I/*Xho*I fragment and inserted into *Kpn*I/*Xho*I opened pBlueScriptII SK(+) to obtain plasmid pJN8.1. Subsequently, the 1.9 kb *Xho*I-*Hind*III fragment containing *phleo* expression cassette, obtained from plasmid pMA299, or the 3.1 kb *Xho*I-*Hind*III fragment containing *hygB* expression cassette, obtained from plasmid pΔ2380 (Damveld *et al.* 2008), together with the *Hind*III/*Not*I isolated 3'flank of *nicB*, were ligated into *Xho*I/*Not*I opened pJN8.1, resulting in the *nicB::phleo* disruption plasmid pJN10.1 or *nicB::hygB* disruption plasmid pJN9.1. The *nicB* gene deletion cassettes were amplified by PCR using pJN10.1 or pJN9.1 as template with primer NicBKO1 and NicBKO4 and used for transformation to *A. niger* strain MA169.4 (*ku70*$^{-}$, *pyrG*$^{-}$) to give OJP3.1 (*ku70*$^{-}$*, pyrG*$^{-}$, *ΔnicB::phleo*) or to *A. niger* strain AW8.4 (*ku70*$^{-}$, *ΔolvA::AOpyrG*), resulting in JN6.2 (*ku70*$^{-}$*, ΔolvA::AOpyrG*, *ΔnicB::hygB*).

To construct the disruption cassette of *adeA* gene (An11g10150), the flanking regions of the gene were amplified by PCR from N402 genomic DNA with primers Fw_adeA_5' and

Rev_adeA_5' to obtain the 0.9 kb 5'flanking region and Fw_adeA_3'and Rev_adeA_3' to obtain the 0.7 kb 3'flanking region (Supplementary Table 1). The 1.8 kb *A. nidulans pyrG* selection marker was amplified by PCR from the plasmid *pCRpyrGAN* (Ouedraogo *et al.* 2015) with the primers Fw_pyrG_adeA and Rev_pyrG_adeA which contain complementary sequence of Rev_adeA-5'and Fw_adeA-3' respectively (Supplementary Table 1). The *adeA::Anid_pyrG* deletion cassette was obtained by a fusion PCR reaction of the three purified PCR products, followed by cloning of the 3.4 kb fusion PCR product into pJet1.2, resulting to plasmid pOJP1. Proper deletion of the *nicB, adeA* and *argB* genes was confirmed by Southern blot analysis (Supplementary Figures 1-3).

For complementation studies, *argB*, *nicB* and *adeA* genes, including their promoter and terminator regions, were amplified from wild type *A. oryzae* and *A. niger* genomic DNA with appropriate primer pairs described in the Supplementary Table 1. The respective complementing gene fragments were cloned into pJet1.2 (Thermo Scientific) and sequenced (Table 2). The plasmids pOJP5 (pJet1.2_*Anig.argB*), pOJP4 (pJet1.2_*Anig.nicB*), pOJP3 (pJet1.2_*Anig.adeA*), pJN29 (pJet1.2_*Aory.argB*), pJN30 (pJet1.2_*Aory.nicB*) and pJN31 (pJet1.2_*Aory.adeA*) were used to complement the respective auxotrophic mutants.

Recyclable split marker strategy for creation of a strain with multiple auxotrophies

To construct an *A. niger* strain with multiple auxotrophies, it was necessary to use a recyclable split marker approach. Therefore, auxotrophic marker specific direct repeats (DR) surrounding the *AOpyrG* selection marker were introduced by PCR. By selecting on 5-FOA the *AOpyrG* marker was removed. The recyclable split marker approach is outlined in Figure 1; see Supplementary Table 1 for primer sequences. Strain MA169.4 (*ku70⁻, pyrG⁻*) was used as starting strain to first delete the *nicB* gene, and subsequently, *adeA* and the *argB* marker. All strains containing single, double, triple and the quadruple auxotrophic strain (MA335.5 *ku70⁻,* Δ*nicB,* Δ*adeA,* Δ*argB, pyrG⁻*) are listed in Table 1. Correct integration of split marker fragments and successful loop out of the *AOpyrG* was confirmed by Southern blot analysis for all strains and shown for MA335.5 in Supplementary Figures 1-3).

Table 2. Plasmids used in this study

Name	Description	Reference/source
pJN3.3	5'flank of *argB* in pBluescript II SK(+)	This study
pΔ2380	*ΔugmB::hygB* deletion cassette	Damveld et al. 2008
pJN4.5	pBluescript_*argB*::*hygB*	This study
pJN8.1	5'flank of *nicB* in pBluescript II SK(+)	This study
pMA299	pBluescript_*phleo*	This study
pJN10.1	pBluescript_*nicB::phleo*	This study
pCRpyrGAN	Containing the full gene of *A. nidulans pyrG*	DuPont Bioscience
pOJP1	pJet1.2_*adeA::pyrG*	This study
pOJP5	pJet1.2_*Anig.argB*	This study
pOJP4	pJet1.2_*Anig.nicB*	This study
pOJP3	pJet1.2_*Anig.adeA*	This study
pJN29	pJet1.2_*Aory.argB*	This study
pJN30	pJet1.2_*Aory.nicB*	This study
pJN31	pJet1.2_*Aory.adeA*	This study
pAO4-13	Containing full *pyrG* gene of *A. oryzae*	(De Ruiter-Jacobs *et al.* 1989)

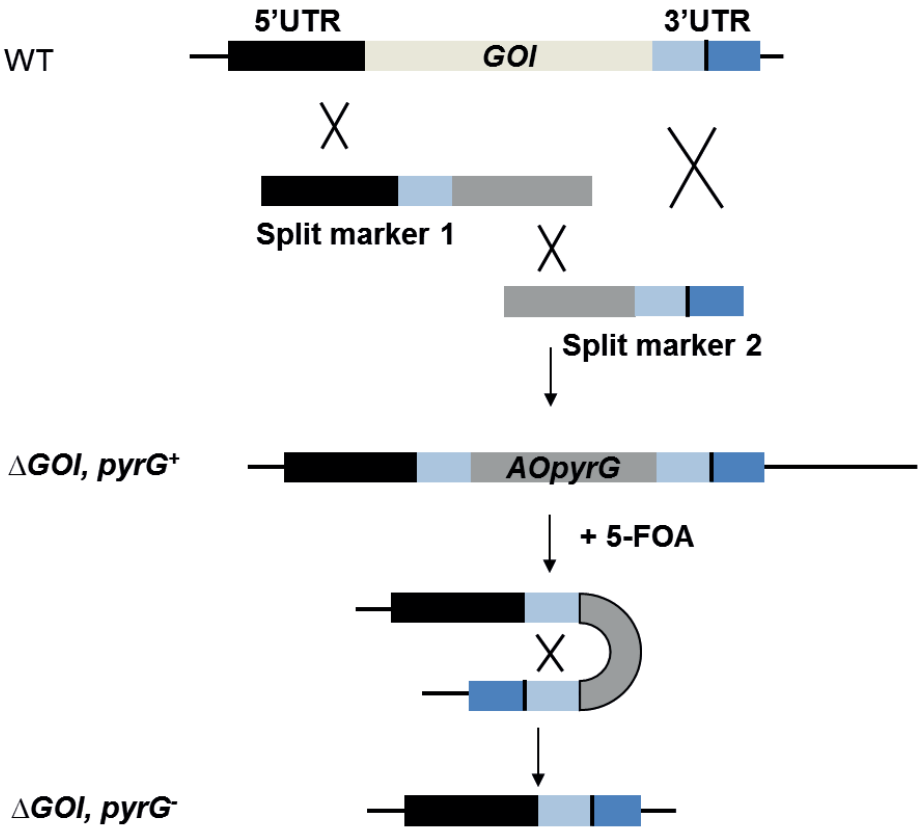


Figure 1. Schematic representation of the recyclable split marker approach for multiple gene deletion mutants. Deletion of the gene of interest (GOI) by split marker approach with recycling of the *Aspergillus oryzae pyrG* marker. The split marker fragments 1 and 2 are used during transformation to knock out the GOI by homologous recombination which generates a uridine prototroph *(pyrG^{+})* strain. The *pyrG* marker is subsequently looped out by 5-FOA selection and the resulting *pyrG*$^{-}$ strain is suitable for a second gene deletion with the *pyrG* marker. The split marker approach is described previously (Arentshorst et al., 2014).

A. niger parasexual cycle

Heterokaryon formation and selection for diploids was performed as described (Pontecorvo *et al.* 1953; Pontecorvo and Sermonti 1953). Segregation of diploids by benomyl was performed essentially as described (Bos *et al.* 1988) with slight modifications (Niu *et al.* 2016).

Sequencing and analysis

Genome sequencing of JN3.2 (*olvA::pyrG*, *argB::hygB*) and JN6.2 (*olvA::pyrG*, *nicB::hygB*) was performed using NGS platform (Illumina GA) as described (Park *et al.* 2014). Sequencing was performed at ServiceXS, Leiden, The Netherlands. SNPs between JN3.2 and JN6.2 were identified using *A. niger* strain ATCC1015 (http://genome.jgi-psf.org/pages/search-for-genes.jsf?organism=Aspni5) as reference genome. For each SNP it was verified whether the SNP was in a predicted protein encoding region using the *A. niger* 3.0 genome at JGI using the SNP coordinates (Park et al., 2014).

Results and Discussion

Construction and characterization of *argB*, *nicB* and *adeA* auxotrophic mutants

Deletion constructs *nicB::hygB*, *argB::phleo* and *adeA::pyrG* were transformed to strain MA169.4 (*ku70⁻, pyrG⁻)* and hygromycin, phleomycin resistant or uridine prototrophic transformants were obtained and purified. Proper deletion of the respective markers was verified by diagnostic PCRs (data not shown) and by testing the growth on MM-plates containing the relevant supplements. As shown in Figure 2, the *nicB*, *argB* and *adeA* mutants required the addition of the nicotinamide, L-arginine or adenine to allow growth.

To determine the minimal concentrations of nicotinamide, arginine or adenine for full supplementation, spores of the auxotrophic mutants were spotted on plates containing a concentration series of the respective supplements and the growth was monitored over time. The results in Figure 2 show the necessity to use at least 800 mg/L of arginine and 1.25 mg/L of nicotinamide to fully supplement the Δ*argB* and Δ*nicB* strains respectively. For the Δ*adeA* mutant, the supplementation test shows that a concentration of adenine between 10 and 50 mg/L lead to the accumulation of red pigment. At this range of adenine concentrations, the strain is not forming conidia. Further analysis showed that this red pigment was accumulated into the vacuole when cells were grown in liquid medium (data not shown). To fully supplement the Δ*adeA* mutant, at least 150 mg/L of adenine in the growth medium was required.

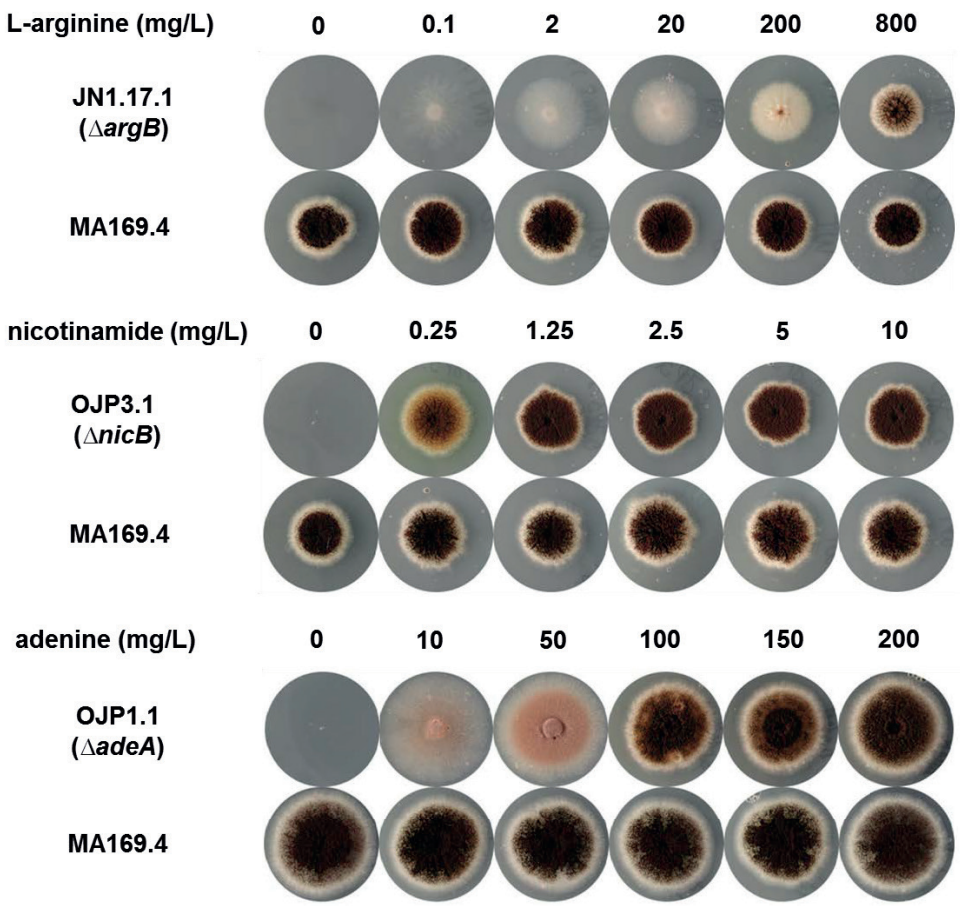


Figure 2. Supplementation test of the auxotrophic A. niger mutants. 10 µl of a spore stock ($1x10^7$ conidia/ml) of each auxotrophic strain and the parental strain (MA169.4) was inoculated on an MM plate without and with serial concentrations of the respective supplement and incubated at 30 °C for 3 days for arginine and nicotinamide supplementation test and for 4 days for adenine supplementation test.

Construction and characterization of a quadruple auxotrophic strain (Δ*nicB*, Δ*argB*, Δ*adeA*, *pyrG*⁻)

We have constructed a quadruple auxotrophic strain based on the recyclable split marker approach described in Figure 1 and in materials and methods. This approach allows iterative construction of gene knockouts in *A. niger* by subsequent recycling of the *pyrG* marker using counter selection on 5-FOA, due to the presence of the direct repeated sequences flanking the selection marker. The proper deletion and absence of ectopic copies of the deletion cassettes in the quadruple auxotrophic strain MA335.3 was confirmed by Southern blot analysis (Supplemental Figures 1-3) and characterized by the inability to growth in absence of arginine, nicotinamide, adenine or uridine (Figure 3). This quadruple auxotrophic strain offers the possibility to delete multiple genes without the need to recycle the selection marker.

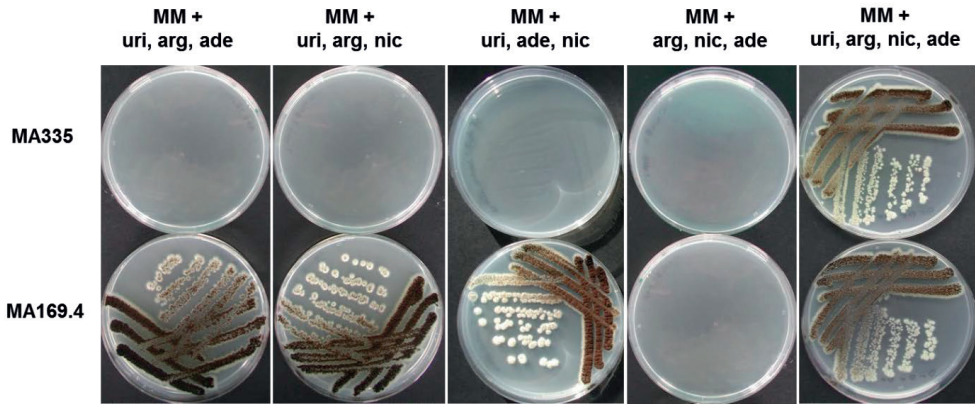


Figure 3. Growth analysis of the quadruple auxotrophic *A. niger* strain. MA335.3 *(ΔnicB, ΔargB, ΔadeA, pyrG⁻)* was plated on solid MM with and without the different supplements at 30 °C and growth was analyzed after 3 days. The parental strain MA169.4 was taking along the analysis for comparison.

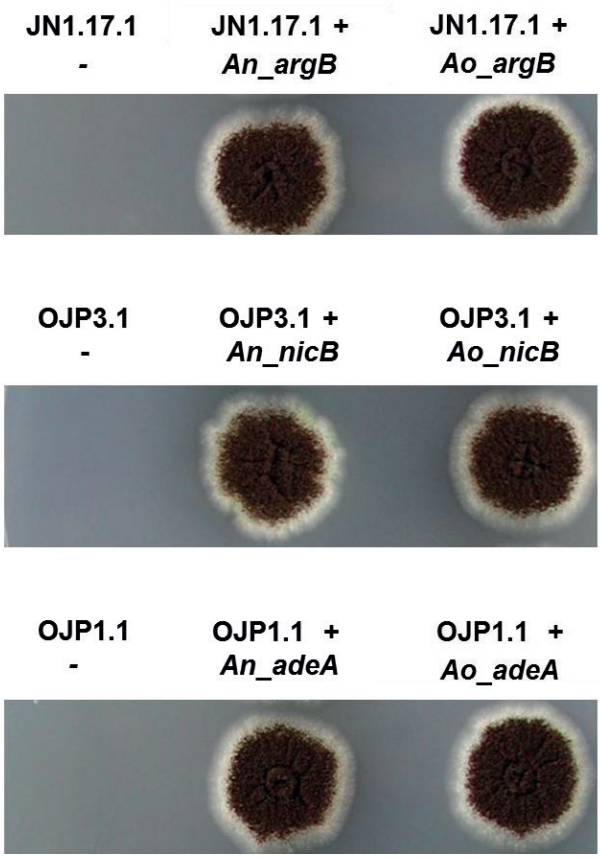


Figure 4. Growth analysis of the complemented transformants. Spores of JN1.17.1 (*ΔargB, pyrG*⁻) OJP3.1 (*ΔnicB, pyrG*⁻) and OJP1.1 (*ΔadeA*) and complemented strains were spotted on selective medium to test complementation of the *argB*, *nicB* and *adeA* respectively from *A. niger* (An) or *A. oryzae* (Ao). Pictures were taken after 3 days of growth at 30 °C.

The *nicB*, *argB* and *adeA* genes from *A. oryzae* are suitable markers for *A. niger* transformation

To prove that auxotrophic mutants can be complemented by heterologous and homologous markers, DNA fragments containing the *argB*, the *nicB* and the *adeA* genes from *A. oryzae and A. niger*, including their promoters and 3' untranslated sequences, were used for the complementation of the respective *A. niger* auxotrophic mutants. Protoplasts of JN1.17.1 (*ΔargB::hygB*), OJP3.1 (*ΔnicB::phleo*) and OJP1.1 (*ΔadeA::pyrG*) were transformed with plasmids containing the corresponding marker genes from *A. oryzae* or *A. niger*. Transformants were obtained for the *A. oryzae* heterologous markers, which demonstrated

that *nicB, argB* and *adeA* of *A. oryzae* complemented the auxotrophy and therefore are suitable markers for *A. niger* transformations. As expected, also all *A. niger* genes (*argB*, *nicB* and *adeA)* were able to complement the respective auxotrophic *A. niger* mutants. The obtained transformants were further analysed to determine whether the *A. oryzae* marker also complemented the auxotrophies. As shown in Figure 4, all heterologous genes complement similarly to the homologous *A. niger* genes. A heterologous marker for gene disruption experiments is preferred as it reduces the homologous integration of the marker gene in the disruption cassette at the homologous site. We have compared the DNA sequence of the different genes markers of *A. niger* to those of *A. oryzae* by BLASTN (http://blast.ncbi.nlm.nih.gov/) using standard settings. The identity of the coding regions between the different gene markers was 73.3%, 72.0% and 77.8% for *argB*, *nicB* and *adeA* genes respectively. These values are comparable to the value obtained when comparing the *pyrG* genes markers of both *Aspergillus* species. The *pyrG* gene of *A. oryzae* is identical to the *pyrG* gene of *A. niger* at 78.6% and has been so far successfully used to transform *A. niger* and vice versa (Mattern *et al.* 1987; Carvalho *et al.* 2010). It should be noted that complementation analysis in the Δ*ku70* background is not efficient because of the low frequencies of ectopic integration the complementing fragment. To circumvent this limitation, we constructed a curable *ku70* deletion strategy (Carvalho et al., 2010). The presence of *ku70* repeats around the AmdS selection marker used to disrupt the *ku70* gene allow efficient loop out of the AmdS marker via fluoro-acetamide counter selection as described (Arentshorst *et al.* 2012). An alternative method for easy complementation, which omits the need for curing the *ku70* locus, is the use of a second auxotrofic marker which can be used to target the complementing gene to this locus. For the *pyrG* marker, an efficient gene targeting method has recently become available (Arentshorst *et al.* 2015a) with allows targeted integration when the complementing fragment is cloned in the *pyrG* targeting vector. For example, one could start with a *nicB*$^{-}$, *pyrG*$^{-}$ strain and use the *nicB* selection marker for initial deletion of the gene of interest, followed by a complementation experiment in which the complementing fragment is cloned in the *pyrG* targeting vector which is that transformed to the deletion strain.

Isogenic auxotrophic colour mutants for parasexual crossing in *A. niger*

Combining mutations by crossing strains is a powerful genetic tool for strain construction. In *Aspergillus nidulans* this method is well established and used in many studies to construct double mutants (Todd *et al.* 2007). The lack of a sexual cycle *in A. niger* has limited the use

of crossings to combine mutations. However, the use of the parasexual cycle in *A. niger* (Pontecorvo *et al.* 1953) has been used extensively for linkage studies in *A. niger* and can be used to combine mutations (Bos *et al.* 1988). Straightforward crossing in *A. niger* requires complementing auxotrophies to select for a heterokaryotic mycelium and preferably colour makers to select for a diploid strain. The frequency by which *A. niger* forms diploids is generally very low (1 in 10^6 to 10^7 spores) and diploids are not easily detected if wild-type strains are used that produce black conidia. By using complementing colour markers a diploid can be selected as only this diploid will produce black spores, whereas a heterokaryotic mycelium will produce a mix of heterogeneously coloured spores (Pontecorvo *et al.* 1953). By combining colour mutants (*fwnA* and *olvA*) with complementary auxotrophic markers such as *pyrG*, *nicB* or *argB*, heterokaryons and diploids can be easily selected. We constructed several auxotrophic colour mutant strains including MA100.1 (*fwnA::hygB*, *pyrG*$^-$, JN3.2 (*olvA::pyrG*, *argB::hygB*), and JN6.2 *(olvA::pyrG, nicB::hygB* (Table 1). In a recently conducted study JN3.2 has been used for parasexual crossings to obtain haploid segregants (Niu *et al.* 2016). With these segregants, a bulk segregant analysis was performed to identify SNPs that are closely linked or responsible for the mutant phenotypes (Niu *et al.* 2016).

To test the isogenicity between two auxotrophic colour mutants JN3.2 (*olvA::pyrG, argB::hygB*) and JN6.2 (*olvA::pyrG, nicB::hygB*), the genomes of these strains were sequenced and compared to the genome of the reference ATCC strain. In total, 155 SNPs were found for JN3.2 and JN6.2 respectively when compared to the ATCC reference strain (Suppl. Table 2). Two SNPs were found to be specific for JN3.2, and two SNPs were specific for JN6.2. None of them were found in predicted open reading frames (Table 3), demonstrating that JN3.2 and JN6.2 are likely to have no mutation affected its phenotype and that they are near isogenic.

Table 3. SNP comparison JN6.2 and JN3.2

	Position	Allel ATCC	JN6.2	JN3.2	Details mutation
chr_1_2	726573	T	T	C	Intergenic
chr_3_4	45864	T	T	A	Intergenic
chr_8_2	2725044	G	A	G	Intergenic
chr_8_2	2725045	T	A	T	Intergenic

In conclusion, new auxotrophic strains carrying targeted deletions in the *argB, nicB* and *adeA* genes of *A. niger* were constructed. The orthologous genes *argB*, *nicB* and *adeA* of *A. oryzae* complemented the arginine, nicotinamide and adenine auxotrophic mutants similar to the endogenous genes and are therefore suitable selection markers for *A. niger* transformations. The quadruple auxotrophic strain MA335 (*argB*$^-$*, nicB*$^-$, *adeA*$^-$ and *pyrG*$^-$) allows rapid deletion of multiple genes deletion without the need to recycle selection markers. The targeted deletion of auxotrophic markers instead of selection of auxotrophic strains after UV mutagenesis significantly reduces the occurrence of mutations as genome sequencing of two auxotrophic mutants (JN3.2 and JN6.2) revealed only four SNP between them.

Acknowledgements

Jing Niu is supported by a grant from the China Scholarship Council.

References

Andersen, M. R., J. B. Nielsen, A. Klitgaard, L. M. Petersen, M. Zachariasen *et al.*, 2013 Accurate prediction of secondary metabolite gene clusters in filamentous fungi. Proc Natl Acad Sci U S A 110**:** E99-107.

Arentshorst, M., E. L. Lagendijk and A. F. Ram, 2015a A new vector for efficient gene targeting to the pyrG locus in *Aspergillus niger*. Fungal Biology and Biotechnology 2:2.

Arentshorst, M., J. Niu and A. F. J. Ram, 2015b Efficient Generation of *Aspergillus niger* Knock Out Strains by Combining NHEJ Mutants and a Split Marker Approach, pp. 263-272 in *Genetic Transformation Systems in Fungi*, edited by M. A. van den Berg and K. Maruthachalam. Springer International Publishing, Cham, Switzerlands.

Arentshorst, M., A. F. Ram and V. Meyer, 2012 Using non-homologous end-joining-deficient strains for functional gene analyses in filamentous fungi. Methods Mol Biol 835**:** 133-150.

Bennet, J. W., and L. L. Lasure, 1991 *More Gene Manipulations in Fungi*. Academic Press, San Diego.

Bos, C. J., A. J. Debets, K. Swart, A. Huybers, G. Kobus *et al.*, 1988 Genetic analysis and the construction of master strains for assignment of genes to six linkage groups in *Aspergillus niger*. Curr Genet 14**:** 437-443.

Carvalho, N. D., M. Arentshorst, M. Jin Kwon, V. Meyer and A. F. Ram, 2010 Expanding the ku70 toolbox for filamentous fungi: establishment of complementation vectors and recipient strains for advanced gene analyses. Appl Microbiol Biotechnol 87**:** 1463-1473.

Caspeta, L., and J. Nielsen, 2013 Toward systems metabolic engineering of *Aspergillus* and *Pichia* species for the production of chemicals and biofuels. Biotechnol J 8**:** 534-544.

Damveld, R. A., A. Franken, M. Arentshorst, P. J. Punt, F. M. Klis *et al.*, 2008 A novel screening method for cell wall mutants in *Aspergillus niger* identifies UDP-galactopyranose mutase as an important protein in fungal cell wall biosynthesis. Genetics 178**:** 873-881.

de Ruiter-Jacobs, Y. M., M. Broekhuijsen, S. E. Unkles, E. I. Campbell, J. R. Kinghorn *et al.*, 1989 A gene transfer system based on the homologous pyrG gene and efficient expression of bacterial genes in *Aspergillus oryzae*. Curr Genet 16**:** 159-163.

Goswami, R. S., 2012 Targeted gene replacement in fungi using a split-marker approach. Methods Mol Biol 835**:** 255-269.

Jacobs, D. I., M. M. Olsthoorn, I. Maillet, M. Akeroyd, S. Breestraat *et al.*, 2009 Effective lead selection for improved protein production in *Aspergillus niger* based on integrated genomics. Fungal Genet Biol 46 Suppl 1**:** S141-152.

Jin, F. J., J. Maruyama, P. R. Juvvadi, M. Arioka and K. Kitamoto, 2004 Adenine auxotrophic mutants of *Aspergillus oryzae*: development of a novel transformation system with triple auxotrophic hosts. Biosci Biotechnol Biochem 68**:** 656-662.

Jorgensen, T. R., J. Park, M. Arentshorst, A. M. van Welzen, G. Lamers *et al.*, 2011 The molecular and genetic basis of conidial pigmentation in *Aspergillus niger*. Fungal Genet Biol 48**:** 544-553.

Kelly, J. M., and M. J. Hynes, 1985 Transformation of *Aspergillus niger* by the amdS gene of Aspergillus nidulans. EMBO J 4**:** 475-479.

Krijgsheld, P., R. Bleichrodt, G. J. van Veluw, F. Wang, W. H. Muller *et al.*, 2013 Development in *Aspergillus*. Stud Mycol 74**:** 1-29.

Lenouvel, F., P. J. van de Vondervoort and J. Visser, 2002 Disruption of the *Aspergillus niger argB* gene: a tool for transformation. Curr Genet 41**:** 425-432.

Mattern, I. E., S. Unkles, J. R. Kinghorn, P. H. Pouwels and C. A. M. J. J. Vandenhondel, 1987 Transformation of *Aspergillus oryzae* Using the *Aspergillus niger pyrG* Gene. Molecular & General Genetics 210**:** 460-461.

Meyer, V., M. Arentshorst, A. El-Ghezal, A. C. Drews, R. Kooistra *et al.*, 2007 Highly efficient gene targeting in the *Aspergillus niger* kusA mutant. J Biotechnol 128**:** 770-775.

Meyer, V., M. Fiedler, B. Nitsche and R. King, 2015 The Cell Factory *Aspergillus* Enters the Big Data Era: Opportunities and Challenges for Optimising Product Formation. Adv Biochem Eng Biotechnol 149**:** 91-132.

Meyer, V., A. F. J. Ram and P. J. Punt, 2010 Genetics, genetic manipulation and approaches to strain improvement of filamentous fungi, pp. 318-329 in *Manual of industrial microbiology and biotechnology*, edited by R. Baltz, A. Demain and Eds. Wiley, NY.

Meyer, V., F. Wanka, J. van Gent, M. Arentshorst, C. A. van den Hondel *et al.*, 2011 Fungal gene expression on demand: an inducible, tunable, and metabolism-independent expression system for *Aspergillus niger*. Appl Environ Microbiol 77**:** 2975-2983.

Nielsen, M. L., L. Albertsen, G. Lettier, J. B. Nielsen and U. H. Mortensen, 2006 Efficient PCR-based gene targeting with a recyclable marker for *Aspergillus nidulans*. Fungal Genet Biol 43**:** 54-64.

Niu, J., M. Arentshorst, P. D. S. Nair, Z. Y. Dai, S. E. Baker *et al.*, 2016 Identification of a Classical Mutant in the Industrial Host *Aspergillus niger* by Systems Genetics: LaeA Is Required for Citric Acid Production and Regulates the Formation of Some Secondary Metabolites. G3-Genes Genomes Genetics 6**:** 193-204.

Ouedraogo, J. P., M. Arentshorst, I. Nikolaev, S. Barends and A. F. Ram, 2015 I-SceI-mediated double-strand DNA breaks stimulate efficient gene targeting in the industrial fungus *Trichoderma reesei*. Appl Microbiol Biotechnol.

Park, N. I., I. Y. Choi, B. S. Choi, Y. S. Kim, M. Y. Lee *et al.*, 2014 EST sequencing and gene expression profiling in *Scutellaria baicalensis*. EXCLI J 13**:** 392-400.

Pel, H. J., J. H. de Winde, D. B. Archer, P. S. Dyer, G. Hofmann *et al.*, 2007 Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. Nat Biotechnol 25**:** 221-231.

Pontecorvo, G., J. A. Roper and E. Forbes, 1953 Genetic Recombination without Sexual Reproduction in *Aspergillus niger*. Journal of General Microbiology 8**:** 198-210.

Pontecorvo, G., and G. Sermonti, 1953 Recombination without Sexual Reproduction in *Penicillium chrysogenum*. Nature 172**:** 126-127.

Punt, P. J., and C. A. van den Hondel, 1992 Transformation of filamentous fungi based on hygromycin B and phleomycin resistance markers. Methods Enzymol 216**:** 447-457.

Sambrook, J., and D. W. Russell, 2001 *molecular cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York

Swart, K., A. J. Debets, C. J. Bos, M. Slakhorst, E. F. Holub *et al.*, 2001a Genetic analysis in the asexual fungus *Aspergillus niger*. Acta Biol Hung 52**:** 335-343.

Swart, K., D. van Heemst, M. Slakhorst, F. Debets and C. Heyting, 2001b Isolation and characterization of sexual sporulation mutants of *Aspergillus nidulans*. Fungal Genet Biol 33**:** 25-35.

Todd, R. B., M. A. Davis and M. J. Hynes, 2007 Genetic manipulation of *Aspergillus nidulans*: meiotic progeny for genetic analysis and strain construction. Nat Protoc 2**:** 811-821.

Ugolini, S., and C. V. Bruschi, 1996 The red/white colony color assay in the yeast *Saccharomyces cerevisiae*: epistatic growth advantage of white ade8-18, ade2 cells over red ade2 cells. Curr Genet 30**:** 485-492.

Verdoes, J. C., P. J. Punt, P. van der Berg, F. Debets, A. H. Stouthamer *et al.*, 1994 Characterization of an efficient gene cloning strategy for *Aspergillus niger* based on an autonomously replicating plasmid: cloning of the nicB gene of A. niger. Gene 146**:** 159-165.

Ward, O. P., 2012 Production of recombinant proteins by filamentous fungi. Biotechnol Adv 30**:** 1119-1139.

Wosten, H. A., G. J. van Veluw, C. de Bekker and P. Krijgsheld, 2013 Heterogeneity in the mycelium: implications for the use of fungi as cell factories. Biotechnol Lett 35**:** 1155-1164.

Supporting information

The supplementary material of this chapter are available via

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5040738/ and comprises the following:

Supplementary Table 1. Primers used in this study.

Supplemental Table 2. Comparison of SNP positions between ATCC1028 and JN3.2/JN6.2.

Supplemental Fig. 1. Verification of the *nicB* deletion in OJP3.1 (*nicB*::*phleo* in MA169.4) and quadruple auxotrophic strain MA335. A) Schematic representation of the *nicB* locus of the wild-type and after *nicB::phleo* deletion and after loop out of the *pyrG* gene after disrupting *nicB* in MA335. Predicted sizes of the DNA fragment hybridizing with the indicated probe are shown. B) Southern blot analysis of genomic DNA of MA169.4 (lane 1), OJP3.1 (lane 2) and MA335.4. (lane 3), and MA335.4 (lane 4). Left panel: agarose gel stained with ethidium bromide. MW = molecular weight marker size (in kb) are indicated. Right panel: Southern blot after hybridization with *nicB* probe.

Supplemental Fig. 2. Verification of the *adeA* deletion in OJP1.1 (*adeA*::*pyrG* in MA169.4) and quadruple auxotrophic strain MA335. A) Schematic representation of the *adeA* locus of the wild-type and after *adeA::pyrG* deletion and after loop out of the *pyrG* gene after disrupting *adeA* in MA335. Predicted sizes of the DNA fragment hybridizing with the indicated probe are shown. B) Southern blot analysis of genomic DNA of MA169.4 (lane 1), MA335.3 (lane 2), MA335.4 (lane 3), and OJP1.1 (lane 4). Left panel: agarose gel stained with ethidium bromide. MW = molecular weight marker size (in kb) are indicated. Right panel: Southern blot after hybridization with *adeA* probe.

Supplemental Fig. 3 Verification of the *argB* deletion in JN17.1 (*arg*::*hygB* in MA169.4) and quadruple auxotrophic strain MA335. A) Schematic representation of the *argB* locus of the wild-type and after *argB::hygB* deletion and after loop out of the *pyrG* gene after disrupting *argB* in MA335. Predicted sizes of the DNA fragment hybridizing with the indicated probe are shown. B) Southern blot analysis of genomic DNA of MA169.4 (lane 1), JN1.17.1 (lane 2), MA335.3 (lane 3), and MA335.4 (lane 4). Left panel: agarose gel stained with ethidium bromide. MW = molecular weight marker size (in kb) are indicated. Right panel: Southern blot after hybridization with *argB* probe.

Chapter 4

The interaction of induction and repression mechanisms in the regulation of galacturonic acid-induced genes in *Aspergillus niger*

Jing Niu, Tim G. Homan, Mark Arentshorst, Ronald P. de Vries, Jaap Visser and Arthur F.J. Ram

Published in Fungal Genetics and Biology. 2015 September; 82: 32-42

Abstract

Aspergillus niger is an important industrial fungus expressing a broad spectrum of pectinolytic genes. The main constituent of pectin, polygalacturonic acid (PGA), is degraded into galacturonic acid (GA) by the combined activity of endo- and exo-polygalacturonases some of which are specifically induced by GA. The regulatory mechanisms that control the expression of genes encoding PGA-degrading enzymes are not well understood. Based on available genome-wide expression profiles from literature, we selected five genes that were specifically induced by GA. These genes include three exo-polygalacturonases (*pgaX, pgxB* and *pgxC*), a GA transporter (*gatA*), and an intracellular enzyme involved in GA metabolism (*gaaB*). These five genes contain a conserved motif (5'-TCCNCCAAT-3') in their promoter regions, which we named GARE (galacturonic acid-responsive element). Promoter deletion studies and site-directed mutagenesis of the conserved motif of the *pgaX* gene showed that the conserved element is required for GA-mediated induction. A set of promoter reporter strains was constructed by fusing the promoter region of the five above-mentioned genes to the *amdS* reporter gene. Expression of the *amdS* gene is quantitatively correlated with ability to utilize acetamide as an N-source, hence higher expression of *amdS* improves growth of the strain on acetamide and therefore can be used as an *in vivo* reporter for gene expression. Growth analysis of the reporter strains indicated that four genes (*pgaX, pgxB, pgxC*, and *gatA*) are specifically induced by GA. The *in vivo* promoter reporter strains were also used to monitor carbon catabolite repression control. Except for *gaaB,* all promoter-reporter genes analysed were repressed by glucose in a glucose concentration-dependent way. Interestingly, the strength of glucose repression was different for the tested promoters. CreA is important in mediating carbon catabolite repression as deletion of the *creA* gene in the reporter strains abolished carbon catabolite repression for most promoters. Interestingly, the *pgxC* promoter was still repressed by glucose even in the *creA* null background, suggesting a role for alternative repression mechanisms. Finally, we showed that low concentrations of GA are required to induce gene expression of *pgaX*, *pgxB*, and *pgxC* even under derepressing conditions. The results obtained are consistent with a model in which a GA-specific transcription factor is activated by GA or a GA-derivative, which binds to the conserved motif, possibly in combination with the HAP-complex, to drive GA-specific gene expression.

Keywords: pectin, pectinolytic genes, gene regulation, reporter genes, acetamidase, carbon catabolite repression, CreA

Introduction

Pectin represents a group of complex heterogeneous polysaccharides that are primarily present in the middle lamella of plant cell walls. The backbone consists mainly of α-1,4-linked D-galacturonic acid residues part of which are methyl-esterified. The polygalacturonic acid backbone can be further modified to form substructures with increasing complexity. Pectin is classified into four substructures: homogalacturonan (HGA), xylogalacturonan (XGA), rhamnogalacturonan-I (RG-I), and rhamnogalacturonan-II (RG-II) (Mohnen 2008). RG-II is the most complex structure and composed of up to twelve different types of sugars in >20 different linkages.

Aspergillus niger is an important industrial micro-organism and used as a cell factory for the production of enzymes and organic acids (Pel *et al.* 2007; Andersen *et al.* 2011). Genome mining has revealed a large array of extracellular pectinolytic enzymes in *A. niger* (Coutinho *et al.* 2009; Martens-Uzunova and Schaap 2009). The majority of the pectinolytic enzymes that act on the polygalacturonic acid backbone belong to the GH28 family of glycoside hydrolases, and include endo- and exo-polygalacturonases (Bussink *et al.* 1992a; Bussink *et al.* 1992b; Benen *et al.* 1996; Parenicova *et al.* 1998; Parenicova *et al.* 2000; Martens-Uzunova *et al.* 2006).

Metabolism of GA in fungi is well described and involves four enzymatic reactions to convert GA into glycerol and pyruvate. The genes encoding these enzymes (*gaaA*, *gaaB*, *gaaC*, and *gaaD* have been identified and the biochemical properties of the enzymes have been determined (Kuorelahti *et al.* 2005; Kuorelahti *et al.* 2006; Liepins *et al.* 2006; Hilditch *et al.* 2007; Mojzita *et al.* 2010; Wiebe *et al.* 2010; Zhang *et al.* 2011; Kuivanen *et al.* 2012). Specific sugar transporters that are able to transport GA over the plasma membrane have recently been identified and characterized in Neurospora crassa (Benz *et al.* 2014) as well as in *A. niger* (Sloothaak et al. 2014) and *Botrytis cinerea* (Zhang *et al.* 2014). The transporters identified in these studies are phylogenetically related and probably represent a subfamily of GA-specific transporters (Zhang *et al.* 2014).

Several studies have focused on understanding the transcriptional regulation of pectinolytic genes (Bussink *et al.* 1990; Bussink *et al.* 1992a; Maldonado and Strasser De Saad 1998; De Vries *et al.* 2002) and have shown that most pectinolytic genes are specifically induced by GA. In combination with a genome-wide expression study in *A. niger* (Martens-Uzunova and Schaap 2008) a conserved promoter element has been identified that is present in the promoter region of GA-induced genes. A promoter deletion study of the *pgaII* gene (encoding an endo-polygalacturonase) showed that this element is important for high level

expression of *pgaII* (Bussink *et al.* 1992b). The element (5'-TCCNCCAAT-3') is present in genes encoding extracellular enzymes that are specific for polygalacturonic acid hydrolysis (both exo- and endo-activities), in putative GA-transporter genes, and in genes encoding the enzymes for GA metabolism (*gaaA* to *gaaD*) (Martens-Uzunova and Schaap 2008). We propose to name this element GARE for galacturonic acid-responsive element. Benen and co-workers identified a second region (CCCTGA), which is present in promoters of many pectinolytic genes that might be important in activation of pectinolytic gene expression and has been named PecR (Benen *et al.* 1996; Coutinho *et al.* 2009). The GARE motif (5'-TCCNCCAAT-3') is similar to the binding site of the HAP2/3/4 complex. The HAP complex (as it is named in *Saccharomyces cerevisiae*) is a conserved multimeric transcription factor that regulates gene expression by binding to the consensus sequence CCAAT. In the filamentous fungi, the HAP complex consists of four subunits (HapB/C/E and HapX). The complex is required for the regulation of gene expression not only related to carbon or nitrogen source utilization (Van Heeswijck and Hynes 1991; Kato *et al.* 1997; Kato *et al.* 1998; Steidl *et al.* 1999) but also to other cellular processes like secondary metabolite production (Litzka *et al.* 1996), iron homeostasis (Hortschansky *et al.* 2015), and oxidative stress responses (Thon *et al.* 2010).

Several studies have shown that GA-induced genes are repressed by the presence of glucose (Bussink *et al.* 1991; De Vries *et al.* 2002) through carbon catabolite repression control (CCR). In filamentous fungi, the C2H2 type transcription factor CreA/CRE1, which is related to Mig1/Mig2/Mig3 proteins that mediate glucose repression in *S. cerevisiae* (Westholm *et al.* 2008), has been shown to act as a repressor mediating CCR (Dowzer and Kelly 1991; Ruijter and Visser 1997). CreA/CRE1 binds to the promoters of the respective target genes via the consensus motif 5'-SYGGRG-3' to repress expression. Disruption mutants in *creA* are viable in *A. nidulans* (Shroff *et al.* 1997) and *A. niger* (Yuan *et al.* 2006) and can be used to analyse transcriptional regulation under derepressed conditions.

For the efficient degradation and utilization of polymeric substrates, a synergistic and coordinated expression of the hydrolysing enzymes, sugar transporters, and enzymes involved in the intracellular metabolism is required. In filamentous fungi several substrate-specific transcription factors have been identified, which function as key regulators to control gene expression in response to the presence of a particular substrate (Kowalczyk *et al.* 2014). However, a specific GA-responsive transcription factor has not yet been identified. We suggest that an as yet unidentified transcription factor is responsible for GA-dependent

induction of the genes encoding the extracellular PGA-degrading enzymes as well as activation of genes that encode the intracellular enzymes involved in GA metabolism.

In this study, we show the importance of the GARE motif (5'-TCCNCCAAT-3') for GA-induced gene expression. We also constructed promoter-reporter constructs to analyse the regulation of these promoters *in vivo*. Using these reporter strains, we show that induction and repression of GA-induced genes is differentially fine-tuned in response to inducing and repressing conditions.

Material and methods

2.1 Strains and growth conditions

The *A. niger* strains used in this study are listed in Table 1. Strains were grown in liquid or on solidified (by addition of 2 % agar) minimal medium (MM), which contained 7 mM KCl, 8 mM KH_2PO_4, 70 mM $NaNO_3$, 2 mM $MgSO_4$ (pH adjusted to pH 5.5) as described by Bennett and L.L. 1991 (Bennett and L.L. 1991). MM was supplemented with a specific carbon source to a final concentration of 50 mM as indicated. Standard complete medium (CM) was also used and consisted of MM supplemented with 0.1% casamino acids and 0.5% $w.v^{-1}$ yeast extract and 50 mM glucose. MM agar plates containing 10 mM acetamide as sole nitrogen source were made as previously described (Arentshorst *et al.* 2012). Transformation of *A. niger* strains was also carried out as described in (Arentshorst *et al.* 2012). Targeted integration of reporter constructs to the *pyrG** locus was carried out as described previously by (Van Gorcom and Van Den Hondel 1988) or through via a recently developed *pyrG*-targeting vector pMA334 (named *pyrG**)* (Arentshorst *et al.* 2015). Fungal chromosomal DNA isolation was performed as described by Meyer *et al*. 2010 (Meyer *et al.* 2010). Two strains were used as a recipient for transformation and include AB4.1 and MA299.2 (Table 1). MA299.2 is derived from the *Δku70* mutant strain MA70.15 (*kusA::amdS*, *pyrG*$^-$) after curing the *amdS* marker by fluoro-acetamide selection (Arentshorst *et al.* 2012). The resulting strain MA299.2 (*kusA*$^-$, *pyrG*$^-$) was checked for proper removal of the *amdS* marker by diagnostic PCR. The growth of transformants was assayed by point-inoculating 4 µl of spore suspension (1×10^5 spores/µl) in the centre of the agar plate and incubating the plates for 7 days at 30°.

Table 1. Strains used in this study

Strain	Genotype	Reference
N402	*cspA*	Bos et al. 1988
AB4.1	*pyrG*$^{-}$ derivatie of N402	van Hartingsveldt et al. 1987
XY1.1	Δ*creA::pyrG* in AB4.1	Yuan et al. 2006
MA70.15	*ΔkusA::amdS, pyrG*$^{-}$ in AB4.1	Meyer et al. 2007
MA299.2	*ΔkusA*$^{-}$*, pyrG*$^{-}$ derivative of MA70.15	This study
MA211.15	*PpgaX(1203)-uidA-pyrG** in AB4.1[1,2]	This study
MA212.2	*PpgaX(1005)-uidA-pyrG** in AB4.1	This study
MA213.10	*PpgaX(688)-uidA-pyrG** in AB4.1	This study
MA214.14	*PpgaX(409)-uidA-pyrG** in AB4.1	This study
MA215.8	*PpgaX(380)-uidA-pyrG** in AB4.1	This study
CR01.6	*PpgaX (1210)* (TTTACCTTT)-*uidA-TtrpC-pyrG** in AB4.1[3]	This study
CR02.8; CR02.9	*PpgaX (688)* (TTTACCTTT)-*uidA-TtrpC-pyrG** in AB4.1	This study
CR03.1	*PpgaX (1210)* (TTTACCTTT)-*amdS-TtrpC-pyrG** in AB4.1	This study
CR04.2	*PpgaX (688)* (TTTACCTTT)-*amdS-TtrpC-pyrG** in AB4.1	This study
CR05.31	*PpgaX (1210)* (TTTACCAAT)-*uidA-TtrpC-pyrG** in AB4.1	This study
CR06.4; CR06.5	*PpgaX (688)* (TTTACCAAT)-*uidA-TtrpC-pyrG** in AB4.1	This study
CR09.1; CR09.20	*PpgaX (1210) (*TCCACCTTT)-*uidA-TtrpC-pyrG** in AB4.1	This study
CR10.1; CR10.6	*PpgaX (688)* (TCCACCTTT)-*uidA-TtrpC-pyrG** in AB4.1	This study
JC1.5; JC1.7; JC1.8	*PpgaX(1210)-amdS-pyrG***in MA299.2[4]	This study
JC3.6; JC3.7; JC3.9	*PpgxB(1250)-amdS-pyrG*** in MA299.2	This study
JC4.2; JC4.3; JC4.4	*PpgxC(1201)-amdS-pyrG*** in MA299.2	This study
JC5.1; JC5.2; JC5.3	*PgaaB(1194)-amdS-pyrG*** in MA299.2	This study
JC6.6; JC6.7; JC6.8	*PgatT(1208)-amdS-pyrG*** in MA299.2	This study
JN11.2	*PabfA(1266)-amdS--pyrG** in AB4.1	This study
JN29.2	*creA::hygB* in JC1.5	This study
JN31.3	*creA::hygB* in JC3.6	This study
JN32.1	*creA::hygB* in JC4.2	This study
JN33.1	*creA::hygB* in JC5.1	This study
JN34.3	*creA::hygB* in JC6.6	This study
JN16.1	*creA::hygB* in JN11.2	This study

[1] number in brackets indicates length of the promoter region.
[2] *pyrG** refers to targeting integration method of DNA constructs using the *pyrG* targeting method described by van Gorcom et al., 1988
[3] Underlined regions indicate mutations compared to the wild-type motif (TCCACCAAT).
[4] *pyrG*** refers to targeting integration method to the *pyrG* locus described by Arentshorst et al., 2015 (Arentshorst *et al.* 2015).

2.2 General DNA procedures

PCR amplifications were performed using phusion DNA polymerase (Finnzymes) and were carried out according to the manual provided by the manufacturer. *Escherichia coli* strain DH5α was used for all recombinant DNA experiments. *E. coli* was transformed using standard heat shock protocols as described by (Inoue *et al.* 1990). All endonuclease restriction enzymes were purchased from Fermentas or Sigma. DNA sequence analysis was carried out by Macrogen, Korea. Ligations were performed using the Rapid DNA ligation kit (Fermentas).

2.3 Construction of recombinant vectors

The plasmids used in this study are listed in Supplemental Table 1. *PgaX* promoter fragments with different lengths (1203, 1005, 688, 409 and 380 bp) were PCR amplified (see Supplemental Table 2 for primers) and cloned as *Not*I/*Eco*RI fragments in PagsA(30-bp)-*uidA-pyrG** (Damveld *et al.* 2005) to give pMA211, pMA212, pMA213, pMA214 and pMA215. Promoter fragments mutated in the CCNCCAAT box were constructed in either the large *pgaX* promoter fragment (1202 bp) or the short promoter (409 bp). Mutations were introduced by using oligonucleotides (Supplemental Table 2). PCR fragments were again cloned as *Not*I/*Eco*RI fragments in PagsA(30-bp)-*uidA-pyrG** to construct the pCRO plasmids (Supplemental Table 1).

Promoter regions of *pgxB*, *pgxC*, *gaaB*, *gatA* and *abfA* were amplified by PCR using the primers listed in Supplemental Table 2 and using *A. niger* N402 genomic DNA as template, and ligated into pJET1.2 (Fermentas). To construct the *pgaX-amdS* reporter constructs with the mutated CCNCCAAT box, primers Mut1-for and Mut1-rev were used (Supplemental Table 2). The promoter fragments containing the mutations were subsequently cloned in pJET1.2. From these intermediate vectors the promoter fragments were re-isolated as *Sal*I/*Eco*RI and subsequently ligated into the 3,662 bp *Sal*I/*Eco*RI-digested backbone of vector PagsA(2010)-AmdS (Damveld *et al.* 2008) to construct the *amdS* reporter constructs. For the *PabfA-amdS* reporter construct, the *pyrG** gene was cloned into this vector as a *Xba*I fragment to produce the final vector (Table 2). For the construction of promoter reporter plasmids for targeted integration through homologous integration via pMA334, a *Not*I site was introduced at the *Sal*I restriction site by inserting a *Not*I restriction site oligonucleotide (Supplemental Table 2) at the *Sal*I site. The promoter fragments were subsequently isolated as *Not*I fragments and cloned into pMA334 and transformed to MA299.2. Proper integration of the constructs was verified by Southern blot. For each promoter reporter construct three independent transformants were analysed by Southern blot and analysed phenotypically. For each construct a single strain was selected and the *creA* gene was deleted using the split marker method (Arentshorst *et al.* 2015) using primers listed in Supplemental Table 2 for amplification of *creA* 5' and 3' split marker fragments. For each reporter construct, three *creA* mutants were purified, analysed by diagnostic PCR and analysed phenotypically. All PCR amplified promoter regions and final constructs were verified by sequencing (Macrogen).

Table 2. Putative GA-responsive and putative CreA-binding site in GA induced genes.

Gene number	Gene name	Promoter length (bp)	Promoter element	
			GA-responsive element (CCNCCAAT)	CreA binding site (SYGGRG)
An12g07500	*pgaX*	1210	-389	3 (+) [*a] 5 (-)
An16g05390	*gaaB*	1194	-327	10 (+) 5 (-)
An03g06740	*pgxB*	1250	-299	7 (+) 3 (-)
An02g12450	*pgxC*	1201	-279	5 (+) 5 (+)
An14g04280	*gatA*	1208	-334[*b]	2 (+) 4 (-)
An01g00330	*abfA*	1266	Not present	8 (+) 6 (-)

*[a] + and - indicates a putative CreA in the forward (+) or reverse (-) orientation. The exact positions of the different putative *creA* binding sites are given in Supplemental Table 3.
*[b] GARE site in *gatA* does not perfectly matches consenses (TCCAGCAAT).

2.4 Analysis of transformants expressing GUS

The P*pgaX-uidA*-T*trpC-pyrG** vector variants (pMA211-pMA215 and pCRO plasmids) were transformed to *A. niger* AB4.1. Transformants were screened based on glucuronidase activity as described (Damveld *et al.* 2005). For each construct, two transformants expressing GUS from the *pyrG* locus were selected for further analysis. To perform expression analysis of *uidA* and to perform β-glucuronidase enzymatic assay (GUS assay). RNA and protein was isolated from mycelium that was grown for 18 hours at 30˚C in 100 ml CM supplemented with 50 mM GA as the sole carbon source. After growth, mycelium was harvested, frozen in liquid nitrogen, grinded and subsequently used for RNA isolation using TRIzol (Damveld *et al.* 2005) or for protein isolation. For protein isolation 200 mg of mycelium was resuspended in 1 ml Z-buffer (see above) to which a protease inhibitor cocktail (Sigma P8215) was freshly added. Fifty µl of the supernatant was mixed with 1 ml Z-buffer and 200 µl p-nitrophenyl α-D-glucuronic acid solution (PNP-Gluc, Sigma N-1377). Samples were incubated at 37˚C for 15 min or one hour after which the reaction was stopped by addition of two ml 100 mM Na_2CO_3. The ODs at 415 and 550 nm were determined using an Ultraspec 2100 pro (Amersham) spectrophotometer. The specific activity was determined in triplicate as described by (Roberts *et al.* 1989).

Results

3.1. A conserved motif in promoters of galacturonic acid-induced genes is required for *pgaX* induction

Genome-wide expression analysis of *A. niger* grown on various carbon sources has revealed specific induction of pectinolytic genes on galacturonic acid (GA) (Martens-Uzunova and Schaap 2008). These genes include among others exo-polygalacturonidases (*pgaX*, *pgxB*, *pgxC*), a putative galacturonic acid transporter (*gatA*) (Sloothaak *et al.* 2014), and genes involved in the intracellular metabolism of GA (*gaaA*, *gaaB* and *gaaC)*. Promoter analysis of these co-regulated genes identified a conserved motif comprising the consensus sequence 5'-YCCNCCAAT-3' (Martens-Uzunova and Schaap 2008) to which we will refer to as GARE (galacturonic acid responsive element) in this manuscript. To assess the importance of GARE, promoter deletion and site-directed mutagenesis studies were performed using the *pgaX* promoter as representative example. As schematically depicted in Figure 1, promoter deletions and promoter mutations were made in th*e pgaX* promoter and these promoter fragments were cloned in front of the β-glucuronidase (GUS) reporter. All promoter reporter constructs were targeted to the *pyrG* locus using the *pyrG** method (Van Gorcom and Van Den Hondel 1988) and Southern blots were performed to show single-copy, targeted integration of the reporter constructs (data not shown). For each reporter, two independently obtained single-copy transformants were analysed for *uidA* expression and β-glucuronidase activity when grown on GA. The promoter deletion studies showed that the conserved GARE motif is required for expression of the *pgaX* gene leading to GUS activity (Figure 1, and data not shown for the *uidA* expression analysis). As GARE (5'-TCCNCCAAT-3') contains a putative HAP-binding site (CCAAT), the role of the element was determined in more detail by introducing mutations in the specific motifs of the element. Mutations that abolish HAP complex binding (CCAAT to CCTTT mutations) eliminated the expression and induction of the *pgaX* gene. Importantly, also mutating the CC motif upstream of the CCAAT motif (CC to TT) abolished *pgaX* expression, indicating that it is not only the HAP binding site which is important for induction. Mutations in GARE were made either in the large *pgaX* promoter fragment (2012 bp) or the short *pgaX* promoter fragment (409), but in both cases no GUS activity could be detected. The results strongly suggest that both the CC motif and the CCAAT motif in the 5'-TCCACCAAT-3' box are necessary for *pgaX* expression. PecR (CCCTGA) is a second element which has been identified in promoter of GA-induced genes (Bussink *et al.* 1992b; Benen *et al.* 1996). The results presented in Fig.1 indicate that the pecR element is not necessary for *pgaX* induction by GA.

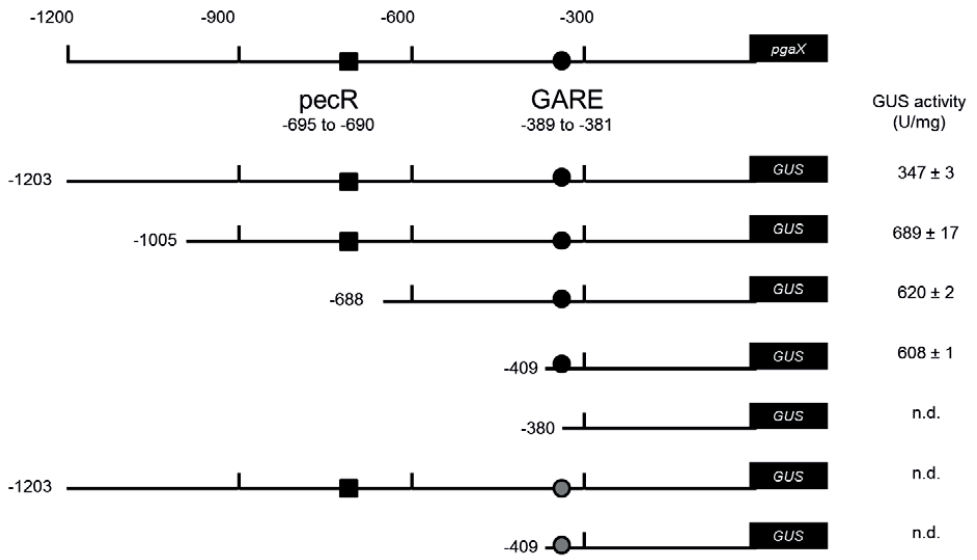


Figure 1. Schematic strategy to analyse the role of conserved promoter elements in the *pgaX* promoter. A) Promoter fragments with decreasing length were PCR-amplified and cloned in front of the β-glucuronidase (GUS) reporter. Mutations in the galacturonic acid-responsive element (GARE) in either the long promoter (1203 bp) or the shortened (409 bp) promoter version are indicated by the gray circle. For either the long or short promoter fragment three mutations were made (CCACCAAG to **TT**ACCAAG, CCACC**TT**G or **TT**ACC**TT**G) which all abolished expression if the *uidA* gene. On the right - GUS activity in mycelia of transformants grown on galacturonic acid as a sole carbon source.

3.2. Construction and analysis of *in vivo* promoter-reporter constructs to assess promoter activity

The acetamidase gene (*amdS*) of *A. nidulans* (Kelly and Hynes 1985) was used as a reporter for gene expression in *A. niger* (e.g. (Damveld *et al.* 2008; Punt *et al.* 2008) To analyse GA-specific expression *in vivo* on agar plates, we constructed several *amdS* reporter constructs, in which promoter sequences of GA-induced genes (*pgaX*, *pgxB*, *pgxC*, *gaaB* and *gatA*) were selected (Table 2). In addition, we included a promoter of *abfA* which is known to be an arabinose-induced gene (De Groot *et al.* 2003). The rationale of the assay is that expression of *amdS* from the various promoters can be assessed by the ability of the transformant to grow on medium containing acetamide as N-source. To allow comparison of promoter activity all promoter-reporter constructs were targeted to the *pyrG* locus either by the *pyrG** method (*abfA*) (Van Gorcom and Van Den Hondel 1988) or through a recently developed *pyrG*** targeting method ((*pgaX, pgxB, pgxC, gaaB* and *gatA*) (Arentshorst *et al.* 2015). Correct integration of the reporter constructs was verified by Southern blot analysis (data not shown). For each construct, three independent single-copy transformants with the expected integration

pattern were analysed for growth. Transformants were point-inoculated on minimal medium agar plates containing acetamide as a sole N-source and various carbon sources (Figure 2). Growth of the transformants containing the AmdS reporter fused to the promoter of *pgaX*, *pgxB*, and *pgxC*, was specific for GA, polygalacturonate (PGA) and pectin. On all other carbon sources these strains did not grow, indicating that *pgaX, pgxB, and pgxC* are specifically induced by GA. The growth of the transformant harbouring the *pgxC-amdS* reporter construct is reduced compared to the other strain, indicating that the expression from the *pgxC* promoter might be lower than expression from the other promoters. Growth of the *gaaB-amdS* reporter strain was less specific for GA, indicating that the *gaaB* gene is also expressed under non-inducing conditions. Growth analysis of *gatA-amdS* reporter suggests that the gene is not only expressed in the presence of GA, but also in the presence of rhamnose, arabinose, and fructose. As described in more detail in the next paragraph, the ability of the *gatA-amdS* reporter strain to grow on rhamnose, arabinose, and fructose is likely due to derepression. Growth of the *abfA-amdS* reporter strain nicely confirmed previous reports that *abfA* is specifically induced by arabinose (De Groot *et al.* 2003). The *abfA* promoter does not contain a GARE or pecR element. The *abfA-amdS* reporter strain also grows well on PGA and pectin indicating the *abfA* is also induced under these conditions. Whether the induction by pectin and PGA is AraR-dependent awaits further studies.

The *pgaX-amdS* reporter was also used to show the importance of the 5'-TCCACCAAT-3' motif. Mutating the GARE promoter in the wild type *pgaX* promoter (5'-TCCACCAAT-3' to 5'-TTTGTTGAT-3') and placing the mutated promoter in front of the *amdS* reporter abolished the ability of transformant containing the mutated reporter construct to grow on GA (data not shown). Collectively, by using both the *uidA* and *amdS* reporter constructs we show that *pgaX* induction is specific on GA and requires the GARE motif.

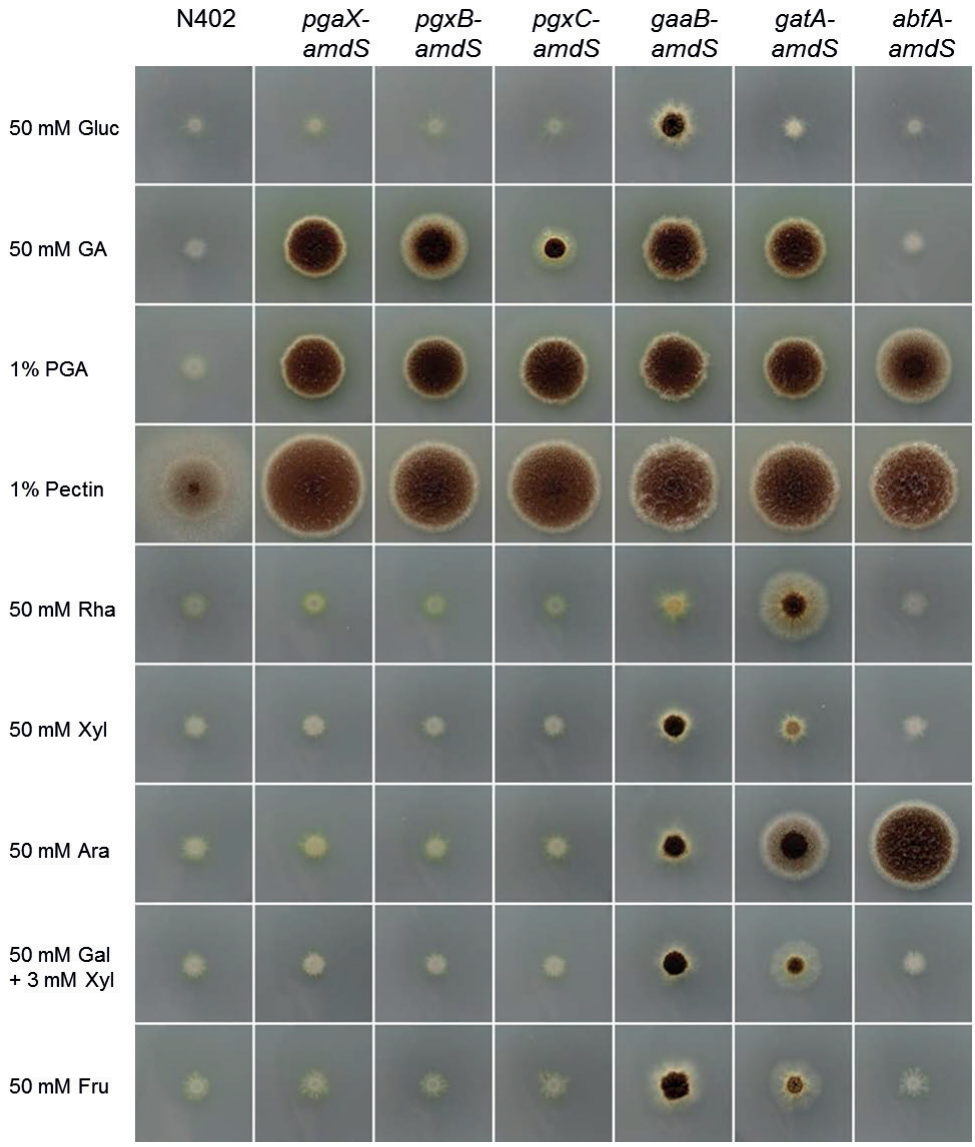


Figure 2. Growth analysis of promoter-reporter strains on various carbon sources. The wild-type (N402) and reporter strains were grown on MM-acetamide supplemented with 50 mM of a monomeric carbon source, or 1% w/v polymeric substrate.

3.3 Expression of galacturonic acid-induced genes is under carbon catabolite repression control and CreA-dependent

The reporter strains described above also offer the possibility to analyse *in vivo* the role of carbon catabolite repression on the GA-induced genes. First, *in silico* analysis of the promoter

sequences of all five reporter constructs for putative CreA binding sites (SYGGRG) showed that all promoters contain several putative CreA binding sites (Table 2 and Supplemental Table 3). To analyse the influence of glucose repression on the expression of the promoter-reporter genes, spores of the reporter strains were inoculated on MM-acetamide plates containing 50 mM GA and increasing concentrations of glucose (up to 50 mM) (Figure 3, columns A, C, E, G, I). Expression of *pgaX, pgaxB*, and *pgxC* was strongly repressed by the addition of glucose. Whereas expression of *pgxB* and *pgxC* was strongly repressed by the addition of 0.5 mM glucose, this concentration had less effect on *pgaX* expression. At glucose concentrations between 5 mM and 50 mM, however, *pgaX*, *pgxB*, and *pgxC* were all three fully repressed. The repression of *gatA* was only partial at the higher glucose concentrations (>5 mM) and the expression of *gaaB* was apparently not affected as growth of the reporter strains was not reduced on GA plates supplemented with glucose (Figure 4, column G).

To investigate whether CreA is responsible for carbon catabolite repression of GA-induced genes, the *creA* gene was deleted in the different AmdS-reporter strains. For efficient deletion of the *creA* gene in the reporter strains a *ku70* mutant (*kusA*) was used as a recipient strain (see material and methods). Deletion of *creA* in the reporters was very efficient (80-90%) and proper deletion of *creA* in all five reporter strains was verified through diagnostic PCR (data not shown). Growth and glucose repression of the strains were analysed as described above by inoculating spores on a GA/acetamide plate containing increasing concentrations of glucose. As shown in Figure 3 (columns B, D, F, H and J) deletion of *creA* abolished glucose repression in most of the reporter strains. Glucose repression of the *pgaX*, *pgxB and gatA* genes was completely lost (Figure 3; columns B, D, and J, respectively). Repression was still observed, although only at higher glucose concentrations (≥10 mM) in the *pgxC* reporter strain. The reduced growth of the *gaaB* reporter in the *creA* background at glucose concentration higher than 5 mM (compare in Figure 3, column G and H) is probably not directly related to *amdS* expression from the *gaaB* promoter but results from the general reduction in growth of the *creA* mutant compared to the wild-type strain as previously reported (Yuan *et al.* 2006). The analysis of the reporter strains indicates that for the *pgaX*, *pgxB* and *gatA* promoters CreA is required for glucose repression. The expression of *amdS* from the *pgxC* promoter is more complex as repression is still observed, only at high glucose concentration (Figure 3, column F). It will be of interest to discover the molecular basis for glucose repression in the *pgxC* gene at high concentrations as its repression seems to be independent of CreA.

3.4 GA-induced gene expression requires a specific activator

The promoter-*amdS* reporter strains in the *creA* mutant background also made it possible to assess the role of inducer molecules. In most cases, transcriptional regulation of networks involved in utilization of a specific carbon source is mediated via a network-specific transcription factor such as XlnR, AraR or AmyR (see introduction). These transcriptional activators are often activated via specific sugars derived from the substrate. For GA, it has been suggested that GA itself or a derivative of GA acts as an inducer (Mojzita *et al.* 2010). To establish if the presence of GA as an inducer was required to support growth, spores of the reporter strains in which the *creA* gene was deleted were point-inoculated on glucose medium (Figure 4, upper row). The reporter strains expressing *amdS* from the *pgaX*, *pgxB*, and *gaaB* promoters, still required the presence of GA as an inducer to activate gene expression. The *pgxC-amdS* reporter strain did not grow, indicating that the expression from the *pgxC* promoter is rather low, which was also observed in the wild-type background (Figure 2). The *gatA-amdS* reporter strain also grows on glucose/acetamide plates even without inducer indicating that derepression by deleting *creA* results in sufficient expression of *amdS* from the *gatA* promoter to support growth. As shown in Figure 4, an extracellular concentration of 100 µM GA was sufficient to induce expression of the GA promoter reporter constructs and to support growth.

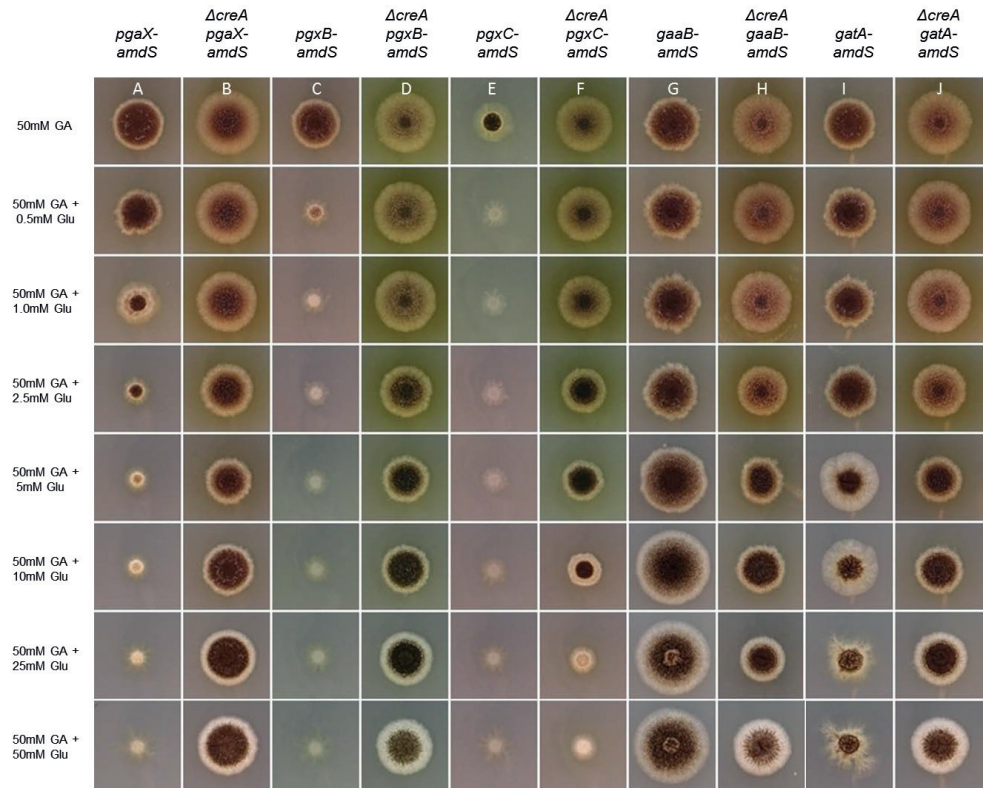


Figure 3. Galacturonic acid-induced gene expression is repressed by glucose in a *creA* dependent way. Growth of *A. niger* reporter strains was monitored on MM containing galacturonic acid/acetamide and increasing concentrations of glucose. Columns A, C, E, G, I represent promoter constructs in the wild-type background; columns B, D, F, H, J represent promoter constructs in the *ΔcreA* null background.

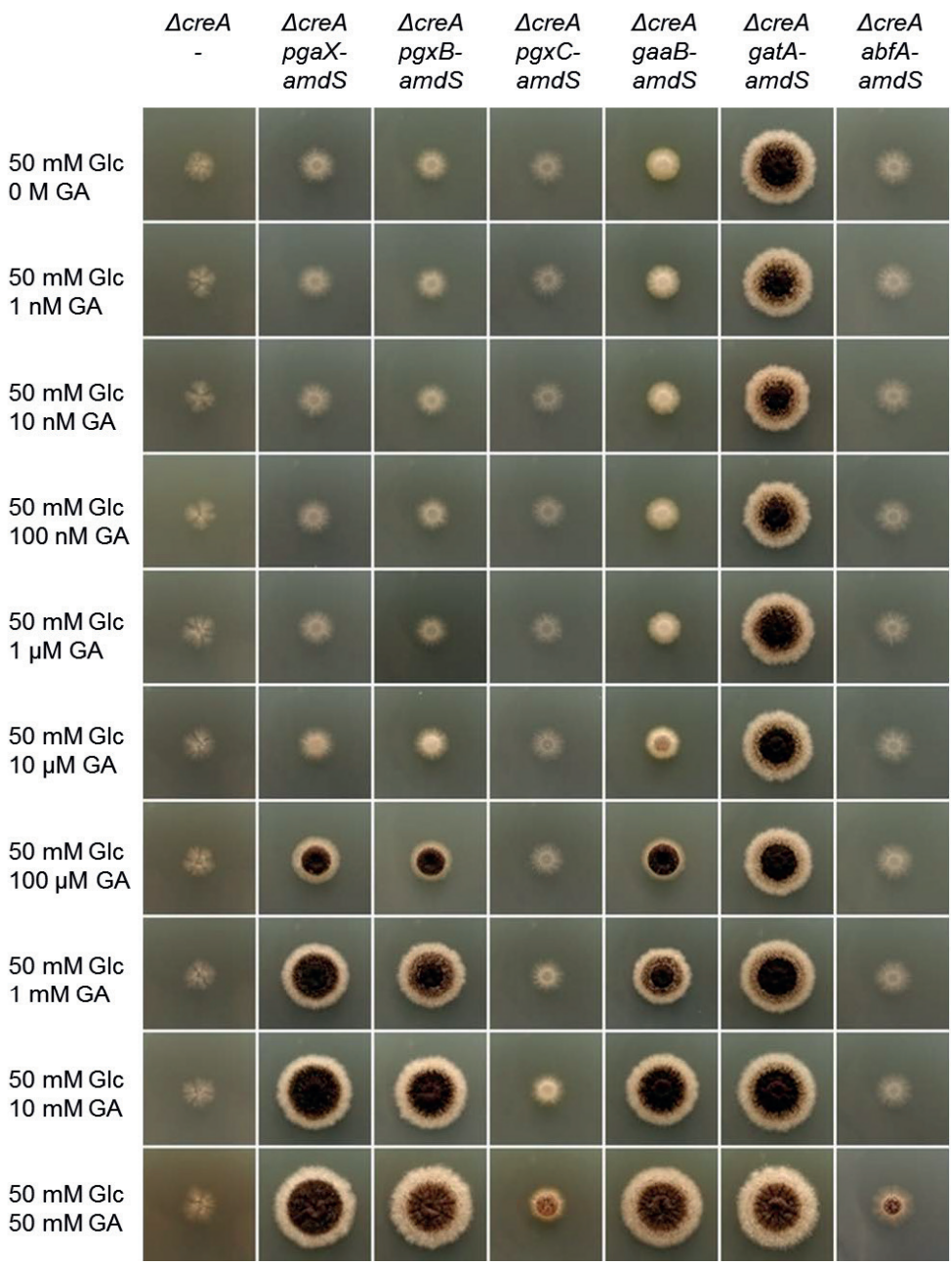


Figure 4. Galacturonic acid is required as an inducer. Growth analyses of the reporter strains on MM containing glucose/acetamide and increasing concentrations of galacturonic acid. No growth is observed for the reporter strains on MM containing glucose/acetamide without the addition of galacturonic acid, except for the *PgatA-amdS* reporter strain.

Discussion

The regulation of pectinolytic enzymes is likely to be complex and searches for transcription factor mutants that are unable to utilize galacturonic acid by targeted deletion of selected transcription factors have so far been unsuccessful (our unpublished data). The genome of *A. niger* harbours around 660 genes encoding transcription factors (Pel *et al.* 2007). We recently constructed a *A. niger* transcription factor knock-out mutants in which 240 transcription factor encoding gene were deleted. None of the transcription factor mutants in our collection showed a specific growth defect on GA, PGA or pectin (unpublished results). Besides the possibility that we missed so far a GA specific transcription factor, a possible explanation for why it is so difficult to obtain such mutants is genetic redundancy and/or overlapping GA degradation and utilization networks. In both cases, deletion of a single transcription factor does not result in strongly reduced growth on GA, PGA, or pectin. The co-regulation of a particular set of pectinolytic genes in *A. niger* supports however the idea that a specific GA-responsive transcriptional activator is involved (De Vries *et al.* 2002; Martens-Uzunova and Schaap 2008).

To study the regulation of pectinolytic genes, several promoters of GA-induced genes were selected based on their expression profile in microarray studies (Martens-Uzunova and Schaap 2008). By fusing these promoter regions to either the *uidA* or *amdS* marker gene, we could show that the induction of most of these genes was specific for GA and that the promoter was not induced on other monomeric carbon sources including those that are found in pectin, such as xylose, rhamnose and arabinose. Thus, it is likely that these promoters are not under control of e.g. the XlnR, RhaR and AraR transcription factors. The reduced growth of the *pgxC-amdS* reporter on GA/acetamide plates suggests a lower expression of *pgxC* compared to *pgaX* and *pgxB* (Figure 2). As the expression in liquid cultures of *pgaX*, *pgxB* and *pgxC* are comparable (Martens-Uzunova and Schaap 2008) it is well possible that the expression of *pgxC* is also dependant on the mode of cultivation (plate vs submerged) as has been shown for other genes encoding extracellular enzymes (Te Biesebeke *et al.* 2005) or sugar transporters (Fekete *et al.* 2012).

An important finding for future research is that the *pgaX*, *pgxB* and *pgxC* genes require GA as an inducer for expression (Figure 4). Even under derepressing conditions (in the *creA* mutant), the expression of *amdS* was not sufficient to allow growth. The *pgaX, pgxB* or *pgxC* reporter strains in the *creA* background can now be used in a genetic screen to isolate mutants that no longer require the presence of an inducer. Such mutants might contain a mutation that

causes constitutive activation of a GA-specific transcription factor. Identification of transcription factors involved in the activation of GA-responsive genes is important to optimize production of pectin-degrading enzymes in filamentous fungi.

Previous studies related to GA-induced gene expression identified two conserved promoter elements that could be involved in induction. The pecR site (CCCTGA) was identified by Benen and co-workers and it has been suggested that this element might bind a pectin-specific regulator (Benen *et al.* 1996; Coutinho *et al.* 2009). This element is present in the *pgaX* promoter at position -695 to -690. Our promoter deletion analysis showed, however, that this element is not required for the induction by GA. We also noticed the 50% reduction in GUS activity between transformants with the largest transcript (1203) and the transformants with the shorter transcripts (1005, 688 and 409 bp (Fig. 1). The reason for this difference and whether it has biological significance is currently not known.

Both the deletion study and the site-directed mutagenesis of the second conserved element (5'-TCCNCCAAT-3') indicate that this promoter element is important for induction and therefore we named this element GARE for galacturonic acid-responsive element. As previously noted (Bussink *et al.* 1992b; Martens-Uzunova and Schaap 2008), this motif includes a putative binding site (CCAAT) for the CCAAT-binding factor (CBF) complex, also known as AnCF or the HAP-complex(Kato *et al.* 1998; Steidl *et al.* 1999). The HAP complex is an important factor in activating gene expression as the complex is involved in recruiting specific transcription factors to their target sites. In *A. nidulans* the AmdR transcription factor is a $Zn(II)_2Cys_6$ binuclear cluster DNA- binding protein that is required for omega-amino acid induction of the *amdS* gene and the genes for omega-amino acid utilization (Andrianopoulos and Hynes 1988). Deletion of the *hapB* or *hapE* genes results in loss of omega-amino acid induction of *amdS* expression. The most likely explanation for the interaction between AmdR and the HAP-complex is that the HAP-complex facilitates AmdR binding to DNA *in vivo*. It has been suggested that HAP binding to its target sequence is a prerequisite for a change in chromatin structure necessary for AmdR binding (Steidl *et al.* 1999). A similar mode of action can be attributed to the regulation of *agdA* and *taaG2* in *A. nidulans* by AmyR. The gene *amyR* encodes a $Zn(II)_2Cys_6$ transcription factor specific for the induction of starch-degrading enzymes. Unlike AmdR, the AnCF-binding site and AmyR-binding site (starch-responsive element; SRE) are spaced by 100 nucleotides (Tani *et al.* 2001a; Tani *et al.* 2001b), suggesting that no direct interaction between AnCF and AmyR is necessary for activation. The binding of AmyR to the SRE is necessary for induction of gene

expression of both genes, whereas AnCF binding seems to be responsible for high affinity binding of AmyR. Our observation that a mutation in the CCAAT motif completely abolishes GA-induced gene expression, favours a mechanism similar to the AmdR example. In the case of GA-induced gene expression, the binding of a hypothetical GA-specific transcriptional activator (tentatively named GaaR) might require HAP for DNA binding. To examine the role of the HAP complex in *A. niger*, we disrupted the *hapB* gene. However, the deletion mutant has a very severe growth phenotype on a variety of different carbon and nitrogen sources tested (data not shown). Because of the very poor growth of the *hapB* mutant, no expression studies could be performed to analyse whether *pgaX* (or other pectinolytic genes) was induced in the *hapB* mutant. A additional example in which HAP binds to another transcription factor that binds in close proximity of the HAP complex, has recently been discovered for the HAP2/3/5-HapX complex. Both the HAP complex and the HapX protein interact with each other as well as with their DNA binding sites and both the protein-protein and DNA-protein interactions are required to regulate gene expression in order to maintain iron homeostasis (Hortschansky *et al.* 2015). HapX requires the minimal motif 5'-GAT-3', which is localized in a distance of 11 to 12 base pairs downstream of the respective CCAAT box (Hortschansky *et al.* 2015). To examine the possibility that additional conserved elements surrounding GARE were previously missed (Martens-Uzunova and Schaap 2008), we aligned promoter sequences of GA-induced genes and looked for conserved regions in regions including 45 bp up- and 45 bp downstream of the GAREs. The results are shown in Supplemental Figure 1A. As shown in this alignment, no additional motif was identified. We noted however, the preference of a TT motif 2 nucleotides upstream of the 5'-TCCnCCAAT-3 region. However, this motif was not entirely conserved, but could be relevant. To analyse the possible relevance of this TT motif, a cross *Aspergillus* analysis was performed. Since orthologous genes for *gaaA*, *gaaB* and *gaaD* are readily identified in *Aspergillus* species analysed (*A. nidulans*, *A. fumigatus* and *A. oryzae*), we performed our analyses on the promoter for these genes. Please note that *gaaA* and *gaaC* are expressed form the same promoter and therefore *gaaC* was not included separately. The analysis is shown in Supplemental Figure 1B and shows that the TT motif is conserved in the *gaaB* promoters among the four Aspergilli, but not conserved in the others genes (*gaaA* and *gaaD*). It should be noted that the GARE alignment as shown by Martens-Uzunova and Schaap 2008 (Martens-Uzunova and Schaap 2008), includes a putative GARE site in which the CCAAT motif in not completely conserved. For example, the GARE in the promoter of *gatA* 5'-TCCNGCAAT-3' or *pelA* 5'-TCCNCCTAT-3' do not match perfectly with the 5'-TCCNCCAAT-3' consensus. This suggests that small variants in CCAAT motif

are tolerable. Although the *abfA* promoter does not contain a perfect 5'-TCCNCCAAT-3' match, it does contain three variants (- 214 5'-TCCTCCACT, -470 5'-CCCTCCATT, -5'-899 TCCTCCGAT). Given the observation that the *amdS* gene us not expressed from the *abfA* promoter on GA (Fig. 2) we assume these variations are not allowed to mediate GA-induced gene expression. Further mutagenesis studies are required to clarify whether the TT motif, variations in the CCAAT motif, or other sequences surrounding the 5'-TCCNCCAAT-3' are required to function as GA-responsible elelements.

To address the question whether the 5'-TCCNCCAAT-3' motif was sufficient to induce GA-specific gene expression, the 5'-TCCACCAAT-3' motif, including four additional base pairs flanking the motif, was cloned upstream of the *A. nidulans gpdA* minipromoter. This system has been successfully used to identify a benzoate-responsive element from the *cprA* promoter (Van Den Brink *et al.* 2000). Two constructs were made consisting of either the empty mini-promoter (pMini-GUS) or the mini-promoter containing the motif 5'-TCCACCAAT-3' and transformed to the *pyrG* locus using the *pyrG** method. Transformants with targeted integration of the reporter constructs were verified by Southern blot analysis (data not shown). Expression analysis of the transformants revealed no specific induction of the GUS reporter in the presence of GA, indicating that the 5'-TCCACCAAT-3' domain is not sufficient to induce GA-specific gene expression.

Recently, the transcriptomes of A. niger grown on lignocellulose and during carbon starvation have been studied in detail (Delmas et al. 2012; Van Munster et al. 2014). This showed that a subset of genes involved in the degradation of lignocellulose is not only induced in the presence of an inducer (in this case xylose), but also by carbon starvation. The response to carbon starvation was shown to be dependent on CreA derepression, and independent of the xylanolytic regulator XlnR. The authors proposed a model in which carbon starvation induced the expression of scouting enzymes, independently of carbon source-specific inducers (Van Munster *et al.* 2014). These scouting enzymes are expressed with the purpose to sense the presence of plant cell walls by releasing small amounts of inducing sugars, to which the fungus can then respond by secreting additional hydrolases. Our *in vivo* reporter strains could well be used to verify the predictions based on the transcriptomic data. From our analysis using the promoter reporters, it is interesting to note that the GA-transporter gene is expressed under derepressing conditions even in the absence of GA as the inducer (Figure 4). The other promoters (*pgaX*, *pgxB*, *pgxC*) are not highly enough expressed

under derepressing conditions to sustain growth and this suggests that the corresponding enzymes will most likely not act as scouting enzymes.

As a working model (Figure 5), we postulate the existence of a GA-specific transcription factor (GaaR) that binds to the promoter of GA-induced genes. This transcription factor is most likely post-translationally activated by GA or a GA-derivative. Extensive co-expression network analysis based on a large collection of microarray data did not identify co-regulated transcription factors that upon disruption affected growth on GA or pectin (Homan, Alazi, de Vries and Ram, unpublished data), indicating that the activity of the proposed transcription factor is post-translationally regulated. The HAP complex is most likely involved DNA binding of the hypothetical GaaR transcription factor. Whether this directly involves the GARE box (5'-TCCNCCAAT-3') or whether GaaR requires HAP binding to bind to another motif(s) is currently unknown. Since the CC motif is essential for GA-induced gene expression, we considered it unlikely that induction by GA depends only on the HAP complex. We can not exclude the possibility that HAP interacts with some other non-DNA binding protein that gives specificity towards HAP to mediate GA-induced gene expression.

We also provided further evidence for the role of CreA to mediate glucose repression of pgaX, pgxB and to some extent of pgxC. As pgxC is still partially repressed by glucose even in de *creA* mutant, an alternative repression mechanism might play a role. Although we could not link the TupA general repressor protein to glucose repression in *A. niger* (Schachtschabel *et al.* 2013), we cannot exclude a possible involvement of TupA in *pgxC* repression and this possibility will be addressed in future experiments. Finally, the promoter reporter strains in the *ΔcreA* background that still require an inducer for growth (*pgaX*, *pgxB* and *pgxC)* allow forward genetic screens for inducer-independent mutants. One possible explanation for such a mutant is a mutant in which a galacturonic acid sensor protein or a galacturonic acid specific transcription factor is constitutively active, and experiments are ongoing to isolate and characterize such mutants.

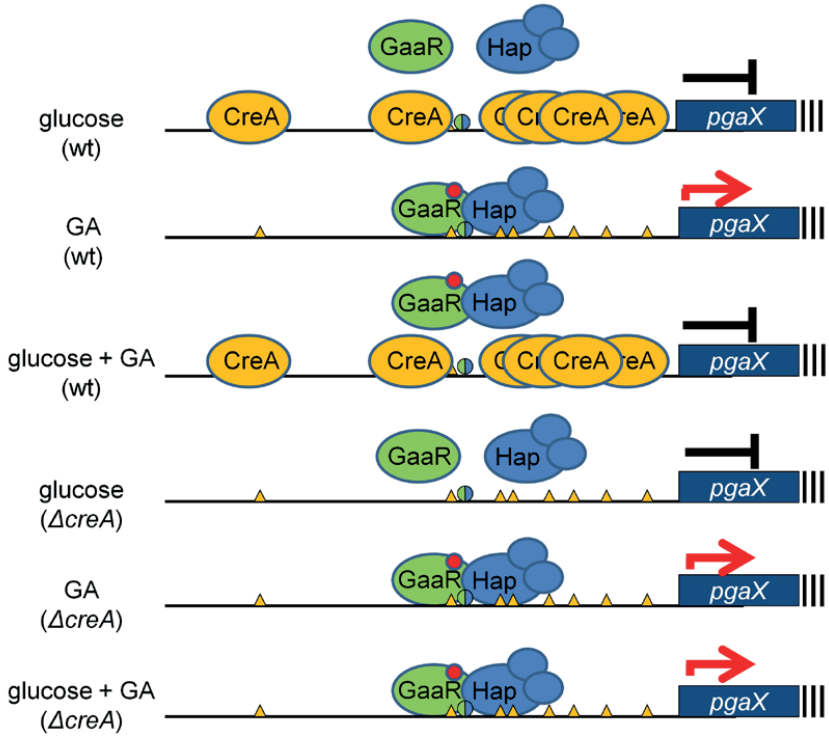


Figure 5. Speculative model for the regulation of galacturonic-acid induced gene expression in *A. niger*. GA-induced genes, poly-galacturonase X (*pgaX*) as an example, are regulated via competing inducing and repressing mechanisms. Induction of GA-responsive genes requires the presence of an inducer molecule (GA, or a metabolic derivative thereof, red circle) which is required for the activation of the putative GA-specific transcription factor (GaaR). The presence and importance of the HAP binding site as part of the GA responsive element (GARE, green/blue circle) suggests that HAP is also required for the expression, possibly by interacting with GaaR. Induction of *pgaX* is repressed by the presence of glucose and possibly other repressing sugars via CreA and putative CreA binding sites in the promoter region (brown triangles). Induction or repression of the *pgaX* gene under the conditions indicated in either the wild type (WT) or *ΔcreA* mutant are indicated by the red arrow or black symbol respectively.

Acknowledgements

We acknowledge Jiachen Ciu, Sakia duPre, Mirjam Hillman, Christan Rohe, Bas Schenkeveld and Ebru Alazi for their experimental contributions and discussions. We thank Frans Klis for critically reading and his valuable comments on the manuscript. Jing Niu is supported by a grant from the China Scholarship Council. Part of this research is supported by the Dutch Technology Foundation (STW), which is part of the Netherlands Organisation for Scientific Research (NWO).

References

Andersen, M. R., M. P. Salazar, P. J. Schaap, P. J. van de Vondervoort, D. Culley *et al.*, 2011 Comparative genomics of citric-acid-producing *Aspergillus niger* ATCC 1015 versus enzyme-producing CBS 513.88. Genome Res 21**:** 885-897.

Andrianopoulos, A., and M. J. Hynes, 1988 Cloning and analysis of the positively acting regulatory gene amdR from *Aspergillus nidulans*. Mol Cell Biol 8**:** 3532-3541.

Arentshorst, M., E. L. Lagendijk and A. F. Ram, 2015 A new vector for efficient gene targeting to the pyrG locus in *Aspergillus niger*. Fungal Biology and Biotechnology 2:2.

Arentshorst, M., A. F. Ram and V. Meyer, 2012 Using non-homologous end-joining-deficient strains for functional gene analyses in filamentous fungi. Methods Mol Biol 835**:** 133-150.

Benen, J. A., L. Pařenicová, M. Kusters-van Someren, Kester H. and Visser J., 1996 Molecular genetic and biochemical aspects of pectin degradation in *Aspergillus* in *Pectins and Pectinases*, edited by J. Visser and A. G. J. Voragen. Elsevier Science.

Bennett, J. W., and L. L.L., 1991 Growth media, pp. 441-458 in *More gene manipulations in fungi*, edited by Bennett J.W. Academic Press, San Diego.

Benz, J. P., R. J. Protzko, J. M. Andrich, S. Bauer, J. E. Dueber *et al.*, 2014 Identification and characterization of a galacturonic acid transporter from Neurospora crassa and its application for *Saccharomyces cerevisiae* fermentation processes. Biotechnol Biofuels 7**:** 20.

Bussink, H. J., F. P. Buxton, B. A. Fraaye, L. H. de Graaff and J. Visser, 1992a The polygalacturonases of *Aspergillus niger* are encoded by a family of diverged genes. Eur J Biochem 208**:** 83-90.

Bussink, H. J., F. P. Buxton and J. Visser, 1991 Expression and sequence comparison of the *Aspergillus niger* and Aspergillus tubigensis genes encoding polygalacturonase II. Curr Genet 19**:** 467-474.

Bussink, H. J., H. C. Kester and J. Visser, 1990 Molecular cloning, nucleotide sequence and expression of the gene encoding prepro-polygalacturonaseII of *Aspergillus niger*. FEBS Lett 273**:** 127-130.

Bussink, H. J., J. P. T. W. van den Hombergh, P. R. L. A. Van den Ijssel and V. J, 1992b Characterization of polygalacturonase-overproducing *Aspergillus niger* transformants. Applied Microbiology and Biotechnology 37.

Coutinho, P. M., M. R. Andersen, K. Kolenova, P. A. vanKuyk, I. Benoit *et al.*, 2009 Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. Fungal Genet Biol 46 Suppl 1**:** S161-S169.

Damveld, R. A., M. Arentshorst, A. Franken, P. A. vanKuyk, F. M. Klis *et al.*, 2005 The *Aspergillus niger* MADS-box transcription factor RlmA is required for cell wall reinforcement in response to cell wall stress. Mol Microbiol 58**:** 305-319.

Damveld, R. A., A. Franken, M. Arentshorst, P. J. Punt, F. M. Klis *et al.*, 2008 A novel screening method for cell wall mutants in *Aspergillus niger* identifies UDP-galactopyranose mutase as an important protein in fungal cell wall biosynthesis. Genetics 178**:** 873-881.

de Groot, M. J., P. J. van de Vondervoort, R. P. de Vries, P. A. vanKuyk, G. J. Ruijter *et al.*, 2003 Isolation and characterization of two specific regulatory *Aspergillus niger* mutants shows antagonistic regulation of arabinan and xylan metabolism. Microbiology 149**:** 1183-1191.

de Vries, R. P., J. Jansen, G. Aguilar, L. Parenicova, V. Joosten *et al.*, 2002 Expression profiling of pectinolytic genes from *Aspergillus niger*. FEBS Lett 530**:** 41-47.

Delmas, S., S. T. Pullan, S. Gaddipati, M. Kokolski, S. Malla *et al.*, 2012 Uncovering the genome-wide transcriptional responses of the filamentous fungus *Aspergillus niger* to lignocellulose using RNA sequencing. PLoS Genet 8**:** e1002875.

Dowzer, C. E., and J. M. Kelly, 1991 Analysis of the creA gene, a regulator of carbon catabolite repression in *Aspergillus nidulans*. Mol Cell Biol 11**:** 5701-5709.

Fekete, E., L. Karaffa, B. Seiboth, E. Fekete, C. P. Kubicek *et al.*, 2012 Identification of a permease gene involved in lactose utilisation in *Aspergillus nidulans*. Fungal Genet Biol 49**:** 415-425.

Hilditch, S., S. Berghall, N. Kalkkinen, M. Penttila and P. Richard, 2007 The missing link in the fungal D-galacturonate pathway: identification of the L-threo-3-deoxy-hexulosonate aldolase. J Biol Chem 282**:** 26195-26201.

Hortschansky, P., E. Ando, K. Tuppatsch, H. Arikawa, T. Kobayashi *et al.*, 2015 Deciphering the combinatorial DNA-binding code of the CCAAT-binding complex and the iron-regulatory basic region leucine zipper (bZIP) transcription factor HapX. J Biol Chem 290**:** 6058-6070.

Inoue, H., H. Nojima and H. Okayama, 1990 High efficiency transformation of *Escherichia coli* with plasmids. Gene 96**:** 23-28.

Kato, M., A. Aoyama, F. Naruse, T. Kobayashi and N. Tsukagoshi, 1997 An *Aspergillus nidulans* nuclear protein, AnCP, involved in enhancement of Taka-amylase A gene expression, binds to the CCAAT-containing taaG2, amdS, and gatA promoters. Mol Gen Genet 254**:** 119-126.

Kato, M., A. Aoyama, F. Naruse, Y. Tateyama, K. Hayashi *et al.*, 1998 The *Aspergillus nidulans* CCAAT-binding factor AnCP/AnCF is a heteromeric protein analogous to the HAP complex of *Saccharomyces cerevisiae*. Mol Gen Genet 257**:** 404-411.

Kelly, J. M., and M. J. Hynes, 1985 Transformation of *Aspergillus niger* by the amdS gene of *Aspergillus nidulans*. EMBO J 4**:** 475-479.

Kowalczyk, J. E., I. Benoit and R. P. de Vries, 2014 Regulation of plant biomass utilization in *Aspergillus*. Adv Appl Microbiol 88**:** 31-56.

Kuivanen, J., D. Mojzita, Y. Wang, S. Hilditch, M. Penttila *et al.*, 2012 Engineering filamentous fungi for conversion of D-galacturonic acid to L-galactonic acid. Appl Environ Microbiol 78**:** 8676-8683.

Kuorelahti, S., P. Jouhten, H. Maaheimo, M. Penttila and P. Richard, 2006 L-galactonate dehydratase is part of the fungal path for D-galacturonic acid catabolism. Mol Microbiol 61**:** 1060-1068.

Kuorelahti, S., N. Kalkkinen, M. Penttila, J. Londesborough and P. Richard, 2005 Identification in the mold *Hypocrea jecorina* of the first fungal D-galacturonic acid reductase. Biochemistry 44**:** 11234-11240.

Liepins, J., S. Kuorelahti, M. Penttila and P. Richard, 2006 Enzymes for the NADPH-dependent reduction of dihydroxyacetone and D-glyceraldehyde and L-glyceraldehyde in the mould Hypocrea jecorina. FEBS J 273**:** 4229-4235.

Litzka, O., K. Then Bergh and A. A. Brakhage, 1996 The *Aspergillus nidulans* penicillin-biosynthesis gene aat (penDE) is controlled by a CCAAT-containing DNA element. Eur J Biochem 238**:** 675-682.

Maldonado, M. C., and A. M. Strasser de Saad, 1998 Production of pectinesterase and polygalacturonase by *Aspergillus niger* in submerged and solid state systems. J Ind Microbiol Biotechnol 20**:** 34-38.

Martens-Uzunova, E. S., and P. J. Schaap, 2008 An evolutionary conserved d-galacturonic acid metabolic pathway operates across filamentous fungi capable of pectin degradation. Fungal Genet Biol 45**:** 1449-1457.

Martens-Uzunova, E. S., and P. J. Schaap, 2009 Assessment of the pectin degrading enzyme network of *Aspergillus niger* by functional genomics. Fungal Genet Biol 46 Suppl 1**:** S170-S179.

Martens-Uzunova, E. S., J. S. Zandleven, J. A. Benen, H. Awad, H. J. Kools *et al.*, 2006 A new group of exo-acting family 28 glycoside hydrolases of *Aspergillus niger* that are involved in pectin degradation. Biochem J 400**:** 43-52.

Meyer, V., A. F. J. Ram and P. J. Punt, 2010 Genetics, genetic manipulation and approaches to strain improvement of filamentous fungi, pp. 318-329 in *Manual of industrial microbiology and biotechnology*, edited by R. Baltz, A. Demain and Eds. Wiley, NY.

Mohnen, D., 2008 Pectin structure and biosynthesis. Curr Opin Plant Biol 11**:** 266-277.

Mojzita, D., M. Wiebe, S. Hilditch, H. Boer, M. Penttila *et al.*, 2010 Metabolic engineering of fungal strains for conversion of D-galacturonate to meso-galactarate. Appl Environ Microbiol 76**:** 169-175.

Parenicova, L., J. A. Benen, H. C. Kester and J. Visser, 1998 pgaE encodes a fourth member of the endopolygalacturonase gene family from *Aspergillus niger*. Eur J Biochem 251**:** 72-80.

Parenicova, L., H. C. Kester, J. A. Benen and J. Visser, 2000 Characterization of a novel endopolygalacturonase from *Aspergillus niger* with unique kinetic properties. FEBS Lett 467**:** 333-336.

Pel, H. J., J. H. de Winde, D. B. Archer, P. S. Dyer, G. Hofmann *et al.*, 2007 Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. Nat Biotechnol 25**:** 221-231.

Punt, P. J., F. H. Schuren, J. Lehmbeck, T. Christensen, C. Hjort *et al.*, 2008 Characterization of the *Aspergillus niger* prtT, a unique regulator of extracellular protease encoding genes. Fungal Genet Biol 45**:** 1591-1599.

Roberts, I. N., R. P. Oliver, P. J. Punt and C. A. van den Hondel, 1989 Expression of the Escherichia coli beta-glucuronidase gene in industrial and phytopathogenic filamentous fungi. Curr Genet 15**:** 177-180.

Ruijter, G. J., and J. Visser, 1997 Carbon repression in Aspergilli. FEMS Microbiol Lett 151**:** 103-114.

Schachtschabel, D., M. Arentshorst, B. M. Nitsche, S. Morris, K. F. Nielsen *et al.*, 2013 The transcriptional repressor TupA in *Aspergillus niger* is involved in controlling gene expression related to cell wall biosynthesis, development, and nitrogen source availability. PLoS One 8**:** e78102.

Shroff, R. A., S. M. O'Connor, M. J. Hynes, R. A. Lockington and J. M. Kelly, 1997 Null alleles of creA, the regulator of carbon catabolite repression in *Aspergillus nidulans*. Fungal Genet Biol 22**:** 28-38.

Sloothaak, J., M. Schilders, P. J. Schaap and L. H. de Graaff, 2014 Overexpression of the *Aspergillus niger* GatA transporter leads to preferential use of D-galacturonic acid over D-xylose. AMB Express 4**:** 66.

Steidl, S., P. Papagiannopoulos, O. Litzka, A. Andrianopoulos, M. A. Davis *et al.*, 1999 AnCF, the CCAAT binding complex of *Aspergillus nidulans*, contains products of the hapB, hapC, and hapE genes and is required for activation by the pathway-specific regulatory gene amdR. Mol Cell Biol 19**:** 99-106.

Tani, S., T. Itoh, M. Kato, T. Kobayashi and N. Tsukagoshi, 2001a In vivo and in vitro analyses of the AmyR binding site of the *Aspergillus nidulans* agdA promoter; requirement of the CGG direct repeat for induction and high affinity binding of AmyR. Biosci Biotechnol Biochem 65**:** 1568-1574.

Tani, S., Y. Katsuyama, T. Hayashi, H. Suzuki, M. Kato *et al.*, 2001b Characterization of the amyR gene encoding a transcriptional activator for the amylase genes in *Aspergillus nidulans*. Curr Genet 39**:** 10-15.

te Biesebeke, R., N. van Biezen, W. M. de Vos, C. A. van den Hondel and P. J. Punt, 2005 Different control mechanisms regulate glucoamylase and protease gene transcription in *Aspergillus oryzae* in solid-state and submerged fermentation. Appl Microbiol Biotechnol 67**:** 75-82.

Thon, M., Q. Al Abdallah, P. Hortschansky, D. H. Scharf, M. Eisendle *et al.*, 2010 The CCAAT-binding complex coordinates the oxidative stress response in eukaryotes. Nucleic Acids Res 38**:** 1098-1113.

van den Brink, J. M., P. J. Punt, R. F. van Gorcom and C. A. van den Hondel, 2000 Regulation of expression of the *Aspergillus niger* benzoate para-hydroxylase cytochrome P450 system. Mol Gen Genet 263**:** 601-609.

van Gorcom, R. F., and C. A. van den Hondel, 1988 Expression analysis vectors for *Aspergillus niger*. Nucleic Acids Res 16**:** 9052.

van Heeswijck, R., and M. J. Hynes, 1991 The amdR product and a CCAAT-binding factor bind to adjacent, possibly overlapping DNA sequences in the promoter region of the *Aspergillus nidulans amdS* gene. Nucleic Acids Res 19**:** 2655-2660.

van Munster, J. M., P. Daly, S. Delmas, S. T. Pullan, M. J. Blythe *et al.*, 2014 The role of carbon starvation in the induction of enzymes that degrade plant-derived carbohydrates in *Aspergillus niger*. Fungal Genet Biol 72**:** 34-47.

Westholm, J. O., N. Nordberg, E. Muren, A. Ameur, J. Komorowski *et al.*, 2008 Combinatorial control of gene expression by the three yeast repressors Mig1, Mig2 and Mig3. BMC Genomics 9**:** 601.

Wiebe, M. G., D. Mojzita, S. Hilditch, L. Ruohonen and M. Penttila, 2010 Bioconversion of D-galacturonate to keto-deoxy-L-galactonate (3-deoxy-L-threo-hex-2-ulosonate) using filamentous fungi. BMC Biotechnol 10**:** 63.

Yuan, X. L., C. Goosen, H. Kools, M. J. van der Maarel, C. A. van den Hondel *et al.*, 2006 Database mining and transcriptional analysis of genes encoding inulin-modifying enzymes of *Aspergillus niger*. Microbiology 152**:** 3061-3073.

Zhang, L., C. Hua, J. H. Stassen, S. Chatterjee, M. Cornelissen *et al.*, 2014 Genome-wide analysis of pectate-induced gene expression in *Botrytis cinerea*: identification and functional analysis of putative d-galacturonate transporters. Fungal Genet Biol 72**:** 182-191.

Zhang, L., H. Thiewes and J. A. van Kan, 2011 The D-galacturonic acid catabolic pathway in *Botrytis cinerea*. Fungal Genet Biol 48**:** 990-997.

Supporting information

The supplementary material of this chapter are available via

http://www.sciencedirect.com/science/article/pii/S1087184515001516?via%3Dihub and comprises the following:

Supplemental Table 1. Plasmids used in this study

[1] *pyrG** refers to targeting integration method of DNA constructs using the pyrG targeting method described by van Gorcom et al., 1988, [2] *pyrG*** refers to targeting integration method to the pyrG locus described by Arentshorst et al., 2015.

Supplemental Table 2. Primers used in this study

*Restriction sites are underlined; **Letters in bold indicate mutations in the GARE element ((Mut1-3), ***bold letter indicate overlapping sequences for fusion PCR.

Supplemental Table 3. Positions of putative CreA bindingsites (SYGGRG) in the GA-induced promoters and *abfA*. Number of putative CreA binding sites in given in brackets. Distance from the ATG start codon in bp) is given.

Supplemental Figure 1. A) Promoter alignment of sequences surrounding the GARE motif of GA-induced genes in A. niger. GA-motifs were taken from Martens-Uzunova et al., 2008. B) Promoter alignments of gaaA, gaaB and gaaD orthologs in A. niger, A. nidulans, A. fumigatus and A. *oryzae*. Orthologs were taken from AspGD. (+) or (-) indicate whether the GARE motif was in forward or reverse orientation. The GARE motif, as well as the partially conserved TT-motif 5' upstream of GARE, are highlighted in yellow. * indicates that no CCNCCAAT motive was found in these promoter regions.

Chapter 5

The transcriptional activator GaaR of *Aspergillus niger* is required for release and utilization of D-galacturonic acid from pectin

Ebru Alazi*, Jing Niu*, Joanna E. Kowalczyk*, Mao Peng, Maria Victoria Aguilar Pontes, Jan A.L. van Kan, Jaap Visser, Ronald P. de Vries, Arthur F.J. Ram

* These authors contributed equally to this work

Published in FEBS Lett. 2016 Jun; 590(12): 1804-1815

Abstract

We identified the D-galacturonic acid (GA) responsive transcriptional activator GaaR of the saprotrophic fungus *Aspergillus niger*, which was found to be essential for growth on GA and polygalacturonic acid (PGA). Growth of the *ΔgaaR* strain was reduced on complex pectins. Genome-wide expression analysis showed that GaaR is required for the expression of genes necessary to release GA from PGA and more complex pectins, to transport GA into the cell and to induce the GA-catabolic pathway. Residual growth of *ΔgaaR* on complex pectins is likely due to expression of pectinases acting on rhamnogalacturonan and subsequent metabolism of the monosaccharides other than GA.

Keywords

polygalacturonic acid, pectinase, Zn_2Cys_6 transcription factor, gene regulation, transcriptomics

Abbreviations

AP apple pectin
CM Complete medium
CP citrus pectin
GA D-galacturonic acid
MM minimal medium
PGA polygalacturonic acid
RG rhamnogalacturonan
SBP sugar beet pectin
TF transcription factor
XGA xylogalacturonan

Introduction

Pectins are complex heterogeneous polysaccharides found in plant cell walls. Four sub-structures of pectin have been identified and include polygalacturonic acid (PGA) also known as homogalacturonan, xylogalacturonan (XGA), rhamnogalacturonan I (RG-I), and rhamnogalacturonan II (RG-II) (Mohnen 2008). The backbones of PGA, XGA and RG-II are made up of α-1,4-linked D-galacturonic acid (GA) residues. PGA, a linear polymer of GA, is the most abundant polysaccharide present in pectin (Mohnen 2008). In XGA, β-D-xylose residues are β-1,3-linked to GA residues of the PGA backbone. The backbone of RG-I is made up of alternating GA and L-rhamnose residues (Mohnen 2008; Leijdekkers *et al.* 2015). Side chains of RG-II contain at least 12 different types of monosaccharides, whereas the side chains of RG-I are mainly arabinan and arabinogalactan comprising of L-arabinose and D-galactose residues (Mohnen 2008).

In nature, pectin is an important carbon source for many saprotrophic fungi, such as *Aspergillus niger*. Previous studies demonstrated that *A. niger* can produce more pectin degrading enzymes than other more specialized fungi such as *Podospora anserina* or *Neurospora crassa* (Espagne *et al.* 2008; Coutinho *et al.* 2009; Martens-Uzunova and Schaap 2009). GA is the main product of pectin degradation. In *A. niger*, GA is transported into the cell by a GA-induced sugar transporter named GatA (Sloothaak *et al.* 2014). GA is then catabolized into pyruvate and glycerol (Martens-Uzunova and Schaap 2008), through a pathway consisting of four enzymes: GaaA, D-galacturonate reductase, GaaB, L-galactonate dehydratase, GaaC, 2-keto-3-deoxy-L-galactonate aldolase, and GaaD, L-glyceraldehyde reductase (Martens-Uzunova and Schaap 2008). Deletion of *gaaA*, *gaaB* or *gaaC* abolished growth on GA as the sole carbon source (Mojzita *et al.* 2010b; Wiebe *et al.* 2010; Kuivanen *et al.* 2012). *gaaD*, also known as the L-arabinose reductase gene, *larA*, is involved in the L-arabinose catabolic pathway and the *ΔlarA* strain showed a reduced growth on L-arabinose as the sole carbon source (Mojzita *et al.* 2010a).

The production of extra- and intracellular enzymes in *A. niger* is regulated by a network of transcription factors (TFs) (Kowalczyk *et al.* 2014). Small sugar molecules (mono- and disaccharides) act as inducers and stimulate TFs which can bind to conserved motifs in the promoters of their target genes and activate or repress their expression. Expression of pectinase genes is highly controlled and depends on both induction and carbon catabolite repression (De Vries *et al.* 2002; Niu *et al.* 2015). Induction of the genes required for pectin degradation, GA transport and GA catabolism requires the presence of GA and it has been

shown that GA or a derivative of GA induces the expression of pectinase genes (De Vries *et al.* 2002; Wiebe *et al.* 2010; Kuivanen *et al.* 2012).

Coordination of the induction of genes encoding extracellular enzymes and sugar uptake systems in fungi are often mediated by Zn_2Cys_6 TFs that bind to conserved promoter elements in the co-regulated genes (Chang and Ehrlich 2013; Kowalczyk *et al.* 2014; Tani *et al.* 2014). TFs inducing the genes required for the utilization of L-rhamnose (RhaR), arabinan/L-arabinose (AraR), xylan/D-xylose (XlnR), D-galactose (GalX) and cellulose (XlnR, ClrA and ClrB) have been identified in *A. niger* (Van Peij *et al.* 1998; Battaglia *et al.* 2011; Gruben *et al.* 2012; Gruben *et al.* 2014; Raulo *et al.* 2016). Although L-rhamnose, L-arabinose, D-xylose, and D-galactose are also present in complex pectins, knock out mutants in these TFs display no signs of reduced growth on pectin (Battaglia *et al.* 2011; Gruben *et al.* 2012; Gruben *et al.* 2014), suggesting that the utilization of GA, the main component of this substrate, is not affected.

Martens-Uzunova and Schaap have previously identified a set of GA-induced genes in *A. niger*, containing several pectinases (*pgaX*, *pgxA*, *pgxB*, *pgxC*, *paeA*, *pelA* and *abfC*), sugar transporter encoding genes (*gatA*, An03g01620 and An07g00780) and the GA catabolic pathway genes (*gaaA-D)* (Martens-Uzunova and Schaap 2008). These genes were suggested as the GA-regulon and contain a common GA responsive element (GARE) in their promoter regions. The consensus element was defined as CCNCCAA (Martens-Uzunova and Schaap 2008). Deletion and mutational analysis of GARE showed that the element is required for GA-induced gene expression in both *A. niger* and *Botrytis cinerea* (Niu *et al.* 2015; Zhang *et al.* 2016). A yeast one-hybrid study using a GA-responsive promoter in *B. cinerea* recently identified a novel Zn_2Cys_6 TF (BcGaaR) required for GA utilization (Zhang *et al.* 2016). In this study, the GA-responsive transcriptional activator GaaR of *A. niger* was identified by homology to BcGaaR. Deletion analysis and transcriptomic profiling studies performed in this study showed that the *A. niger* GaaR ortholog is required for growth on GA and PGA and for the induction of the GA-regulon when grown on sugar beet pectin (SBP).

Materials and Methods

Strains, media and growth conditions

A. niger strains MA234.1 (*cspA1*, *kusA::DR-amdS-DR*) and N593.20 (*cspA1*, *pyrG⁻*, *kusA::amdS*) were used to create the *ΔgaaR* strains. N593.20 was made by transformation of N593 (Goosen *et al.* 1987) with a deletion construct (*kusA::amdS*) (Meyer *et al.* 2007) resulting in the deletion of *kusA*. Strain FP-1132.1 (*cspA1, pyrG⁻::AOpyrG, kusA::amdS)* was

obtained by transformation of N593.20 with *pyrG* from *Aspergillus oryzae.* MA234.1 was obtained by transformation of MA169.4 (*kusA*$^{-}$*, pyrG*$^{-}$) (Carvalho *et al.* 2010) with a 3.8 kb *Xba*I fragment containing the *A. niger pyrG* gene, resulting in the full restoration of the *pyrG* locus.

Complementation studies were performed with JN35.1 (*cspA1, kusA::DR-amdS-DR, gaaR::hygB*). To restore functionality of the *kusA* gene to allow ectopic integration of the complementing fragment, the *amdS* marker was looped out of JN35.1 by FAA counter-selection as described (Arentshorst *et al.* 2012) to give JN36.1. The *gaaR* complemented strain JN37.4 was created using JN36.1, by transformation of the *gaaR* gene including promoter and terminator regions (see below). All strains used are listed in Table S1.

Media were prepared as described (Arentshorst *et al.* 2012). For growth phenotype analyses, strains were grown on minimal medium (MM) with 1.5% (w/v) agar and various sole carbon sources: 25 or 50 mM glucose (VWR International), GA (Chemodex), L-rhamnose (Fluka), L-arabinose (Sigma-Aldrich) or D-xylose (Merck), and 1% (w/v) PGA (Sigma), SBP (Pectin Betapec RU301 Herbstreith & Fox KG), citrus pectin (CP) (Acros Organics) or apple pectin (AP) (Pectin Classic AU2022 Herbstreith & Fox KG). pH was adjusted to 5.8 with NaOH or HCl buffer. The plates were inoculated with 2 µl containing 1000 freshly harvested spores and cultivated at 30 °C for 4 days. For gene expression analyses, freshly harvested spores were inoculated with a final concentration of 10^6 spores/ml in 100 ml complete medium (CM) (pH 5.8) with 2% (w/v) D-fructose (Sigma-Aldrich) and were pre-grown for 16h. For Northern blot analysis, mycelium was harvested by filtration through sterile myracloth, washed twice with MM with no carbon sources (pH 4.5) and 1.5 g (wet weight) mycelium was transferred and grown in 50 ml MM (pH 4.5) with 50 mM GA or 50 mM D-fructose for 2, 4 and 6h. For RNA-seq analysis, 2.5 g of pre-grown mycelia were transferred to 50 ml MM (pH 4.5) with 25 mM GA and incubated for 2h or to 50 ml MM with 1% SBP and incubated for 2, 8 or 24h. All incubations were performed in rotary shaker at 30°C and 250 rpm.

Construction of gene deletion and complementation strains

Protoplast-mediated transformation of *A. niger*, purification of the transformants and genomic DNA extraction were performed as described (Arentshorst *et al.* 2012). To construct the deletion cassettes, 5' and 3' flanks of the *gaaR* gene were PCR-amplified using the primer pairs listed in Table S2 and N402 genomic DNA as template. To create JN35.1 strain, the split marker fragments with *hygB* selection were created using fusion PCR (Arentshorst *et al.*

2015) and transformed to MA234.1. To create FP-1126.1 strain, the flanking regions were fused with a fragment containing the *A. oryzae pyrG* gene using GoTaq® Long polymerase (Promega) and transformed into N593.20 strain. Parental strains and *gaaR* deletion mutants were deposited at the Centraal Bureau Schimmelcultures (CBS) under accession numbers indicated in Table S1. To complement the *gaaR* gene, the *gaaR* gene together with its 5' and 3' flanks was PCR-amplified using the primer pairs listed in Table S2, ligated into pJET1.2/blunt cloning vector (Fermentas), amplified in the *E. coli* strain DH5*α* and transformed in to strain JN36.1 together with plasmid pMA357. pMA357 contains the *A. nidulans amdS* gene, cloned behind the *A. nidulans gdpA* promoter (Mark Arentshorst, unpublished vector). Deletion and complementation of *gaaR* were confirmed via Southern blot analysis or diagnostic PCR.

Gene expression analysis

For Northern blot analysis, strains MA234.1 (reference strain) and JN35.1 (Δ*gaaR*) were pre-grown in CM with D-fructose. At the time of transfer (t = 0) and 2, 4 and 6h after the transfer to MM with GA or D-fructose, mycelium was harvested from cultures by filtration through sterile myracloth and frozen immediately in liquid nitrogen. Mycelium samples were stored at -80 °C. Total RNA was extracted from frozen mycelium samples after grinding in liquid nitrogen, using NucleoSpin RNA Kit (Macherey-Nagel) following the protocol provided by the supplier, including the rDNase treatment. Total RNA samples were stored at -80 °C. Quantification and purity assessment of total RNA was done by spectrophotometric method (NanoDrop 2000, Thermo Scientific). Standard molecular techniques were applied as described (Sambrook and Russell 2001). 3.5 µg RNA was loaded per sample and hybridized with [α-32P]-dCTP labelled probes after blotting (DecaLabel DNA Labelling Kit, Thermo Scientific). Probes were PCR-amplified using the N402 genomic DNA and the primer pairs are listed in Table S2. For RNA-seq analysis, the mycelium of FP-1132.1 (reference strain) and FP-1126.1 (Δ*gaaR*) was ground in Tissue Lyser II (Qiagen) and RNA was extracted using TRIzol reagent (Invitrogen) and purified with NucleoSpin RNA Clean-up kit (Macherey-Nagel) with rDNase treatment. RNA quantity of the samples was checked with a NanoDrop-1000 spectrophotometer and the quality by RNA gel electrophoresis. Single-read samples were sequenced using Illumina HiSeq^{TM} 2000 platform (http://illumina.com). Purification of mRNA, synthesis of cDNA library and sequencing reactions were conducted in the BGI Tech Solutions Co., Ltd. (Hong Kong). Transfer experiments and subsequent RNA-sequencing were performed in duplicates.

Bioinformatics

Raw reads were produced from the original image data by base calling. On average, ~13 million read of 51 bp per sample were obtained. After data filtering, the adaptor sequences, highly 'N' containing reads (> 10% of unknown bases) and low quality reads (more than 50% bases with quality value of < 5%) were removed. After data filtering, in average, ~97.5% clean reads remained in each sample. Clean reads were then mapped to the genome of *Aspergillus niger* NRRL3 (http://genome.jgi.doe.gov/Aspni_NRRL3_1) using Bowtie2 (Langmead *et al.* 2009) and BWA software (Li and Durbin 2009). In average, 63.8% total mapped reads to the genome was achieved. The gene expression level was measured in "fragments per kilobase of exon model per million mapped reads" (FPKM) (Trapnell *et al.* 2010) using RSEM tool (Li and Dewey 2011). Genes with expression value lower than 14 were considered low-expressed (approximately bottom 50%) and differential expression was identified by Student's t-test with a P-value cut-off 0.05. The RNA-seq data have been submitted to Gene Expression Omnibus (GEO) (Edgar *et al.* 2002) with accession number: GSE80227. Homology searches were performed using the blastp algorithm from NCBI against the non-redundant database and proteins with an E-value ≤ 1E-50 were defined as homologous (Altschul *et al.* 1990). Hierarchical clusters using the average expression values of genes were made via Genesis 1.7.7 (Sturn *et al.* 2002) with Pearson correlation and complete linkage. Low-expressed pectinases in all conditions were not included.

Results and Discussion

Identification of the *A. niger* GaaR by homology to *B. cinerea* BcGaaR

A putative *A. niger* GA-responsive transcriptional activator was identified by homology to the recently identified *B. cinerea* Zn_2Cys_6 TF (BcGaaR) (Zhang *et al.* 2016). The *A. niger* ortholog (named GaaR) is a 740 amino acid long protein encoded by *gaaR* (An04g00780/NRRL3_08195) and the bidirectional best blast hit of the 817 amino acid long BcGaaR (Bcin09g00170). Analysis of the presence of GaaR among 20 *Aspergillus* species using the *Aspergillus* genome database (http://www.aspgd.org/) revealed that all Aspergilli, except *Aspergillus glaucus* contain a GaaR ortholog in their genome (data not shown). Interestingly, *A. glaucus* is not able to grow on GA as the sole carbon source (http://www.fung-growth.org), indicating the requirement of GaaR for GA utilization. *A. niger* GaaR and BcGaaR show 50.3% identity on the amino acid level throughout the entire protein sequence (Figure S1). GaaR contains a typical Zn_2Cys_6 DNA binding domain with the pattern of $CX_2CX_6CX_6CX_2CX_6C$ close to its NH_2-terminal end (residues 26-56) and a

fungal specific TF domain (residues 139-518). Amino acid alignment and phylogenetic analysis of GaaR revealed no significant similarity (an E-value cutoff < 1E-50) of GaaR to other TFs involved in plant cell wall utilization such as XlnR, AraR, RhaR, GalX, ClrA and ClrB or to any other TF in *A. niger* (data not shown).

Deletion and complementation of *gaaR* and growth analysis of the Δ*gaaR* in *A. niger*

To assess the function of *gaaR* in *A. niger*, several deletion strains (*ΔgaaR*) were created and verified by Southern blot analysis (Figure S2 and data not shown). The growth phenotype of the *ΔgaaR* strains was analyzed on different monomeric and polymeric carbon sources (Figure 1A). Deletion of *gaaR* in the AB4.1 background (MA234.1, Figure S2) and N593 background (N593.20, Figure 1A) resulted in an identical phenotype. Disruption of *gaaR* resulted in a strongly reduced growth on GA and PGA and in a reduced growth and sporulation on SBP, CP and AP. No significant differences in growth and sporulation were observed on other carbon sources tested (Figure 1A, Figure S2). The strongly reduced growth of *ΔgaaR* on GA and PGA was fully complemented by reintroducing the *gaaR* gene ectopically (Figure S2).

GaaR is required for the induction of genes related to D-galacturonic acid utilization

The presence of GA has been shown to induce genes involved in PGA degradation (e.g. *pgxB, pgxC*), GA transport (*gatA*) and catabolism (*gaaA-D*) (Martens-Uzunova and Schaap 2008; Niu *et al.* 2015). As a first indication for the involvement of GaaR in the induction of a subset of these genes on GA, a Northern blot analysis was performed. The reference strain and *ΔgaaR* made the AB4.1 background were pre-grown in D-fructose medium and transferred to either GA or D-fructose medium. For the reference strain, transfer of mycelium to GA resulted in a rapid induction of *pgxB*, *pgxC*, *gatA*, *gaaB* and *gaaC*, whereas this induction was not observed in *ΔgaaR* (Figure 1B).

To analyze the expression of a larger number of genes involved in pectin degradation, GA transport and catabolism, a genome-wide gene expression analysis was performed using RNA-seq. The reference strain and *ΔgaaR* in the N593 background were again pre-grown in D-fructose medium and transferred to GA medium. RNA-seq analysis indicated that the GA-induced expression of all genes that were previously identified as part of the GA-regulon (Martens-Uzunova and Schaap 2008) is dependent on GaaR (Table 1 and Figure 2). The only exception is a putative GA transporter (An03g01620) that is expressed more than 3-fold less in *ΔgaaR* for which the p-value did not pass our significance level (0.05). In general, these

observations show that the genes in the suggested GA-regulon (Martens-Uzunova and Schaap 2008) showed a significant reduction in *ΔgaaR* compared to the reference strain on GA (Table 1) and that GaaR is required for the induction of those genes.

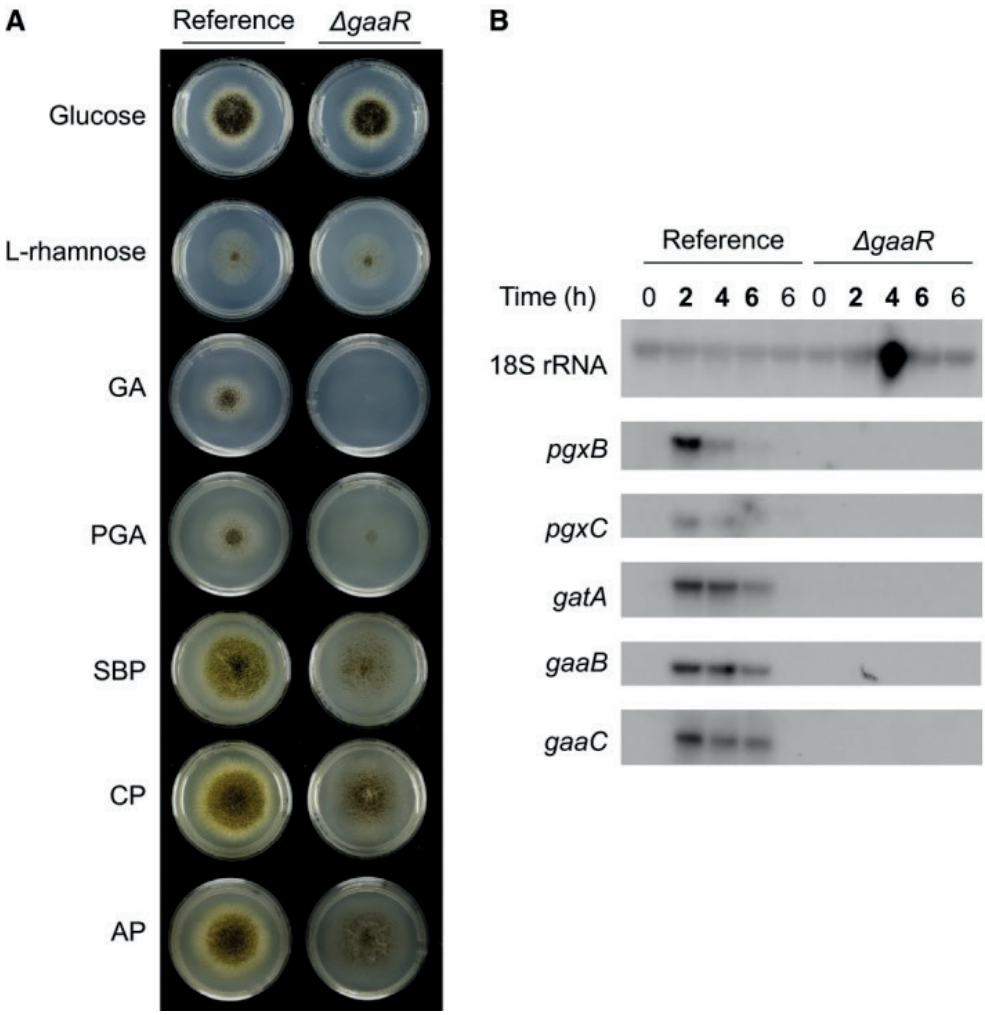


Figure 1 Phenotypic and gene expression analyses of *A. niger ΔgaaR* **A)** Growth profile of the reference strain (FP-1132.1) and *ΔgaaR* (FP-1126.1) on MM with 25mM monomeric and 1% polymeric carbon sources. Strains were grown for 4 days at 30°C. **B)** Northern blot analysis of selected GA-induced genes in the reference strain (MA234.1) and *ΔgaaR* (JN35.1). Mycelia were transferred from D-fructose (pre-culture) to GA or D-fructose. Total RNA was isolated at the time of transfer (0h) from mycelia grown in CM with 2% D-fructose and at different time points (2, 4 and 6h) after the transfer from mycelia grown in MM containing 50 mM GA (in bold) or D-fructose.

To identify additional pectinase genes controlled directly or indirectly by GaaR, the expression of all 58 pectinolytic genes (Martens-Uzunova and Schaap 2009) was examined (Table S3). An overview of the gene abbreviations and their (putative) function is given in Martens-Uzunova and Schaap, 2009 (Martens-Uzunova and Schaap 2009). This analysis resulted in the identification of several additional pectinase genes for which the expression on GA is dependent of GaaR (Table 1 and Figure 2, Figure S3). This difference could be caused by higher sensitivity of the RNA-seq analysis compared to the previously used Affymetrix microarrays. In general, these newly identified genes were lower expressed compared to the genes in the GA-regulon described previously (Martens-Uzunova and Schaap 2008). The gene encoding the putative pectin methylesterase C (*pmeC*) was missing on the Affymetrix chips, and therefore missed previously, but the RNA-seq study clearly indicated that induction of *pmeC* on GA is GaaR dependent. Inspection of the promoter regions of the newly identified members of the GA-regulon indicated the presence of putative GaaR binding sites in the promoter regions of most genes (Table 1), enabling us to expand the GA-regulon to a larger set of genes.

Table 1 RNA-seq analysis on GA of the genes that depend on GaaR for induction. Expression values (FPKM) are averages of duplicates. Fold changes ≥2 and p-values ≤0.05 are highlighted. GARE position is given with respect to the transcription start site. Genes identified as the GA-regulon by Martens-Uzunova and Schaap [7] are marked with an asterisk (*[1])

NRRL3 protein ID	CBS 513.88 gene ID	Gene name	Ref GA 2h	*ΔgaaR* GA 2h	Fold change Ref/*ΔgaaR* GA 2h	p-value	GARE (CCNCCAA) position
NRRL3_00958	An14g04280	*gatA**[1]	888.35	13.32	66.69	1.54E-03	+ strand -360
NRRL3_08663	An03g01620	GA transporter (putative)*[1]	106.09	30.34	3.50	1.25E-01	+ strand -673
NRRL3_04281	An07g00780	GA transporter (putative)*[1]	90.41	1.86	48.74	7.77E-03	- strand -42 and -994
NRRL3_05650	An02g07710	*gaaA* *[1]	2599.98	117.53	22.12	1.69E-04	+ strand -414 and -100
NRRL3_06890	An16g05390	*gaaB* *[1]	11309.00	344.03	32.87	1.88E-03	+ strand -326
NRRL3_05649	An02g07720	*gaaC* *[1]	5658.32	106.21	53.27	2.98E-04	- strand -292 and -606
NRRL3_10050	An11g01120	*gaaD* *[1]	8104.43	506.79	15.99	7.01E-03	- strand -538, -583, -801 and -813
NRRL3_03144	An12g07500	*pgaX* *[1]	698.90	24.27	28.80	1.19E-02	+ strand -388
NRRL3_09810	An11g04040	*pgxA* *[1]	10.65	0.34	31.32	9.10E-03	- strand -594
NRRL3_08281	An03g06740	*pgxB* *[1]	200.31	12.39	16.17	2.62E-02	- strand -298 and -823
NRRL3_05260	An02g12450	*pgxC* *[1]	99.93	4.10	24.40	6.24E-04	+ strand -268 and - strand -642
NRRL3_06053	An02g02540	*paeA* (putative) *[1]	522.81	22.99	22.75	4.57E-03	+ strand -1238
NRRL3_04916	An07g08940	*paeB* (putative) *[3]	13.41	10.57	1.27	7.42E-01	
NRRL3_08325	An03g06310	*pmeA*	6.54	0.42	15.75	1.18E-02	+ strand -983 and - strand -308
NRRL3_07470	An04g09690	*pmeB* (putative)	30.16	4.67	6.46	1.41E-02	+ strand -389
NRRL3_05252	An02g12505	*pmeC* (putative) *[2]	558.37	24.68	22.62	4.20E-03	+ strand -275, -246 and -35
NRRL3_02571	An01g11520	*pgaI*	56.38	6.56	8.59	6.96E-04	+ strand -221
NRRL3_05859	An02g04900	*pgaB* *[3]	15.10	3.11	4.86	6.74E-02	- strand -753 and -934
NRRL3_08805	An05g02440	*pgaC*	5.26	0.59	8.99	3.65E-02	+ strand -374, -196 and -865
NRRL3_02835	An01g14670	*pgaE* *[3]	4.26	2.40	1.78	4.12E-01	
NRRL3_00965	An14g04370	*pelA* *[1]	56.54	9.74	5.80	2.12E-04	
NRRL3_09811	An11g04030	*pelC*	0.51	0.00	NA	4.77E-03	
NRRL3_01237	An19g00270	*pelD*	18.95	0.34	55.74	6.03E-04	- strand -409 and -465
NRRL3_04153	An15g07160	*pelF* *[3]	35.48	37.02	0.96	8.73E-01	- strand -644
NRRL3_10559	An18g04810	*rgxC* (putative)	20.00	0.90	22.22	1.26E-02	+ strand -880 and -852 and - strand -250

NRRL3_00684	An14g01130	*rglA*	5.77	1.03	5.63	9.23E-03	- strand -188
NRRL3_01606	An01g00330	*abfA* *[3]	87.96	59.62	1.48	5.81E-01	
NRRL3_10865	An08g01710	*abfC* (putative) *[1]	201.62	67.16	3.00	4.04E-02	
NRRL3_07094	An16g02730	*abnD* (putative)	4.57	1.53	2.99	2.66E-02	+ strand -246
NRRL3_02479	An01g10350	*lacB* (putative)	137.63	41.24	3.34	1.36E-02	
NRRL3_11738	An06g00290	*lacC* (putative)	28.91	9.08	3.19	3.55E-02	+ strand -267
NRRL3_10643	An18g05940	*galA*	105.64	29.24	3.61	3.53E-02	+ strand -307

*[1] Genes identified as the GA-regulon by Martens-Uzunova and Schaap [7]

*[2] *pmeC* not present on the Affymetrix microarray

*[3] Genes not significantly differentially expressed on GA, but differentially expressed on SBP (see Table 2)

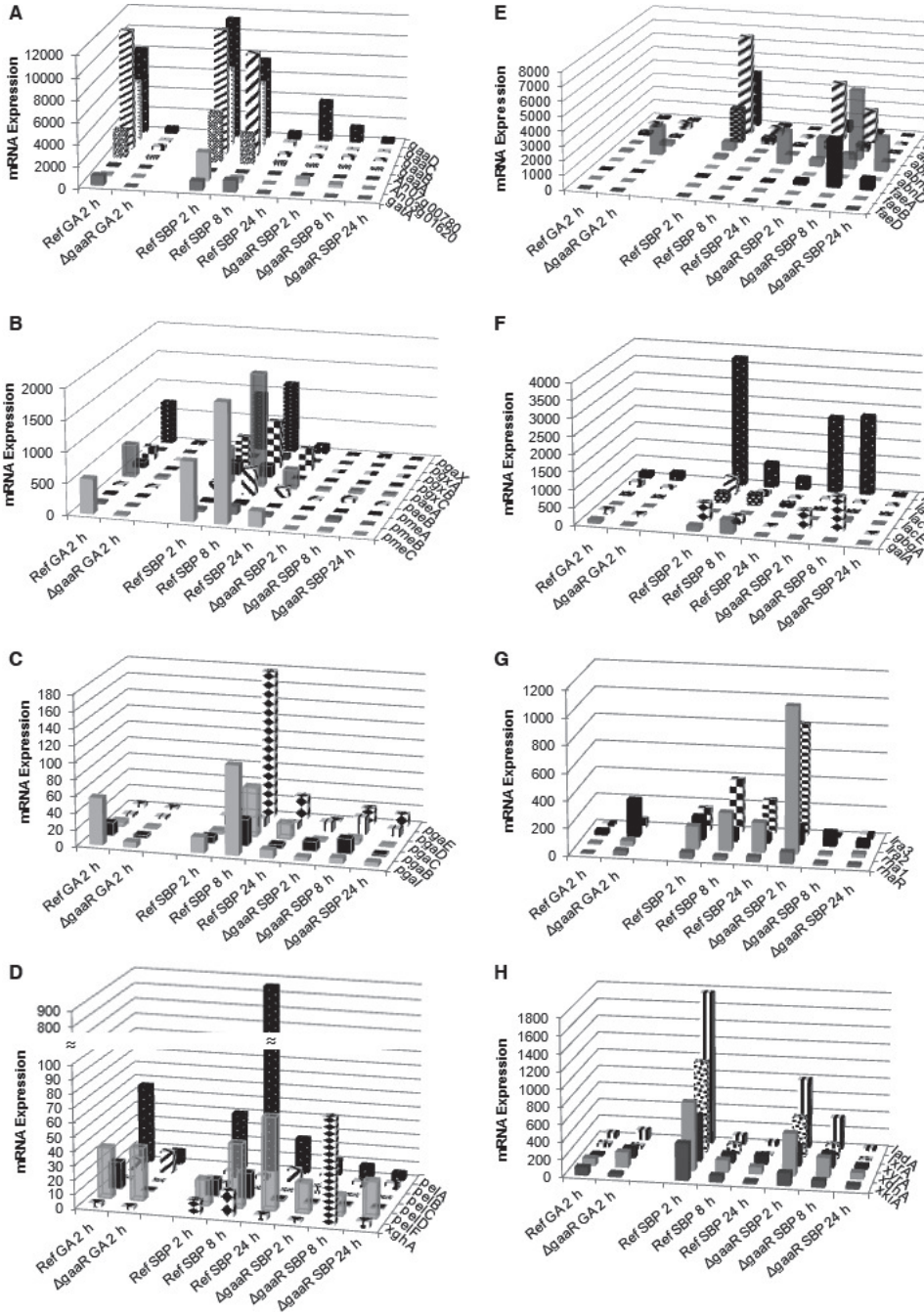

A
mRNA Expression
12000
10000
8000
6000
4000
2000
0
gaaD
gaaA
gaaB
gaaC
xghA
An08g00780
galA
An01g01650
Ref GA 2 h
ΔgaaR GA 2 h
Ref SBP 2 h
Ref SBP 8 h
Ref SBP 24 h
ΔgaaR SBP 2 h
ΔgaaR SBP 8 h
ΔgaaR SBP 24 h
B
mRNA Expression
2000
1500
1000
500
0
pgaX
pgaA
pgaB
pgaC
pgaA
pmeA
pmeB
pmeC
C
mRNA Expression
180
160
140
120
100
80
60
40
20
0
pgaE
pgaD
pgaC
pgaB
pgaI
D
mRNA Expression
900
800
100
90
80
70
60
50
40
30
20
10
0
pelA
pelB
pelC
pelD
pelE
xghA
E
mRNA Expression
8000
7000
6000
5000
4000
3000
2000
1000
0
abfA
abfB
abnA
abnB
abnC
faeA
faeB
faeD
F
mRNA Expression
4000
3500
3000
2500
2000
1500
1000
500
0
lacA
lacB
lacC
lacD
gdgA
galA
G
mRNA Expression
1200
1000
800
600
400
200
0
rgx3
rhg1
rhg2
rhaA
rhaR
H
mRNA Expression
1800
1600
1400
1200
1000
800
600
400
200
0
lacA
xlnA
xlnB
xlnD
xlnR

Figure 2 Transcript levels of pectin utilization genes in *A. niger* reference and *ΔgaaR* on GA or SBP. **A)** GA transporters and GA catabolic pathway enzymes, **B)** exo-polygalacturonases and pectin acetyl- and methylesterases, **C)** endo-polygalacturonases, **D)** pectin lyases and endo-xylogalacturonan hydrolase, **E)** α-L-arabinofuranosidases, arabinan endo-1,5-α-L-arabinofuranosidase, endo-arabinanases, ferulic acid esterases and feruloyl esterase D, **F)** β-galactosidases, galactan 1.3-β-galactosidase and β-1.4-ensogalactanase, **G)** the L-rhamnose regulator *rhaR* and L-rhamnose catabolic pathway enzymes and **H)** the L-arabinose and D-xylose catabolic pathway enzymes. Mycelia of the reference strain (FP-1132.1) and *ΔgaaR* (FP-1126.1) were pre-grown in CM with 2% D-fructose, washed and transferred to MM with 25mM GA or 1% SBP in and incubated for 2, 8 or 24h.

GaaR is required for the induction of genes related to polygalacturonic acid degradation and D-galacturonic acid utilization on complex pectin

Both the strongly reduced growth phenotype on GA and PGA and the expression analysis in *ΔgaaR* suggest that that GaaR is required for GA utilization in *A. niger*. Growth and sporulation of *ΔgaaR* on complex pectins such as SBP was also reduced, but not as severe as on GA and PGA (Figure. 1A). This could be explained by two (not mutually exclusive) hypotheses. The first explanation could be that *A. niger* has alternative mechanisms (independent of GaaR) to induce genes involved in GA utilization. The second possibility is that additional sugars such as L-arabinose, D-galactose, D-xylose or L-rhamnose that are present in SBP are metabolized and used for growth. To gain insight in the expression of pectinase genes in *ΔgaaR* on complex pectin, the reference strain and *ΔgaaR* were transferred from D-fructose to SBP and grown for 2, 8 and 24h before harvesting mycelia and extraction of RNA.

Expression profiles of pectinase genes in the reference strain and *ΔgaaR* were pairwise compared for identical time points (Table 2 and Figure 2, Figure S3). Most of the genes in the GA-regulon, including those required for GA transport and catabolism, are dependent on GaaR for induction on SBP (Figure 2A-D). This observation strongly suggests that *ΔgaaR* is not utilizing GA from SBP. FThe expression of *gaaD*/*larA* can be explained by the dual activity of the enzyme encoded by this gene as both an L-glyceraldehyde reductase and an L-arabinose reductase (Mojzita *et al.* 2010a) and the utilization of L-arabinose from SBP in *ΔgaaR* (see below). The expression profile of exo-polygalacturonases, pectin acetyl- and methylesterases, endo-polygalacturonases and pectin lyases (Table 2 and Figure 2B-D) all acting on the PGA backbone support the conclusion that the GaaR target genes are not induced during growth on SBP in *ΔgaaR*.

Table 2 RNA-seq analysis on SBP of the genes that depend on GaaR for induction. Expression values (FPKM) are averages of duplicates. Fold changes ≥2 and p-values ≤0.05 are highlighted. Genes identified as the GA-regulon by Martens-Uzunova and Schaap [7] are marked with an asterisk (*[1])

NRRL3 protein ID	CBS 513.88 gene ID	Gene name	Ref SBP 2h	*ΔgaaR* SBP 2h	Fold change Ref/*ΔgaaR* SBP 2h	p-value	Ref SBP 8h	*ΔgaaR* SBP 8h	Fold change Ref/*ΔgaaR* SBP 8h	p-value	Ref SBP 24h	*ΔgaaR* SBP 24h	Fold change Ref/*ΔgaaR* SBP 24h	p-value
NRRL3_00958	An14g04280	*gatA**[1]	849.85	12.60	67.45	6.18E-05	1077.70	21.92	49.17	2.84E-03	57.92	1.82	31.91	5.06E-02
NRRL3_08663	An03g01620	GA transporter (putative)*[1]	2647.36	642.68	4.12	1.01E-04	387.81	273.28	1.42	2.41E-01	2.69	3.20	0.84	7.16E-01
NRRL3_04281	An07g00780	GA transporter (putative)*[1]	33.99	7.43	4.58	1.89E-01	14.34	6.24	2.30	5.60E-02	34.49	2.63	13.11	8.63E-03
NRRL3_05650	An02g07710	*gaaA* *[1]	4649.59	70.19	66.24	2.16E-03	2785.77	78.98	35.27	1.73E-03	215.02	24.06	8.94	9.85E-02
NRRL3_06890	An16g05390	*gaaB* *[1]	11722.91	113.85	102.97	2.38E-03	9634.93	229.60	41.96	2.72E-04	208.56	65.62	3.18	8.85E-02
NRRL3_05649	An02g07720	*gaaC* *[1]	7306.08	92.88	78.66	2.95E-05	6041.04	78.17	77.29	2.63E-03	527.35	34.38	15.34	1.02E-01
NRRL3_10050	An11g01120	*gaaD* *[1]	11412.45	3807.08	3.00	5.06E-03	7573.39	1434.68	5.28	8.61E-04	621.40	409.01	1.52	4.46E-01
NRRL3_03144	An12g07500	*pgaX* *[1]	948.28	5.71	166.22	1.16E-02	1154.06	21.72	53.13	4.37E-03	133.41	4.97	26.87	1.59E-03
NRRL3_09810	An11g04040	*pgxA* *[1]	19.61	0.11	186.71	1.72E-02	72.87	0.60	121.44	2.09E-03	3.23	0.00	NA	NA
NRRL3_08281	An03g06740	*pgxB* *[1]	483.28	3.15	153.66	1.81E-02	784.73	13.84	56.72	6.23E-02	364.35	3.40	107.32	4.37E-02
NRRL3_05260	An02g12450	*pgxC* *[1]	206.37	3.08	67.11	8.92E-03	191.37	4.23	45.29	2.90E-04	7.90	8.21	0.96	7.98E-01
NRRL3_06053	An02g02540	*paeA* (putative) *[1]	585.00	14.43	40.55	1.48E-02	1836.37	30.62	59.97	1.75E-02	300.70	11.83	25.43	1.46E-02
NRRL3_04916	An07g08940	*paeB* (putative) *[3]	166.46	7.51	22.16	1.47E-02	427.08	62.09	6.88	9.60E-02	151.00	10.65	14.18	9.63E-04
NRRL3_08325	An03g06310	*pmeA*	15.92	1.30	12.24	3.81E-02	27.43	0.99	27.70	3.18E-03	0.90	0.12	7.50	2.43E-01
NRRL3_07470	An04g09690	*pmeB* (putative)	41.34	4.84	8.55	3.34E-02	130.79	52.94	2.47	1.77E-02	33.31	2.06	16.17	2.20E-01
NRRL3_05252	An02g12505	*pmeC* (putative) *[2]	957.92	7.59	126.29	1.59E-02	1917.18	26.77	71.63	1.09E-02	249.81	0.40	624.53	1.67E-02

NRRL3_02571	An01g11520	*pgaI*	18.75	4.38	4.29	1.44E-02	106.37	6.59	16.14	8.39E-03	9.88	4.83	2.04	4.97E-01
NRRL3_05859	An02g04900	*pgaB* *[3]	7.84	11.01	0.71	2.52E-01	31.37	17.51	1.79	3.24E-02	4.29	2.07	2.07	3.62E-01
NRRL3_08805	An05g02440	*pgaC*	2.97	0.65	4.60	3.05E-01	58.43	6.22	9.40	2.60E-02	17.67	0.41	43.09	2.87E-01
NRRL3_02835	An01g14670	*pgaE* *[3]	3.16	1.35	2.34	8.30E-02	177.86	19.69	9.03	3.46E-03	28.57	15.13	1.89	1.99E-01
NRRL3_00965	An14g04370	*pelA* *[1]	40.64	11.09	3.67	1.14E-02	861.91	9.02	95.55	2.19E-02	25.43	5.41	4.70	1.82E-01
NRRL3_09811	An11g04030	*pelC*	2.39	4.58	0.52	4.27E-01	0.00	0.00	NA	NA	0.00	0.00	NA	NA
NRRL3_01237	An19g00270	*pelD*	11.72	0.61	19.37	1.10E-02	17.97	0.58	31.24	1.26E-03	1.28	0.42	3.05	8.54E-03
NRRL3_04153	An15g07160	*pelF* *[3]	18.03	21.02	0.86	4.67E-01	44.85	14.09	3.18	2.56E-02	64.95	25.60	2.54	5.26E-04
NRRL3_10559	An18g04810	*rgxC* (putative)	86.41	0.94	92.41	6.26E-02	207.96	3.76	55.38	4.39E-02	116.98	1.11	105.39	1.02E-01
NRRL3_00684	An14g01130	*rglA*	2.46	0.00	NA	NA	6.60	2.32	2.84	2.96E-02	3.98	2.30	1.73	6.60E-01
NRRL3_01606	An01g00330	*abfA* *[3]	3864.76	705.67	5.48	6.85E-03	133.72	176.93	0.76	3.59E-01	1.56	4.72	0.33	1.76E-01
NRRL3_10865	An08g01710	*abfC* (putative) *[1]	2406.40	527.90	4.56	2.46E-02	441.88	483.86	0.91	2.48E-01	97.31	36.96	2.63	2.77E-02
NRRL3_07094	An16g02730	*abnD* (putative)	3.92	2.82	1.39	1.06E-01	14.38	6.20	2.32	4.33E-02	55.45	12.53	4.43	2.40E-01
NRRL3_02479	An01g10350	*lacB* (putative)	487.09	48.89	9.96	1.33E-03	115.20	24.31	4.74	8.71E-03	13.85	8.31	1.67	2.24E-01
NRRL3_11738	An06g00290	*lacC* (putative)	297.53	12.04	24.71	6.23E-03	314.50	8.56	36.74	3.61E-02	113.92	2.81	40.61	3.74E-02
NRRL3_10643	An18g05940	*galA*	154.02	30.49	5.05	1.82E-02	398.55	26.21	15.21	2.85E-04	19.05	20.72	0.92	5.82E-01

*[1] Genes identified as the GA-regulon by Martens-Uzunova and Schaap [7]

*[2] *pmeC* not present on the Affymetrix microarray

*[3] Genes not significantly differentially expressed in on GA, but differentially expressed on SBP

The results described above indicate that the residual growth of *ΔgaaR* on SBP is due to the utilization of other monosaccharides released from SBP. Analysis of the monosaccharide composition of the SBP used in this study was performed as described previously (Santander *et al.* 2013) and showed that it contains 55 mol% GA, as well as 17 mol% L-arabinose, 16 mol% D-galactose and 10 mol% L-rhamnose. Analysis of the expression of the genes involved in the degradation of RG-I such as exo-rhamnogalacturonases (*rgx*), rhamnogalacturonases (*rhg*), rhamnogalacturonan acetyl esterases (*rgae*), rhamnogalacturonyl hydrolases (*urhg*), arabinofuranosidases (*abf*), endo-arabinanases (*abn*), ferulic acid esterases (*fae*) and β-galactosidases (*lac*), and the genes responsible for catabolism of L-rhamnose, L-arabinose and D-xylose showed that these genes were still expressed in *ΔgaaR* (Figure. 2E-H, Figure S3), indicating that the degradation and metabolism of RG-I support the growth of *ΔgaaR* on SBP.

A clustering analysis of the expression of genes encoding the (putative) GA transporters, GA catabolic pathway genes and pectinases provided further insight in the groups of co-regulated genes (Figure 3). Clusters E and G consist of genes that are members of the GA-regulon (Table 1) and represent genes involved in the release and utilization of GA. Cluster F also consists mostly of genes that are part of the GA-regulon (Table 1, 2). Genes in Cluster F, like genes in Clusters E and G, are expressed in the reference strain on GA and SBP at 2 and 8h, but unlike genes in Clusters E and G also expressed in the *ΔgaaR* strain on SBP at 2 and 8h. Cluster F mainly includes pectinases acting on RG-I side-chains. Their expression profile indicates that they are regulated by GaaR as well as other TFs involved in pectin degradation. Genes in Clusters A, B, C and D are generally expressed in a GaaR independent fashion and represent pectinases acting on RG-I and XGA. Pectinase genes of Cluster D are predominantly expressed in the *ΔgaaR* strain on SBP at 2 and 8h. Genes in Clusters A, B and C are expressed predominantly in the reference strain and *ΔgaaR* on SBP at 24h or in *ΔgaaR* on GA, suggesting that these genes are likely induced on starvation or derepressed conditions.

In conclusion, in this paper we showed that the conserved Zn_2Cys_6 TF GaaR of *A. niger* is required for the utilization of GA and PGA. We also showed that GaaR is essential for GA utilization from complex pectic substrates and that residual growth of *ΔgaaR* on complex pectins is likely due to induction of pectinases releasing L-rhamnose from the RG-I backbone and L-arabinose and D-galactose from the RG-I “hairy regions”. These monosaccharides are metabolized independently of *gaaR*. With the identification of the GaaR in *A. niger*, we identified the missing link to further understand the interplay between several TFs involved in

plant cell wall degradation. Insight in the regulation of pectin degradation and GA utilization in *A. niger* can help in exploiting *A. niger* for more efficient pectinase production.

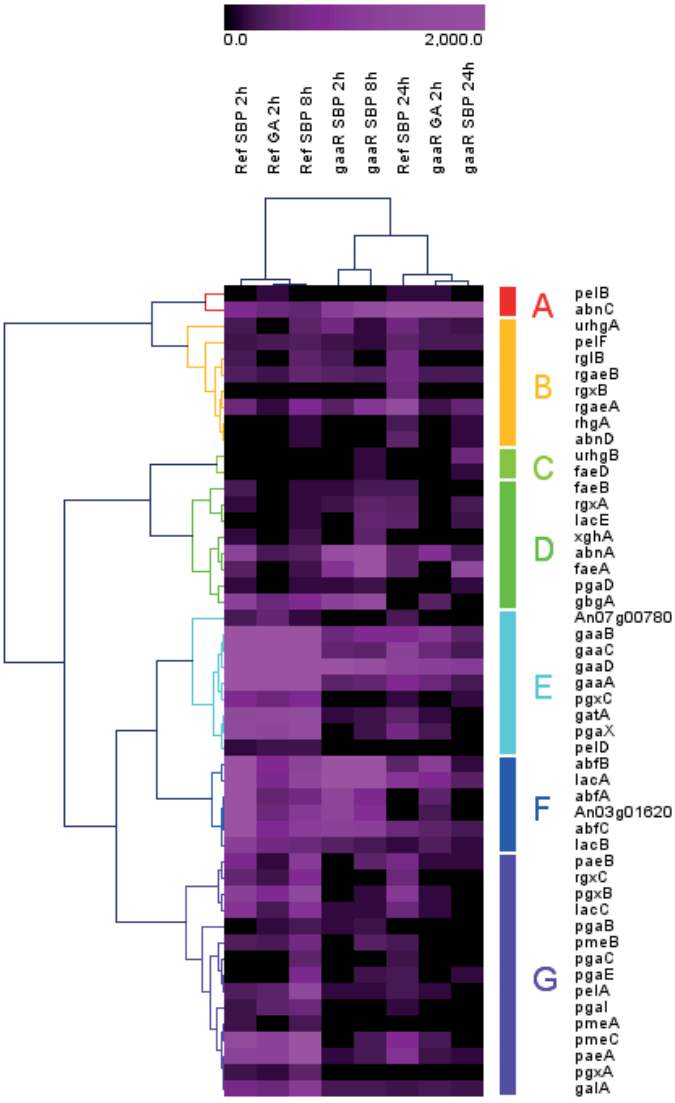


Figure 3 Hierarchical clustering of pectin utilization genes according to their expression in the reference strain (FP-1132.1) and *ΔgaaR* (FP-1126.1) on GA and SBP. The color code displayed represents the transcript levels of the genes. Clusters E and G include genes that are members of the GA-regulon.

Acknowledgements

EA and JN were supported by a grant from BE-Basic (Flagship 10) and the China Scholarship Council, respectively. JEK and MVAP were supported by a grant of the Dutch Technology Foundation STW, Applied Science division of NWO and the Technology Program of the Ministry of Economic Affairs 016.130.609 to RPdV. We thank Patricia vanKuyk and Mark Arentshorst for constructing the *kusA* deletion strains. The pectin composition analysis was conducted at Complex Carbohydrate Research Centre and supported by the Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, U.S. Department of Energy grant (DE-FG02-93ER20097) to Parastoo Azadi.

References

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

Arentshorst, M., J. Niu and A. F. J. Ram, 2015 Efficient Generation of *Aspergillus niger* Knock Out Strains by Combining NHEJ Mutants and a Split Marker Approach, pp. 263-272 in *Genetic Transformation Systems in Fungi*, edited by M. A. van den Berg and K. Maruthachalam. Springer International Publishing, Cham, Switzerlands.

Arentshorst, M., A. F. Ram and V. Meyer, 2012 Using non-homologous end-joining-deficient strains for functional gene analyses in filamentous fungi. Methods Mol Biol 835**:** 133-150.

Battaglia, E., L. Visser, A. Nijssen, G. J. van Veluw, H. A. Wosten *et al.*, 2011 Analysis of regulation of pentose utilisation in *Aspergillus niger* reveals evolutionary adaptations in Eurotiales. Stud Mycol 69**:** 31-38.

Carvalho, N. D., M. Arentshorst, M. Jin Kwon, V. Meyer and A. F. Ram, 2010 Expanding the ku70 toolbox for filamentous fungi: establishment of complementation vectors and recipient strains for advanced gene analyses. Appl Microbiol Biotechnol 87**:** 1463-1473.

Chang, P. K., and K. C. Ehrlich, 2013 Genome-wide analysis of the Zn(II)(2)Cys(6) zinc cluster-encoding gene family in *Aspergillus flavus*. Appl Microbiol Biotechnol 97**:** 4289-4300.

Coutinho, P. M., M. R. Andersen, K. Kolenova, P. A. vanKuyk, I. Benoit *et al.*, 2009 Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. Fungal Genet Biol 46 Suppl 1**:** S161-S169.

de Vries, R. P., J. Jansen, G. Aguilar, L. Parenicova, V. Joosten *et al.*, 2002 Expression profiling of pectinolytic genes from *Aspergillus niger*. FEBS Lett 530**:** 41-47.

Edgar, R., M. Domrachev and A. E. Lash, 2002 Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30**:** 207-210.

Espagne, E., O. Lespinet, F. Malagnac, C. Da Silva, O. Jaillon *et al.*, 2008 The genome sequence of the model ascomycete fungus *Podospora anserina*. Genome Biol 9**:** R77.

Goosen, T., G. Bloemheuvel, C. Gysler, D. A. de Bie, H. W. van den Broek *et al.*, 1987 Transformation of *Aspergillus niger* using the homologous orotidine-5'-phosphate-decarboxylase gene. Curr Genet 11**:** 499-503.

Gruben, B. S., M. Zhou and R. P. de Vries, 2012 GalX regulates the D-galactose oxido-reductive pathway in *Aspergillus niger*. FEBS Lett 586**:** 3980-3985.

Gruben, B. S., M. Zhou, A. Wiebenga, J. Ballering, K. M. Overkamp *et al.*, 2014 *Aspergillus niger* RhaR, a regulator involved in L-rhamnose release and catabolism. Appl Microbiol Biotechnol 98**:** 5531-5540.

Kowalczyk, J. E., I. Benoit and R. P. de Vries, 2014 Regulation of plant biomass utilization in *Aspergillus*. Adv Appl Microbiol 88**:** 31-56.

Kuivanen, J., D. Mojzita, Y. Wang, S. Hilditch, M. Penttila *et al.*, 2012 Engineering filamentous fungi for conversion of D-galacturonic acid to L-galactonic acid. Appl Environ Microbiol 78**:** 8676-8683.

Langmead, B., C. Trapnell, M. Pop and S. L. Salzberg, 2009 Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10**:** R25.

Leijdekkers, A. G., J. H. Huang, E. J. Bakx, H. Gruppen and H. A. Schols, 2015 Identification of novel isomeric pectic oligosaccharides using hydrophilic interaction chromatography coupled to traveling-wave ion mobility mass spectrometry. Carbohydr Res 404**:** 1-8.

Li, B., and C. N. Dewey, 2011 RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12**:** 323.

Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25**:** 1754-1760.

Martens-Uzunova, E. S., and P. J. Schaap, 2008 An evolutionary conserved d-galacturonic acid metabolic pathway operates across filamentous fungi capable of pectin degradation. Fungal Genet Biol 45**:** 1449-1457.

Martens-Uzunova, E. S., and P. J. Schaap, 2009 Assessment of the pectin degrading enzyme network of *Aspergillus niger* by functional genomics. Fungal Genet Biol 46 Suppl 1**:** S170-S179.

Meyer, V., M. Arentshorst, A. El-Ghezal, A. C. Drews, R. Kooistra *et al.*, 2007 Highly efficient gene targeting in the *Aspergillus niger* kusA mutant. J Biotechnol 128**:** 770-775.

Mohnen, D., 2008 Pectin structure and biosynthesis. Curr Opin Plant Biol 11**:** 266-277.

Mojzita, D., M. Penttila and P. Richard, 2010a Identification of an L-arabinose reductase gene in *Aspergillus niger* and its role in L-arabinose catabolism. J Biol Chem 285**:** 23622-23628.

Mojzita, D., M. Wiebe, S. Hilditch, H. Boer, M. Penttila *et al.*, 2010b Metabolic engineering of fungal strains for conversion of D-galacturonate to meso-galactarate. Appl Environ Microbiol 76**:** 169-175.

Niu, J., T. G. Homan, M. Arentshorst, R. P. de Vries, J. Visser *et al.*, 2015 The interaction of induction and repression mechanisms in the regulation of galacturonic acid-induced genes in *Aspergillus niger*. Fungal Genet Biol 82**:** 32-42.

Raulo, R., M. Kokolski and D. B. Archer, 2016 The roles of the zinc finger transcription factors XlnR, ClrA and ClrB in the breakdown of lignocellulose by *Aspergillus niger*. AMB Express 6**:** 5.

Sambrook, J., and D. Russell, 2001 *Molecular cloning: a laboratory manual*. Cold Spring Harbor Press, New York, NY, USA.

Santander, J., T. Martin, A. Loh, C. Pohlenz, D. M. Gatlin, 3rd *et al.*, 2013 Mechanisms of intrinsic resistance to antimicrobial peptides of *Edwardsiella ictaluri* and its influence on fish gut inflammation and virulence. Microbiology 159**:** 1471-1486.

Sloothaak, J., M. Schilders, P. J. Schaap and L. H. de Graaff, 2014 Overexpression of the *Aspergillus niger* GatA transporter leads to preferential use of D-galacturonic acid over D-xylose. AMB Express 4**:** 66.

Sturn, A., J. Quackenbush and Z. Trajanoski, 2002 Genesis: cluster analysis of microarray data. Bioinformatics 18**:** 207-208.

Tani, S., T. Kawaguchi and T. Kobayashi, 2014 Complex regulation of hydrolytic enzyme genes for cellulosic biomass degradation in filamentous fungi. Appl Microbiol Biotechnol 98**:** 4829-4837.

Trapnell, C., B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan *et al.*, 2010 Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 28**:** 511-515.

van Peij, N. N., J. Visser and L. H. de Graaff, 1998 Isolation and analysis of xlnR, encoding a transcriptional activator co-ordinating xylanolytic expression in *Aspergillus niger*. Mol Microbiol 27**:** 131-142.

Wiebe, M. G., D. Mojzita, S. Hilditch, L. Ruohonen and M. Penttila, 2010 Bioconversion of D-galacturonate to keto-deoxy-L-galactonate (3-deoxy-L-threo-hex-2-ulosonate) using filamentous fungi. BMC Biotechnol 10**:** 63.

Zhang, L., R. J. Lubbers, A. Simon, J. H. Stassen, P. R. Vargas Ribera *et al.*, 2016 A novel Zn2 Cys6 transcription factor BcGaaR regulates D-galacturonic acid utilization in *Botrytis cinerea*. Mol Microbiol 100**:** 247-262.

Supporting information

The supplementary material of this chapter are available via

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5111758/ and comprises the following:

Table S1 Strains used in this study

Table S2 Primers used in this study. Overlapping sequences for fusion PCR are written in bold.

Table S3A RNA-seq analysis of pectinases on GA and SBP. Expression values (FPKM) are averages of duplicates. Fold changes ≥2 and p-values ≤0.05 are highlighted.

Table S3B RNA-seq analysis of pectinases on GA and SBP. Expression values (FPKM) are averages of duplicates. Fold changes ≥2 and p-values ≤0.05 are highlighted.

Table S3C RNA-seq analysis of pectinases on GA and SBP. Expression values (FPKM) are averages of duplicates. Fold changes ≥2 and p-values ≤0.05 are highlighted.

Table S3D RNA-seq analysis of pectinases on GA and SBP. Expression values (FPKM) are averages of duplicates. Fold changes ≥2 and p-values ≤0.05 are highlighted.

Figure S1 Alignment of AnGaaR and BcGaaR using EMBOSS Needle with standard settings (http://www.ebi.ac.uk/Tools/psa/emboss_needle/)

Figure S2 Verification of the gaaR deletion strain in the MA234.1 background. A) Strategy and primer design for disruption the gaaR gene using the split marker method [27]. Primers P1 till P8 correspond to primers gaaRP1f, gaaRP2r, gaaR3Pf, gaaRP4r, hygP6f, hygP9r, hygP8f and hygP7r in Table S2. B) Schematic representation of the gaaR locus in the reference strain and after gaaR deletion. Predicted sizes of the DNA fragment hybridizing with the indicated probes are shown. C) Southern blot analysis of genomic DNA of MA234.1 (lane 1), JN35.1 (lane 2), JN35.2 (lane 3), JN36.1 (lane 4), JN37.4 (lane 5) and JN37.5 (lane 6). Left panel: agarose gel stained with ethidium bromide. Marker size (M, in kb) is indicated. Right panel: Southern blot after hybridization. D) Growth analysis of the reference strain MA234.1, gaaR deletion strains and complemented strains on MM with 50 mM or 1% different carbon sources.

Figure S3 Transcript levels of pectinases acting on RG-I backbone in A. niger reference and ΔgaaR on GA or SBP. A) exo-rhamnogalacturonases and rhamnogalacturonase A B) rhamnogalacturonan lyases, rhamnogalacturonan acetyl esterases and rhamnogalacturonyl hydrolases. Mycelia of the reference strain (FP-1132.1) and ΔgaaR (FP-1126.1) were pre-grown in CM with 2% D-fructose, washed and transferred to MM with 25mM GA or 1% SBP in and incubated for 2, 8 or 24h.

Chapter 6

An evolutionarily conserved transcriptional activator-repressor module controls expression of genes for D-galacturonic acid utilization in *Aspergillus niger*

Jing Niu, Ebru Alazi, Ian D. Reid, Mark Arentshorst, Peter J. Punt, Jaap Visser, Adrian Tsang, Arthur F.J. Ram

Published in Genetics. 2017 Jan; 205(1): 169-183

Abstract

The expression of genes encoding extracellular polymer-degrading enzymes and the metabolic pathways required for carbon utilization in fungi are tightly controlled. The control is mediated by transcription factors that are activated by the presence of specific inducers, which are often monomers or monomeric derivatives of the polymers. A D-galacturonic acid-specific transcription factor named GaaR was recently identified and shown to be an activator for the expression of genes involved in galacturonic acid utilization in *Botrytis cinerea* and in *Aspergillus niger*. Using a forward genetic screen, we isolated mutants that constitutively express GaaR-controlled genes. Reasoning that mutations in the *gaaR* gene would lead to a constitutively activated transcription factor, the *gaaR* gene in eleven of the constitutive mutants was sequenced, but no mutations in *gaaR* were found. Full genome sequencing of five constitutive mutants revealed allelic mutations in one particular gene encoding a previously uncharacterized protein (NRRL3_08194). The protein encoded by NRRL3_08194 shows homology to the repressor of the quinate utilization pathway identified previously in *Neurospora crassa (qa-1S)* and *Aspergillus nidulans* (QutR). Deletion of NRRL3_08194 in combination with RNA-seq analysis showed that the NRRL3_08194 deletion mutant constitutively expresses genes involved in galacturonic acid utilization. Interestingly, NRRL3_08194 is located next to *gaaR* (NRRL3_08195) in the genome. The homology to the quinate repressor, the chromosomal clustering, and the constitutive phenotype of the isolated mutants suggest that NRRL3_08194 is likely to encode a repressor, which we name GaaX. The GaaR-GaaX module and its chromosomal organization is conserved among ascomycetes filamentous fungi, resembling the quinate utilization activator-repressor module in amino-acid sequence and chromosomal organization.

Key words: gene regulation; galacturonic acid; repressor protein; genomics; transcriptomics; pectin

Introduction

The filamentous fungus *Aspergillus niger* is an important producer of pectin-degrading enzymes that are used in industrial applications including in food and feed processing (Kashyap *et al*. 2001; Khan *et al*. 2013). In nature, *A. niger* is a saprotrophic fungus that feeds on organic matter from decaying plants. The major carbon sources in plant cells are the storage polysaccharides starch, and less frequently inulin, as well as the cell wall polymers cellulose, hemicelluloses and pectin. Of the different plant polysaccharides, pectin has the most complex structure. Pectin is made up of four substructures including homogalacturonan, xylogalacturonan, rhamnogalacturonan I and rhamnogalacturonan II. The abundance of each substructure varies with plant species, but typically homogalacturonan is the most abundant polysaccharide in pectin (65%) followed by rhamnogalacturonan I (25-30%). Xylogalacturonan and rhamnogalacturonan II comprise less than 10% of the total pectin (Mohnen 2008).

Utilization of plant polysaccharides by fungi, including *A. niger*, is accomplished by tightly controlled secretion of extracellular enzymes that degrade the polymers into monosaccharides or oligosaccharides that are taken up and catabolised by the fungus. The controlled regulation is not only confined to the expression of genes encoding extracellular proteins. It also includes the controlled expression of genes encoding specific sugar transporters to guarantee efficient uptake of the liberated sugars and the intracellular catabolic pathway enzymes. The precise induction of the network of genes encoding substrate-specific enzymes, transporters and catabolic pathway enzymes has so far been shown to be mediated via $Zn(II)_2Cys_6$ transcription factors. Specific transcription factors in *A. niger* regulating the utilization of the major polysaccharides have been characterized. They include AmyR, the regulator for starch utilization (Petersen *et al*. 1999; Yuan *et al*. 2008a, vanKuyk *et al.* 2012); InuR for inulin (Yuan *et al*. 2008b); ManR, ClrA and ClrB for cellulose (Raulo *et al*. 2016); XlnR for xylan (Van Peij *et al*. 1998; Battaglia *et al*. 2014); AraR for arabinan (Battaglia *et al*. 2014); RhaR for rhamnose (Gruben *et al*. 2014); and GaaR for polygalacturonic acid (PGA) (Alazi *et al*. 2016). These transcription factors exert coordinated regulation of the target genes by interacting with conserved binding sites that are located upstream of the target genes. Computational analysis has been used to identify the galacturonic acid responsive element (GARE) of GA-induced genes (Martens-Uzunova and Schaap 2008). The predicted sequence (CCNCCAA) was shown to be required for the induction of GA-responsive genes in *A. niger* (Niu *et al*. 2015) and *Botrytis cinerea* (Zhang *et al*. 2016). Furthermore, using the

yeast one-hybrid method, it was shown in *B. cinerea* that the GaaR transcription factor interacts specifically with the GARE (Zhang *et al*. 2016).

Phenotypic characterization of mutants lacking the GA-regulator in both *B. cinerea* and *A. niger* has shown that GaaR is required for growth on GA (Zhang *et al*. 2016; Alazi *et al*. 2016). Expression analysis in both fungi confirmed that GaaR is required for the induced expression of GA-responsive genes. On complex pectins, growth of *B. cinerea* and *A. niger gaaR* deletion mutants was severely reduced and genome-wide expression analysis in *A. niger* revealed that the residual growth on pectin is likely due to the GaaR-independent expression of pectinases acting on arabinofuranosyl- and galactopyranosyl-containing side chains in rhamnogalacturonan (Alazi *et al*. 2016).

In addition to the transcription regulation via GaaR, GA-responsive genes are also under carbon catabolite repression (CCR) control (de Vries *et al*. 1999; de Vries *et al.* 2002). In filamentous fungi, CreA mediates CCR (Dowzer *et al*. 1991; Ruijter and Visser*,* 1997). In *A. niger*, CreA also exerts CCR control on GA-responsive genes (de Vries *et al*. 1999; Niu *et al*. 2015). Using an *'in vivo'* reporter construct consisting of the promoter of the GA-inducible *pgaX* gene (P*pgaX*) and the acetamidase (*amdS*) gene as a reporter, both the specific induction of *pgaX* on GA as well as the carbon repression of *pgaX* via CreA had been demonstrated (Niu *et al*. 2015). In this study, we have used the P*pgaX-amdS* reporter strain to isolate mutants displaying constitutive expression of GA-responsive genes. Analysis of the mutants resulted in the identification of a protein that likely acts as a repressor that specifically inhibits GaaR transcription activation activity under non-inducing conditions.

Materials and Methods

Strains, media and growth conditions

All strains in this study are listed in Table 1. Strains were grown in liquid or solidified (1.5% agar) minimal medium (MM) containing 7 mM KCl, 8 mM KH_2PO_4, 70 mM $NaNO_3$, 2 mM MgSO4 (pH adjusted to pH 5.5) as described (Bennett and Lasure, 1991). MM was supplemented with 50 mM glucose, 50 mM D-galacturonic acid, 50 mM fructose or 50 mM sorbitol as carbon source. Complete medium (CM) was also used and consists of MM supplemented with 0.1% casamino acids and 0.5% $w.v^{-1}$ yeast extract and 50 mM glucose. MM agar plates containing acetamide as sole nitrogen source were made as described previously (Arentshorst *et al*. 2012)

Table 1. *Aspergillus niger* strains used in this study.

Name	Genotype/description	Reference/source
N402	*cspA1*, derivative of N400	Bos *et al.* 1988
AB4.1	*pyrG*$^-$,derivative of N402	van Hartingsveldt *et al.* 1987
MA234.1	*Δku70::DR_amdS_DR* in MA169.4	Alazi *et al.* 2016
MA70.15	*Δku70::amdS* in AB4.1	Meyer et al., 2007
MA299.2	*Δku70* in MA70.15	Niu *et al.* 2015
MA323.1	*Δku70::amdS, ΔnicB*$^-$*, pyrG*$^-$	Niu *et al.* 2016
JC1.5	*pgaX-amdS* in MA299.2, *pyrG*$^+$	Niu *et al.* 2015
JN29.2	*ΔcreA*::*hygB* in JC1.5	Niu *et al.* 2015
JN38	spontaneous mutation S1in JN29.2	This study
JN39	spontaneous mutation S2 in JN29.2	This study
JN42	spontaneous mutation S5 in JN29.2	This study
JN44	spontaneous mutation S7 in JN29.2	This study
JN52	UV1 in JN29.2	This study
JN53	UV2 in JN29.2	This study
JN54	UV3 in JN29.2	This study
JN55	UV4 in JN29.2	This study
JN56	UV5 in JN29.2	This study
JN57	UV6 in JN29.2	This study
JN58	UV7 in JN29.2	This study
JN59	UV8 in JN29.2	This study
JN60	UV9 in JN29.2	This study
JN61	UV10 in JN29.2	This study
JN62	UV11 in JN29.2	This study
JN63	UV12 in JN29.2	This study
JN64	UV13 in JN29.2	This study
JN122.1, JN122.1, JN122.1,	*ΔgaaX::phleo* in JN29.2	This study
JN123.1, JN123.2, JN123.3	*ΔgaaX::hygB* in JC1.5	This study
JN125.1	*ΔgaaX::nicB* in MA323.1	This study
JN126.2, JN126.5, JN126.6	PgaaX::GaaX::GFP::TgaaX in JN125.1	This study
JN127.1, JN127.2, JN127.3	PgaaX::GFP::GaaX::TgaaX in JN125.1	This study

Isolation of mutants with constitutive expression of genes involved in polygalacturonic acid utilization

A. niger strain JN29.2 (Table 1) was used for the selection of mutants with constitutive expression of genes involved in PGA utilization. Spontaneous mutants were obtained by plating out freshly harvested and myracloth filtered conidia (1 x 10^4 conidia per plate) on MM glucose/acetamide plates and incubated at 30℃ for 5 days. In addition, mutants were obtained after mild UV mutagenesis (80% survival) as described (Damveld *et al*. 2008). Individual mutants growing on the primary MM-glucose/acetamide selection plates were purified twice on the MM-glucose/acetamide agar plates. In total, 14 spontaneous mutants and 59 UV-mutants were isolated that grew well on MM-glucose/acetamide agar plates and they were considered to be potential mutants with constitutive expression of genes involved in PGA utilization. To identify mutants constitutively producing PGA degrading enzymes, all 73 mutants were grown by inoculating 5 x 10^7 spores in 50 ml MM-glucose medium for 36 h at

30℃ with shaking (150rpm). Supernatant of each culture was harvested by filtration. The extracellular culture fluid and the mycelia were stored at -80 ℃ for enzymatic assays and RNA extraction, respectively. Ten microlitres supernatant of each sample were spotted on PGA plates made by dissolving 0.2 % PGA (Sigma) in NaAc buffer (pH 4.2) with 1% agarose (Sphaero). The PGA plate assay was modified from the protocol used to detect cellulase activity (Teather and Wood 1982). Plates were incubated at 37℃ for 17 hours after spotting. PGA was stained by flooding the plates with a filter-sterile 0.05% solution of Congo Red (Sigma) dissolved in Milli-Q water for 15 min. The Congo Red solution was then poured off and the plates were washed with Milli-Q water, further treated by flooding with 1 M NaCl for 15 min. The formation of a clear zone of hydrolysis indicated PGA degradation.

The constitutive expression of genes involved in PGA degradation was further determined by northern blot analysis. Total RNA was isolated from eleven UV-mutants and two spontaneous mutants from frozen mycelia using TRIzol® reagent (Invitrogen). Quantification and purity assessment of total RNA was done by spectrophotometric method (NanoDrop 2000, Thermo Scientific). Total RNA, 3.5 μg, was loaded per sample and blotted to a Hybond^{TM}-N^{+} nylon membrane (Amersham, GE Healthcare) followed by hybridization with [α-32P]-dCTP labelled probes (Rediprime II kit, Amersham, GE Healthcare). Probes were PCR-amplified using the N402 genomic DNA and the primer pairs are listed in Table S1. Standard molecular techniques were applied as described (Sambrook and Russell 2001).

DNA sequencing and data analysis

Sequencing of the *gaaR* gene from eleven constitutive mutants was performed by PCR amplification of the *gaaR* gene including 137 bp upstream and 152 bp downstream sequences using genomic DNA of the mutants as template and primers gaaRP7f and gaaRP8r (Table S1). Genomic DNA was isolated as described (Arentshorst *et al*. 2012). The PCR fragment (2765 bp in size) was sequenced in both directions using *gaaR* sequencing primers (Table S1). Sequencing was performed by Macrogen Europe (Amsterdam, The Netherlands).

Genomic DNA of three spontaneous mutants and two UV mutants was isolated as described (Arentshorst *et al*. 2012) and was further purified with DNA Isolation Kit (MO BIO Laboratories) for whole-genome DNA sequencing. The mutant genomes were sequenced at the McGill University Génome Québec Innovation Centre (Montreal) using the Illumina HiSeq platform to about 50-fold coverage. The DNA reads were aligned to the NRRL3

genome with Bowtie2 (Langmead and Salzberg 2012) and sequence differences were detected with Freebayes (Garrison and Marth 2012).

Deletion of *gaaX* gene

Deletion of the *gaaX* gene (NRRL3_08194) in the JC1.5, JN29.2 and MA323.1 backgrounds (Table 1) was carried out using the split marker approach (Arentshorst *et al*. 2015). The 869 bp 5'-flank and 870 bp 3'-flank regions were PCR amplified with the primers listed in Table S1 using N402 genomic DNA as template. These PCR fragments were used in fusion PCRs with hygromycin, phleomycin resistance genes or the *nicB* gene (Niu et al., 2016) to generate the split marker fragments. After amplification, the 5´flank-*hyg* and 3'flank-*hyg* fragments were transformed to the recipient strain JC1.5, the 5´flank-*phleo* and 3'flank-*phleo* fragments were transformed to the recipient strain JN29.2, and the 5´flank-*nicB* and 3'flank-*nicB* fragments were transformed to the recipient strain MA323.1. Putative *gaaX* disruption strains were purified by two consecutive single colony streaks. Genomic DNA was isolated as described (Arentshorst *et al*. 2012) and Southern blot hybridizations, using PCR-amplified fragments generated with primers listed in Table S1 as probes, were performed to confirm proper deletion and to exclude additional integrations.

Bioreactor cultivation

Controlled bioreactor cultivations for *A. niger* MA234.1 and JN123.1 were performed in 6.6-L BioFlo3000 bioreactors (New Brunswick Scientific) as previously described (Jørgensen *et al*. 2010). Briefly, autoclaved bioreactor vessels were filled with 5 L of sterile MM with 0.75% fructose. During cultivation at 30°C, the controller was set to maintain pH 3 by addition of titrants (2 M NaOH or 1 M HCl). Sterile air was supplied at a rate of 1 L min^{-1}. Prior to inoculation, 1.5 ml of 10% (w/v) filter-sterilized yeast extract was added to enhance conidial germination. Cultures were inoculated with freshly harvested spores at a concentration of 7,0 x 10^8 conidia per liter. To reduce the loss of hydrophobic conidia during germination, the stirrer speed was set to 250 rpm and the culture was aerated via the headspace during the first six hours after inoculation. Subsequently, the stirrer speed was increased to 750 rpm, 0.5 ml of polypropyleneglycol P2000 was added as an antifoam agent and air was supplied via the sparger. Cultures broth was harvested at regular intervals from batch cultures and mycelial biomass was retained by vacuum filtration using glass microfiber filters (Whatman). Both biomass and filtrate were quickly frozen in liquid nitrogen and

subsequently stored at -80°C. Dry biomass concentrations were gravimetrically determined from lyophilized mycelia originating from a known mass of culture broth.

Transcriptome analysis

Mycelia grown in bioreactors to mid-exponential phase were used to isolate RNA using TRIzol® reagent (Invitrogen) and purified with NucleoSpin RNA Clean-up kit (Macherey-Nagel) with DNase treatment. Quantity and quality of the RNA samples were determined with a NanoDrop-2000 spectrophotometer and by RNA gel electrophoresis, respectively. RNA sequencing was conducted by Genome scan (Leiden, the Netherlands). Briefly, mRNA was isolated from the total RNA using NEBNext Ultra Directional RNA Library Prep Kit for Illumina according to the manufacturer's protocol. After fragmentation of the mRNA, cDNA was synthesized using random primers; and after a second strand cDNA synthesis reaction, fragments were ligated to the sequencing adapters. Clustering and DNA sequencing was performed using the Illumina NextSeq 500 SR75. Throughout the manuscript we will refer to *A. niger* Gene IDs based on the most up to date and accurate annotation of the *A. niger* NRRL3 genome (http://genome.fungalgenomics.ca/). The RNA-Seq reads were cleaned by correcting sequencing errors with Rcorrector (Song and Florea 2015), trimming sequencing adapters and low quality sequences with Skewer (Jiang *et al.* 2014), and removing ribosomal RNA with SortMeRNA (Kopylova *et al.* 2012). The cleaned reads were mapped to NRRL3 transcripts and counted with Salmon (Patro *et al.* 2016), and the read counts were analyzed for differences in transcript expression between genotypes with DESeq2 (Love *et al.* 2014).

Construction of strains expressing GaaX-GFP or GFP-GaaX fusion proteins

To construct fusions of GFP to the N-terminus or C-terminus of GaaX, PgaaX_GFP::GaaX_TgaaX and PgaaX_GaaX::GFP_TgaaX constructs were generated using a fusion-PCR approach in which N402 genomic DNA as well as plasmid PagsA_eGFP_TtrpC (Damveld *et al.* 2008) were used as template DNA. For constructing the PgaaX_GFP::GaaX_TgaaX construct, the promoter region of *gaaX* was PCR amplified using primers PgaaX_P7f-NotI and PgaaX_P11r, GFP was PCR amplified from plasmid PagsA_eGFP_TtrpC using primers GFP_P1f and GFP_P3r, *gaaX* and the terminator region of *gaaX* was PCR amplified using primers gaaX_P12f and TgaaX_P10r-NotI, and the three fragments were combined together in a two-step fusion PCR. Two amino acids (Gly-Ala) were introduced as spacer between GFP and GaaX. Subsequently, the fusion fragment was

cloned into vector pJet1.2 to give plasmid pJN34. For the PgaaX_GaaX::GFP_TgaaX construct, *gaaX* with the promoter region of *gaaX* was PCR amplified using primers PgaaX_P7f-NotI and gaaX_P8r, GFP was PCR amplified using primers GFP_P1f and GFP_P2r, the terminator region of *gaaX* was PCR amplified using primers gaaX_P9f and TgaaX_P10r-NotI, and the three fragments were combined together in a two-step fusion PCR. Again, a Gly-Ala spacer was introduced between GaaX and GFP. Subsequently, the fusion fragment was cloned into vector pJet1.2 to give plasmid pJN35.

Plasmids pJN34 and pJN35 were digested by *Not*I, and the fragments containing PgaaX_GFP::GaaX_TgaaX and PgaaX_GaaX::GFP_TgaaX were cloned into pMA334 (Arentshorst *et al*. 2015) to generate pJN36 and pJN37, respectively. The pMA334 plasmid has been designed such that the reporter constructs are targeted to the *pyrG* locus. Plasmids pJN36 and pJN37 were linearized by *Asc*I digestion and purified from gel before transformation to *A. niger* strain JN125.1. Proper integration of the PgaaX_GFP::GaaX_TgaaX or PgaaX::GaaX_GFP_TgaaX fragments at the *pyrG* locus was confirmed by Southern blot using PCR-amplified fragments generated with primers listed in Table S1 as probes.

Microscopy

For microscopic analysis, conidia of strains MA323.1, JN125.1, JN126.2 and JN127.3 were inoculated on coverslips in Petri dishes. Liquid minimal medium (MM) supplemented with 50 mM GA or 50 mM fructose as the carbon source was used. After incubation at 30℃ for 16 h, the coverslips with adherent germlings were mounted upside down on glass slides and observed under a confocal laser scanning microscope (Zeiss Imager, Zeiss, Jena, Germany), equipped with a LSM 5 exciter using 63x objectives. Images were processed by ImageJ with the exact same brightness and contrast adjustments and the median filter (radius 1.0).

Data availability

Strains are listed in Table 1 and are available upon request. Table S2 contains SNPs and indels detected in genomes of mutants. Table S3A contains TPM values of NRRL3 gene models in wild type and the *gaaX* mutant, and Table S3B contains their DEseq2 analysis. The DNA reads described in this study are deposited in the Short Read Archive under accession number SRP078415. The RNA reads described in this study are deposited in the Short Read

Archive under accession number SRP078485. The authors state that all data for confirming the conclusions presented in this article are represented fully within the article.

Results

Mutants constitutively expressing genes related to galacturonic acid utilization

To identify mutants that constitutively express genes related to PGA degradation and GA utilization in *A. niger*, we designed a forward screening procedure using a reporter strain containing a P*pgaX-amdS* reporter construct for positive selection of the desired mutants. We recently showed that the *pgaX* gene is specifically induced by GA, PGA and pectin, allowing the reporter P*pgaX-amdS* strain to grow on acetamide as a nitrogen source when GA, PGA or pectin is present as a carbon source (Niu *et al*. 2015). We also showed that deletion of the CCR protein (CreA) did not result in growth of the P*pgaX*-reporter strain on glucose, indicating that derepression via *creA* deletion was not sufficient to drive P*pgaX-amdS* expression to sustain growth. For the mutant screen, we used the P*pgaX-amdS* reporter strain in the Δ*creA* background (JN29.2) to prevent interference with possible CreA pathway-related repression mechanisms. Spores of *A. niger* strain JN29.2 were UV-mutagenized and surviving spores (80%) were plated on MM-glucose-acetamide plates. After mutagenesis, 59 mutants were isolated based on growth on acetamide. In addition to UV-generated mutants, 14 spontaneous mutants were isolated, resulting in a total of 73 mutants that could grow on glucose/acetamide plates.

To determine whether mutations were *cis*- or *trans*-acting, mutants were cultured in glucose medium for 36 h and the medium was analyzed for polygalacturonase activity. Initial experiments showed that cultivation of JN29.2 in glucose medium resulted in very low polygalacturonase levels and no halo was formed on Congo Red stained PGA plates when culture medium was spotted on a PGA plate (Figure 1A). We reasoned that if the mutation is *trans*-acting, the medium should contain increased levels of both exo- and endo-polygalacturonases leading to the formation of a halo. On the other hand, if the mutation is *cis*-acting, thereby only affecting the P*pgaX-amdS* reporter construct, it would not result in a halo on a PGA plate. Based on this assay, we concluded that the mutations of 65 out of the 73 mutants are *trans*-acting, while the remaining eight mutants carry presumed *cis*-acting mutations.

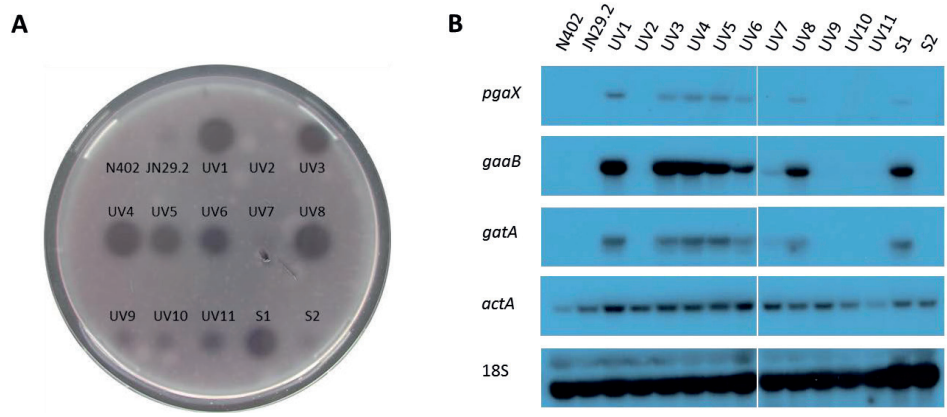


Figure 1. Enzymatic and RNA blot analysis of mutants with constitutive expression of genes involved in polygalacturonic acid utilization. A) 10 µl Supernatant from glucose-grown cultures of reference strains N402 and JN29.2, eleven UV mutants and two spontaneous mutants were spotted on polygalacturonic acid agarose medium to detect polygalacturonase activity. B) Northern blot analysis of selected GA-responsive genes in the reference strains N402 and JN29.2, eleven UV mutants and two spontaneous mutants.

To further demonstrate that the presumed *trans*-acting mutations indeed affected expression of multiple genes related to GA utilization and belonging to the GA-induced genes (Martens-Uzunova *et al*. 2008; Niu *et al*. 2015; Alazi *et al*. 2016), the expression of three GA-induced genes (*pgaX*, *gatA*, *gaaB)* was examined by northern blot analysis in a subset of mutants after growth on glucose. The *pgaX*, *gatA* and *gaaB* genes encode an exo-polygalacturonase, a GA-specific transporter, and the L-galactonic acid dehydratase involved in the GA release from PGA, the uptake of GA, and subsequent metabolism of GA, respectively. As shown in Figure 1B, expression of these genes was not detected in the wild-type (N402) and the parent JN29.2 (*ΔcreA*, P*pgaX-amdS*) strains whereas these genes were expressed in the constitutive mutants that displayed increased polygalacturonase activity. The presumed *cis*-acting mutants from the plate assay (UV2, UV7, UV9, UV10, UV11 and S2) did not constitutively express *pgaX*, *gatA* and *gaaB*, and showed a small halo on PGA plates, indicating the halo assay can be used to discriminate between *cis*- and *trans*-acting mutants. To determine whether a *cis*-acting mutation in the *pgaX* promoter in front of the *amdS* gene was responsible for the ability of this class of mutants to grow on acetamide, the *pgaX* promoter in front of the *amdS* gene of all eight cis-acting mutants was PCR amplified using *pyrG* and *amdS* specific primers (Table S1). This analysis revealed no mutations in the *pgaX* promoter region of any of the eight presumed *cis-acting* mutants. Hence, the nature of the

mutation(s) in these strains which allow growth on acetamide remains unknown. A possibility could be the activation of expression of endogenous *amdS* genes, as at least four *amdS*-like genes are present in the genome of *A. niger*.

Identification of mutations responsible for the constitutive expression of the galacturonic acid utilization genes

A possible explanation for the constitutive expression of GA utilization genes in the mutants is that they carry mutations in the recently identified GaaR transcriptional activator (Alazi *et al*. 2016). We therefore PCR amplified and sequenced the *gaaR* locus of eight constitutive mutants obtained after UV mutagenesis (UV1, UV3, UV4, UV5, UV6, UV8, UV12 and UV13) and three spontaneous mutants (S1, S5 and S7). The *gaaR* coding regions as well as 300 bp flanking regions were sequenced, but no mutation in the *gaaR* gene in any of these eleven mutants was found (data not shown).

To determine whether the constitutive expression of GA-induced genes in these mutants involves a functioning GaaR transcription factor, we deleted the *gaaR* gene in seven of the constitutive mutants (UV1, UV3, UV4, UV5, UV6, UV8 and S1) and analysed constitutive expression using the P*pgaX-amdS* reporter. All seven mutants were unable to grow on glucose/acetamide plates (data not shown), indicating that the constitutive expression of the GA-induced genes requires a functional *gaaR* gene.

To identify mutation(s) in the gene(s) responsible for the constitutive phenotype, the genomes of five mutants (UV1, UV8, S1, S5 and S7) and the parental strain JN29.2 were sequenced. Table 2 summarizes the number of SNPs and indels detected in the five mutant strains and Table S2 lists positions and type of all SNPs and indels detected. Spontaneous mutant S7 contains only eleven SNPs or indels, of which ten are located in intergenic regions and only one SNP mutated a gene, NRRL3_08194. Remarkably, in all four other mutants a mutation was found in the same gene. Two of the mutants carry nonsense mutations (UV1 and S5) and one a frame-shift mutation (S3), all leading to premature stop codons and predicted to result in truncated proteins (Table 2). Mutants S7 and UV8 have missense mutations in the C-terminal part of the protein. These results strongly suggest that the constitutive expression of genes encoding pectin-degrading enzymes in the five mutants is caused by a loss of function of the protein encoded by NRRL3_08194.

Table 2. Mutations in the constitutive mutants as compared to the parental strain JN29.2.

Strain	Total number of SNPs and indels	SNPs or indels in coding region	Mutation in NRRL3_08194	Position of mutation relative to ATG of NRRL3_08194	Mutation in codon (bold)	Amino acid change	Predicted protein length (Full length protein is 697 amino acids)
JN29.2_UV1	40	19	G-T	1372	**G**AG-**T**AG	E to Stop	457 aa
JN29.2_UV8	21	4	G-A	1577	G**G**A-G**A**A	G to E (526)	697 aa
JN29.2_S1	68	34	Extra G	1958	GTT-G**G**T	V to G out of frame	663 aa
JN29.2_S5	48	11	C-T	1105	**C**AA-**T**AA	Q to Stop	368 aa
JN29.2_S7	11	1	T-C	2027	C**T**G-C**C**G	L-P (676)	697 aa

Deletion of NRRL3_08194 results in the constitutive expression of genes required for PGA breakdown and GA-catabolism

To determine whether the constitutive expression of the target genes of the GaaR transcriptional activator (Alazi *et al*. 2016) is caused by a loss-of-function mutation in NRRL3_08194, we deleted this gene in the *pgaX-amdS* reporter strains JC5.1 and JN29.2 (*ΔcreA*), as well in a parental background without reporter constructs (MA323.1) (Table 1). The deletion mutants were purified and deletion of NRRL3_08194 was confirmed by Southern blot analysis for JC5.1, JN29.2 and MA323.1 (Figure S1 and S2). The verified deletion mutants were tested for growth on acetamide plates containing different carbon sources (Figure 2) as well as for constitutive expression of polygalacturonases (see below). Figure 2 shows that deletion of NRRL3_08194 in the *ΔcreA* background (JN122) resulted in the ability to grow on glucose/acetamide, fructose/acetamide and sorbitol/acetamide. The colony size on all three different carbon sources was similar, indicating that the *amdS* gene was expressed regardless of the carbon source used. However, deletion of NRRL3_08194 in the JC1.5 reporter strain (JN123) resulted in similar growth on fructose/acetamide and sorbitol/acetamide plates as JN122, but reduced growth on glucose. This indicates that glucose-mediated carbon catabolite repression repressed P*pgaX* driven *amdS* expression even in the absence of NRRL3_08194. The ability of the ΔNRRL3_08194 strain to grow on acetamide plates strongly suggests that a loss of function of NRRL3_08194 results in constitutive expression of *pgaX* and other pectinolytic genes. Furthermore, deletion of *gaaX* did not result in an altered growth behavior on GA, PGA, apple pectin, glucose, fructose, sorbitol, xylose and arabinose) (data not shown). These results are most easily explained by proposing that NRRL3_08194 encodes a repressor protein, which we name GaaX, that represses the activity of the GaaR transcription factor in the absence of GA. Interestingly, GaaX shows sequence similarity to a previously identified repressor protein, QutR (Grant et al., 1988). Moreover, the transcriptional activator (GaaR, NRRL3_08195) and the repressor (GaaX, NRRL3_08194) are clustered in the genome, similar to the quinic acid utilization transcriptional activator (QutA/Qa-1F) and repressor (QutR/Qa-1S) in *Aspergillus nidulans* and *Neurospora crassa,* respectively (Geever *et al*. 1989; Levesley *et al*. 1996).

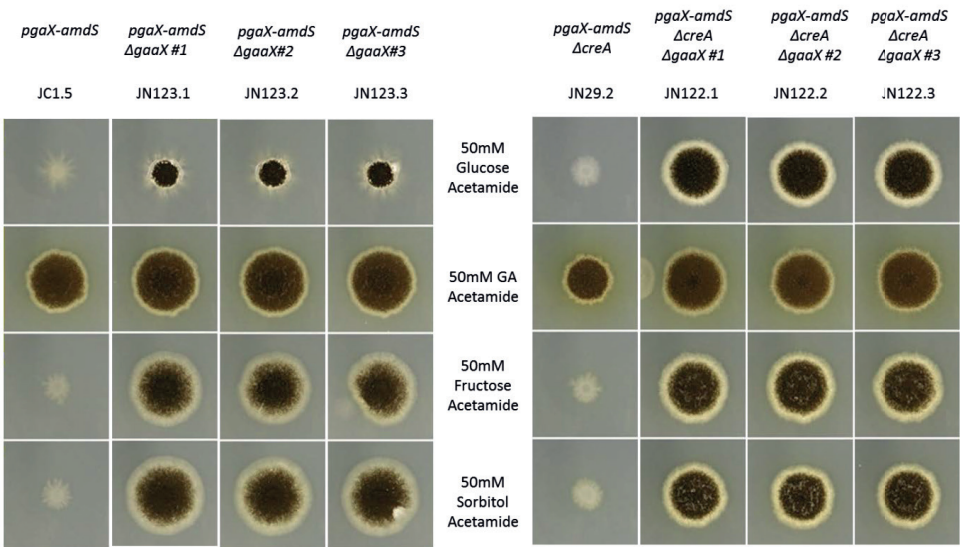


Figure 2. Regulation of the *pgaX* expression is controlled by GaaX and by CreA-mediated glucose repression. Growth of *pgaX-amdS* reporter strains in *gaaX* deletion and their parental strains was examined on various carbon sources. Parental strains and corresponding deletion strains were grown on MM-acetamide supplemented with 50 mM glucose, galacturonic acid (GA), fructose or sorbitol. All strains carry the P*pgaX-amdS* reporter construct. Strain JC1.5 is the parent of three independent transformants (JN123.1, JN123.2 and JN123.3) that contain a deletion in the *gaaX* gene. Strain JN29.2 carries the *ΔcreA* marker and is the parent of three independent transformants (JN122.1, JN122.2 and JN122.3) that contain a deletion in the *gaaX* gene.

The GaaR-GaaX target gene regulon

We posit that the regulation of GA-responsive genes is likely to be negatively controlled by the repressor protein GaaX. A possible mode of action is that the repressor GaaX inhibits the activity of the transcriptional activator GaaR in the absence of an inducer. This would imply that deletion of the repressor or activation of the transcription factor by growth on GA would result in activation of the same set of genes. To show that the loss of function of the repressor activates the GA regulon and to identify the genes repressed by GaaX under non-inducing conditions, RNA-seq profiles of the *ΔgaaX* and its parental strain (MA234.1) were compared after growth on fructose, a non-repressing carbon source. Controlled cultivations in bioreactors showed that the growth rates (μ_{max} parental strain 0.214 ± 0.007 g dry weight kg^{-1} h^{-1} (n=3); μ_{max} *ΔgaaX* 0.223 ± 0.004 g dry weight kg^{-1} h^{-1} (n=2)) as well as biomass yields (Y_{max} parental strain 4.15 ± 0.13 g dry weight kg^{-1} (n=3); Y_{max} *ΔgaaX* 4.29 ± 0.19 g dry weight kg^{-1} (n=2)) of the two strains were highly comparable, indicating the *gaaX* deletion did not result in major physiological changes affecting the growth and biomass yield.

To identify differentially expressed genes in the *ΔgaaX* strain as compared to its parental strain, RNA-seq was performed on RNA isolated from exponentially growing cells at the time point at which about 75% to 80% of the maximum biomass yield was reached. RNA-seq reads were mapped to the NRRL3/N400 genome as this is the parent of the laboratory strain N402 and derivatives used in this study. Transcript per million (TPM) values were calculated using Salmon (Patro *et al,* 2016) (Table S3). Analysis of differential gene expression, based on a stringent False Discovery Rate (FDR)-value of <0.001 and a Fold Change (FC) >4.0, identified 37 upregulated genes (Table 3). Gene Ontology (GO) enrichment analysis using FetGOat (Nitsche *et al*. 2012) and manual inspection of the genes up-regulated in the *ΔgaaX* mutant indicated that genes involved in pectin catabolism were highly enriched. Of the 37 genes, sixteen are predicted to encode extracellular enzymes acting on the GA-backbone of pectin or acting on pectin side chains (Table 3). Nine genes in the group of 37 up-regulated genes in the *ΔgaaX* strain are predicted to encode intracellular proteins. Four of these nine genes (*gaaA*-*gaaD*) are required for the conversion of GA into pyruvate and glycerol (Martens-Uzunova and Schaap 2008). The exact role of the other five genes and their possible role in GA catabolism is currently unknown. The group of 37 up-regulated genes also includes seven genes predicted to encode sugar transporter proteins. Of these seven transporter-encoding genes, only GatA has been studied in detail and shown to be able to transport GA (Sloothaak *et al*. 2014). Apart from the genes encoding extracellular enzymes (16), transporters (7) and enzymes possibly involved in GA catabolism (9), the remaining five genes in this group encode proteins with unknown functions or with similarities to known proteins that for now cannot be directly linked to GA metabolism. The deletion of *gaaX* has the most profound effect on the transcript levels of the genes encoding the first three steps of the GA-utilization pathway (*gaaA*, *gaaB* and *gaaC*) and on the expression of *gatA*. Deletion of *gaaX* resulted in a 1.24-fold (P value 0.000035) increase in *gaaR* gene activity. Since the up-regulation of *gaaR* in the *ΔgaaX* mutant is modest, it is likely that the repressing activity of GaaX is mediated at the protein level (e.g. by interacting with GaaR) rather than by transcriptional control of *gaaR*. Seventeen of the 37 genes up-regulated in the *ΔgaaX* mutant were previously identified as part of the GA regulon (Martens-Uzunova and Schaap 2008; Alazi *et al*. 2016) (Table 3, Figure 3). Sixteen of the 17 genes found in common with previous studies are predicted or demonstrated to encode extracellular pectin-degrading enzymes. These results indicate that loss of function of *gaaX* affects the expression of the GA regulon The other 20 genes were identified as significantly up-regulated in the *gaaX* mutant, but these

were not identified previously as being part of the GA regulon (Table 3, Figure 3). A re-examination of the expression of these twenty genes in the RNA-seq data published earlier (Alazi *et al*. 2016) indicated that 10 of the genes (indicated in Table 3 by the asterisk, Table S4) were also GA induced or GaaR dependent for induction in this previous study. On the other hand, 15 genes identified to be GA-induced in a GaaR-dependent manner in the previous study (Alazi *et al*. 2016) were not significantly up-regulated in the *gaaX* deletion strain (Figure 3). These results therefore suggest that full induction of GA-inducible genes requires more than the loss of GaaX activity, and that an additional induction mechanism plays a role.

An additional GA-induced gene identified in the study of Martens-Uzunova and Schaap (2008) but missing in the GaaR study is *gaaX* itself. Expression of *gaaX* was not examined in the Alazi et al. (2016) study as its function was not yet directly linked to GA utilization. However, re-evaluation of the dataset revealed that the induced expression of *gaaX* on GA is dependent on GaaR (fold change of WT vs *ΔgaaR*: 18.7; P-value 0.003; Table S4). Combining the expression data of the *ΔgaaX* mutant (this study), the *ΔgaaR* mutant (Alazi *et al*. 2016) and the genes induced on GA (Martens-Uzunova and Schaap 2008), we propose a panregulon of 53 GaaR-GaaX controlled genes and a core GaaR-GaaX regulon of at least 27 genes (Figure 3, Table 3 and Table S4). These 27 genes include eleven genes present in the intersection of all three data sets, six genes present in the intersection of the *ΔgaaX* data and the *ΔgaaR* data (Alazi *et al*. 2016), nine genes identified by examining the *gaaX* dataset with supporting evidence from previous studies and *gaaX* (Figure 3). Of these 27 genes, all except NRRL3_00660 (carboxyesterase), NRRL3_10865 (alpha-N-arabinofuranosidase), NRRL3_03342 (short-chain dehydrogenase/reductase), NRRL3_08833 (hypothetical protein), and NRRL3_02479 (beta-galactosidase*)*, have at least one predicted GARE motif in the upstream regions of the coding region (Table 3). It is interesting to note that among the genes listed in Figure 3 and Table 3 that are up-regulated in the *ΔgaaX* some of them (NRRL3_00957 and NRRL3_00958; NRRL3_09862 and NRRL3_09863; NRRL3_03291 and NRRL3_03292) are clustered. Except for NRRL3_00958, which encodes a GA-specific transporter (Sloothaak et al. 2014), the possible role of these genes in pectin degradation is currently unknown.

Table 3. Comparative RNA-seq analysis between wild-type and *ΔgaaX* strains: genes (37) upregulated in *ΔgaaX* strain.

Gene ID (NRRL3)	Gene ID (CBS513.88)	Gene name	Description	Average expression value wt[a]	Average expression value *ΔgaaX*[a]	FC *ΔgaaX* VS WT[b]	False Discovery Rate[b]	predicted localization	GARE element[c]
NRRL3_05649	An02g07720[1,2]	*gaaC*	L-threo-3-deoxy-hexylosonate aldolase	12.54	2283.77	169.39	0.00E+00	intracellular	(-) -292, -606
NRRL3_09863*	An11g03500		alpha-hydroxy acid dehydrogenase	0.82	165.84	160.07	0.00E+00	intracellular	(+)-543, (-)-182
NRRL3_00958	An14g04280[1,2]	*gatA*	MFS-type sugar/inositol transporter	3.47	524.95	140.38	0.00E+00	membrane	(+) -360
NRRL3_06890	An16g05390[1,2]	*gaaB*	L-galactonic acid dehydratase	47.77	6256.70	121.49	0.00E+00	intracellular	(+) -326
NRRL3_03291	An12g05600		heterokaryon incompatibility protein	0.00	6.78	82.98	6.48E-78	intracellular	(+)-737, -325
NRRL3_05650	An02g07710[1,2]	*gaaA*	D-galacturonic acid reductase	19.92	1515.44	71.84	0.00E+00	intracellular	(+) -414, -100
NRRL3_03144	An12g07500[1,2]	*pgaX*	exo-polygalacturonase	1.36	51.63	32.66	2.60E-272	extracellular	(+)-388
NRRL3_05252	An02g12505[2]	*pmeC*	pectin methylesterase	1.06	31.32	25.03	7.10E-189	extracellular	(+)-275, -246, -35
NRRL3_06244*	An02g00140		glycoside hydrolase family 43 protein	0.81	22.19	23.46	8.13E-193	extracellular	(+)-96, (-)-712
NRRL3_08281	An03g06740[1,2]	*pgxB*	exo-polygalacturonase Pgx28B	0.00	2.10	22.95	3.59E-43	extracellular	(-) -298, -823
NRRL3_10559	An18g04810[2]	*rgxC*	glycoside hydrolase family 28 protein	0.08	3.11	17.32	1.33E-45	extracellular	(+)-852, (-) -250
NRRL3_05260	An02g12450[1,2]	*pgxC*	exo-polygalacturonase Pgx28C	0.95	16.77	15.22	1.83E-144	extracellular	(+) -268, (-) -642
NRRL3_08663	An03g01620[1,2]		MFS-type sugar/inositol transporter	0.27	5.62	14.28	8.28E-46	membrane	(+) -673
NRRL3_03342*	An12g04990		short-chain dehydrogenase/reductase	0.81	14.46	13.51	2.10E-62	intracellular	none
NRRL3_10865	An08g01710[1,2]	*abfC*	alpha-N-arabinofuranosidase	0.80	11.21	11.92	1.19E-84	extracellular	none
NRRL3_09862	An11g03510		hypothetical protein	0.00	0.93	11.05	4.32E-18	unknown	(-) -517, -829, -843
NRRL3_10050	An11g01120[1,2]	*gaaD/larA*	NADPH-dependent erythrose reductase Err1	256.41	2732.37	10.16	0.00E+00	intracellular	(-) -538, -583, -801 -813
NRRL3_00957	An14g04260		B3/B4 domain-containing protein	1.53	18.56	10.05	2.46E-64	unknown	none
NRRL3_00502	An09g06200		hypothetical protein	0.93	12.38	9.36	1.08E-32	unknown	(-) -189
NRRL3_10558*	An18g04800		alpha-L-rhamnosidase	0.35	3.85	9.09	2.19E-61	extracellular	(+) -365
NRRL3_11710	An06g00620		MFS-type sugar/inositol transporter	2.86	25.29	8.05	1.39E-118	membrane	(+)-487, (-) -368
NRRL3_06053	An02g02540[1,2]	*paeA*	carbohydrate esterase family 16 protein	2.06	17.62	7.76	5.39E-107	extracellular	(+) -1238
NRRL3_01073	An14g05840		O-methyltransferase, COMT-type	0.54	5.57	7.16	1.06E-21	intracellular	(+) -300
NRRL3_07382	An16g00540		alpha-L-fucosidase	0.04	0.67	7.07	5.92E-14	extracellular	(-) -606

NRRL3_00660*	An14g00860		carboxylesterase	0.21	1.87	6.98	1.47E-31	extracellular	None
NRRL3_08499	An03g03960		uncharacterized protein	0.48	4.58	6.71	3.84E-21	unknown	none
NRRL3_07470	An04g09690[2]	*pmeB*	pectin methylesterase	0.75	5.72	6.36	3.59E-43	extracellular	(+) -389
NRRL3_02479	An01g10350[2]	*lacB*	exo-beta-1,4-galactanase	3.48	22.29	6.03	1.37E-179	extracellular	none
NRRL3_03292	An12g05590		carboxylesterase	0.18	1.86	5.26	1.75E-09	extracellular	(-) -1000
NRRL3_01237	An19g00270[2]	*pelD*	pectin lyase	0.17	1.47	5.25	6.14E-14	extracellular	(-) -409, -465
NRRL3_03467	An12g03550		MFS-type transporter	1.13	6.73	5.05	4.03E-28	membrane	none
NRRL3_01127*	An14g06500	DHAK	dihydroxyacetone kinase	18.90	100.64	4.98	1.61E-102	intracellular	(+) -69
NRRL3_11054*	An08g04040		MFS-type sugar/inositol transporter	6.62	31.89	4.57	2.57E-126	membrane	(+)-778, -726, (-) -646
NRRL3_08833*	n.a.		hypothetical protein	0.31	2.02	4.52	1.58E-12	unknown	none
NRRL3_01398	An13g02090*		MFS-type transporter	2.81	13.10	4.16	5.62E-30	membrane	(-) -270
NRRL3_08325	An03g06310[2]	*pmeA*	pectin methylesterase Pme8A	0.04	0.56	4.10	1.49E-07	extracellular	(+)-983, (-) -308
NRRL3_02770	An01g13880		MFS-type transporter	0.87	4.07	4.06	1.57E-20	membrane	(+) -578

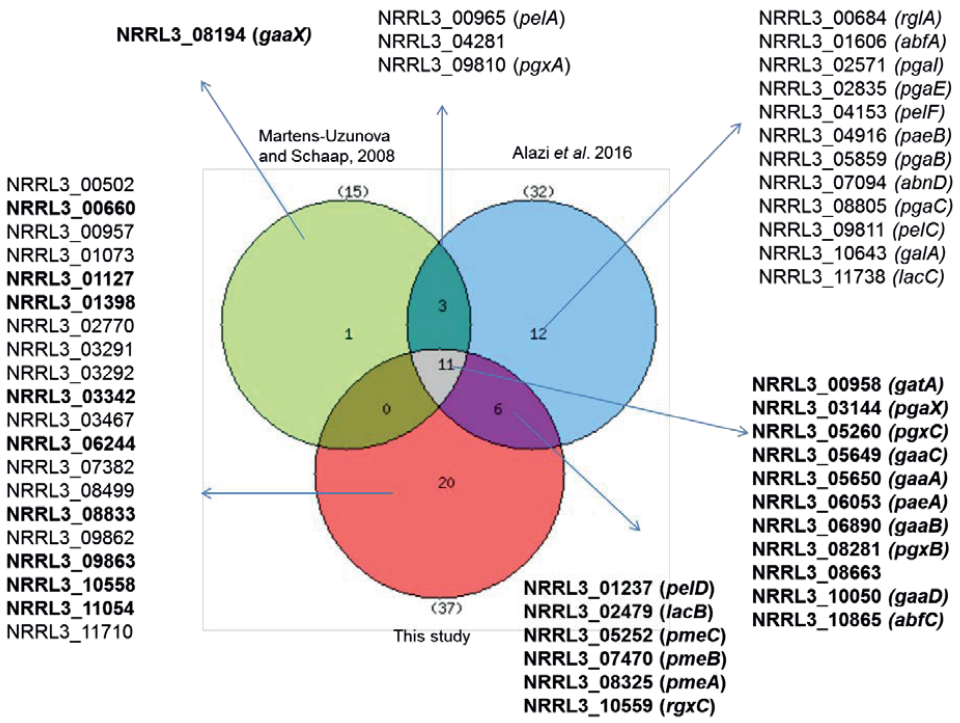


Figure 3. Venn diagram showing the overlaps between up-regulated genes in the wt_glucose versus wt_GA study (Martens-Uzunova and Schaap 2008), the up-regulated genes between *ΔgaaR*-GA versus wt-GA (Alazi *et al*. 2016), and the up-regulated genes in wt_fructose vs *ΔgaaX*-fructose (this study) to identify the GA-regulon. The 27 genes defining the GaaR-GaaX core regulon are indicated in bold.

GaaX is induced on galacturonic acid and localized in the cytosol

GaaX was previously identified as a GA-induced gene with unknown function (Martens-Uzunova and Schaap 2008). To monitor the induction of GaaX and to localize the GaaX protein in the cell, GaaX was fused to GFP at either the N- or C-terminal part of GaaX and expressed from the endogenous GaaX promoter. Fusion constructs were targeted to the *pyrG* locus of *A. niger* in a strain lacking endogenous *gaaX* (JN125.1) to be able to test complementation of the GFP-GaaX and GaaX-GFP fusion proteins (Figure S3). As shown in Figure 4A, JN125.1 (*ΔgaaX::nicB*) constitutively expressed pectinases indicated by the halo on PGA plates, while both the C-terminally tagged as well as the N-terminally tagged versions of GaaX (JN126.2 and JN127.3 respectively) complemented the constitutive expression phenotype, indicating that both fusion proteins are functional. Confocal fluorescent microscopy was performed on GFP-tagged strains to localize GaaX (Figure 4B). Spores were germinated either on GA or on fructose (a non-repressing carbon source) and a

fluorescent signal was only detectable in the GFP-labeled strains after growth on GA. This observation confirms the results from the expression data that indicate that GaaX is lowly expressed under non-inducing conditions and is induced on GA. The expression of GaaX is low on fructose and no GFP signal above the background level was detected on fructose. Based on the fluorescent pictures, GaaX is likely to be localized in the cytosol.

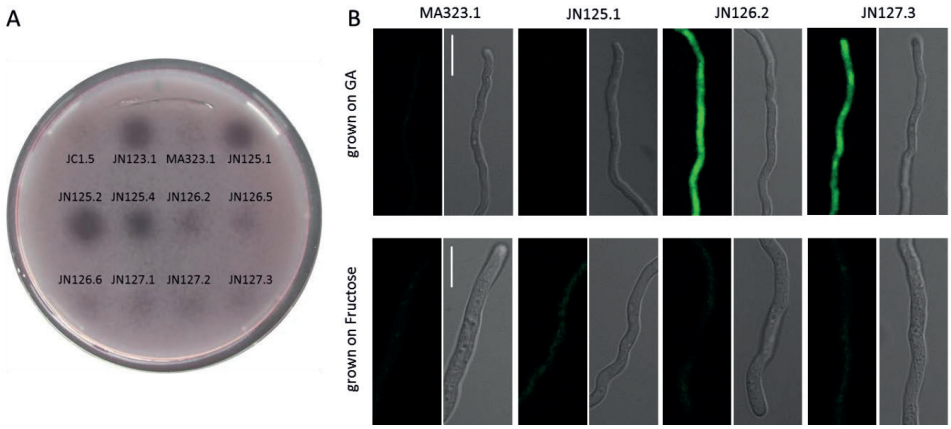


Figure 4. A) Complementation analysis of GaaX-GFP fusions. Polygalacturonase activities of *gaaX* deletion strains, *gaaX-GFP and GFP-gaaX* complementation strains, and their parental strains were detected by spotting 50 µl supernatant from fructose-grown cultures on polygalacturonic acid agarose. B) Subcellular localization of GaaX–GFP and GFP-GaaX in *A. niger* germlings. Strains were grown on coverslips in Petri dishes with minimal medium (pH 5.8) supplemented with either galacturonic acid or fructose as carbon source. Scale bar: 10µm.

Discussion

The forward genetic screen with a positive selection strategy for the isolation of *A. niger* mutants with constitutive expression of genes involved in PGA degradation resulted in the identification of a repressor protein (NRRL3_08194) which we named GaaX. Both the genome sequencing of five independently obtained mutants as well as the analysis of a targeted deletion mutant (*ΔgaaX*), showed that the loss of function of *gaaX* leads to constitutive expression of genes previously identified as GA-induced genes (Martens-Uzunova and Schaap 2008) and genes encoding pectinolytic enzymes that are activated via the transcription factor GaaR (Alazi *et al*. 2016). Deletion of *gaaX* did not result in a growth alteration on any carbon source tested (Figure 2 and data not shown). Transcriptome analysis (Tables S3) strongly suggests that deletion of *gaaX* only affects the expression of genes related to the degradation and metabolism of (poly)galacturonic acid. Genes encoding

enzymes involved in the hydrolysis of non-pectin polysaccharides are not differentially regulated in Δ*gaaX*. In addition, GO enrichment analysis of Δ*gaaX* transcriptome shows a strong correlation only between the activity of GaaX and the expression of GA-induced genes. In agreement with these observations, the phenotype of the *gaaR* deletion mutant was specific for (poly)galacturonic acid with no growth defect observed on other substrates tested (glucuronic acid, rhamnose, xylose, arabinose) (Alazi *et al.* 2016). Taken together, these findings indicate the GaaR and GaaX are specifically involved in the regulation of pectin catabolism.

Interestingly, the *gaaX* gene is located next to the recently identified GA-specific transcriptional activator *gaaR* (NRRL3_08195). The GaaR transcriptional activator is conserved in 19 out of the 20 *Aspergillus* species for which genomic sequences are available via AspGD and only absent in *A. glaucus* (Alazi *et al*. 2016), which corresponds with the inability of *A. glaucus* to grow on GA (http://www.fung-growth.org/). In all nineteen *Aspergillus* species containing GaaR, a GaaX ortholog could be identified adjacent to GaaR. Only in *A. fumigatus* (Figure 5) and *A. wentii* (data not shown) were ORFs predicted to be present in between *gaaX* and *gaaR*. The ORFs between *gaaX* and *gaaR* in *A. fumigatus* are Afu4g06430 and Afu4g06450. Afu4g06430 is predicted to encode a 128 aa long protein which has no ortholog in other aspergilli. According to available expression data (Lind et al. 2015) this gene is not expressed. Whether this predicted gene actually encodes a protein is questionable. Afu4g06450 is predicted to encode a Tan1-related transposase of the DDE family. This type of transposase is found in both *A. nidulans* and *A. niger* as well as in many other organisms. This transposase is lowly expressed in *A. fumigatus* (Lind et al. 2015).

Like *gaaR*, *gaaX* is also missing in *A. glaucus*. BLASTP and synteny analysis between *A. niger* and *A. glaucus* revealed that the GaaR/GaaX encoding genes have been excised, as surrounding genes are conserved. Despite the loss of GaaX and GaaR, *A. glaucus* still possesses the GA-specific catabolic genes *gaaA* (Aspgl1_0124049), *gaaB* (Aspgl1_0091535) and *gaaC* (Aspgl1_0065497).

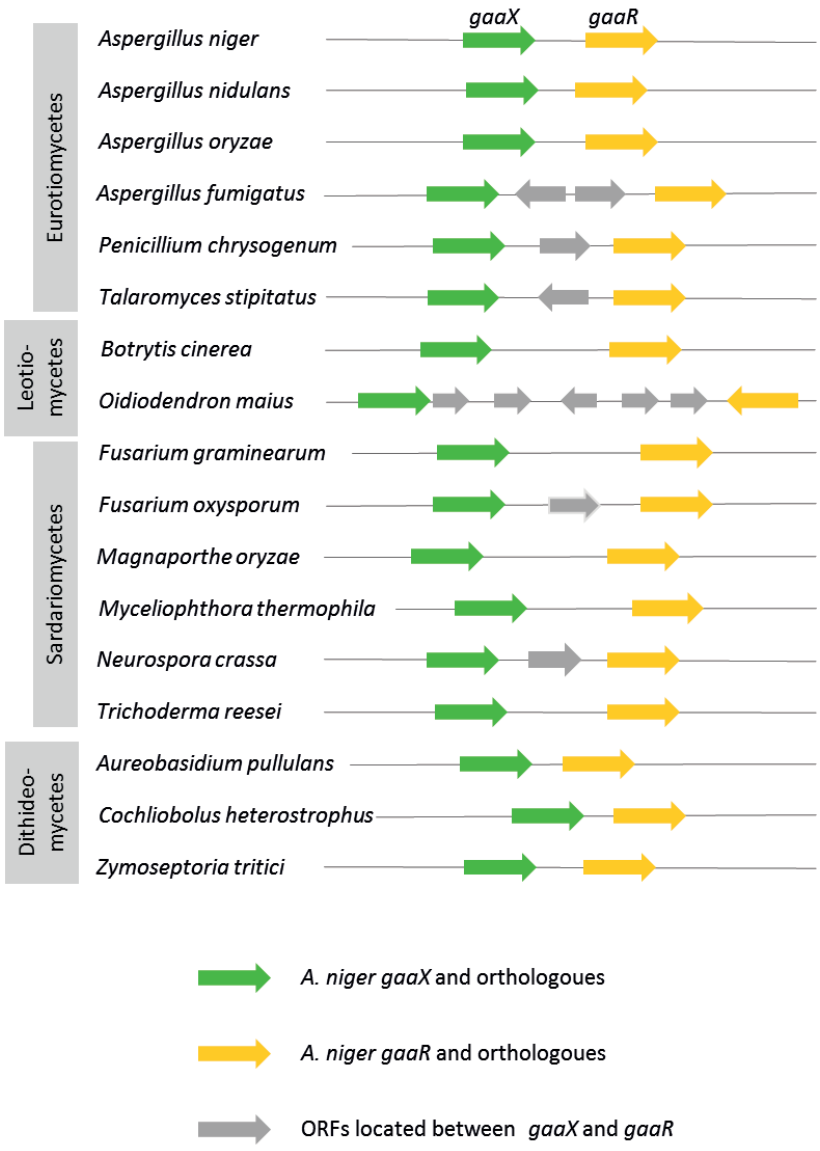


Figure 5. Schematic overview of the conservation of the *gaaX-gaaR* gene pair in 17 *Pezizomycotina* species. GaaX orthologs (green), GaaR orthologs (yellow) and ORFs between *gaaX* and *gaaR* (gray) are indicated. Arrow heads indicate the direction of transcription.

The GaaR transcriptional activator has previously been reported to be conserved in other Ascomycetes belonging to the Pezizomycotina subdivision, including members of the Eurotiomycetes (*Penicillium*, *Talaromyces* spp), Leotiomycetes (*Botrytis*, *Oidiodendron*), Sordariomycetes (*Neurospora*, *Myceliophthora*, *Magnaporthe*, *Trichoderma* and *Fusarium* spp.) and Dothideomycetes (*Zymoseptoria* (*Mycosphaerella*), *Aureobasidium* and *Cochliobolus* spp.) (Zhang *et al*. 2016). Synteny analysis of 17 species belonging to four classes of Pezizomycetes (Eurotiomycetes, Leotiomycetes, Sordariomycetes and Dothideomycetes) revealed a strong conservation of the genomic clustering of *gaaR* and *gaaX* orthologs (Figure 5 and Table S5). For most fungal species analysed, *gaaR* and *gaaX* are next to each other on the chromosome or close to each other and separated by one to five genes (Figure 5). The head to tail orientation of *gaaR-gaaX* driving expression of *gaaR* and *gaaX* from different promoters is conserved in all species except in *Oidiodendron maius*. Like GaaR, GaaX was found only in the Pezizomycotina and not in ascomycete yeasts, zygomycetes or basidiomycetes.

The strategy to identify the responsible mutation by sequencing five independently obtained mutants has been successful and efficient. Clearly, sequencing only a limited number of mutants leads only to successful identification when the mutants isolated in the screen all belong to a single complementation group. If more complementation groups are involved, more mutants would need to be sequenced. It is interesting to note that in addition to mutations in *gaaX* which were present in all five mutants, we noticed that two mutants (S1 and UV1) also contained allelic mutations in NRRL3_06175 (Table S2). The protein encoded by this gene is predicted to encode a cocaine esterase and belongs to a protein subfamily of hydrolases that included cocaine esterase (CocE), several glutaryl-7-ACA acylases, and the putative diester hydrolase NonD of *Streptomyces griseus.* This family shows extensive, low-level similarity to a family of Xaa-Pro dipeptidyl-peptidases. Whether this gene also contributes to the constitutive expression of GA-dependent genes remains to be determined, but this is unlikely as mutants without mutations in this gene display essentially the same constitutive phenotype.

Previous studies have identified genes specifically induced by GA (Martens-Uzunova and Schaap 2008) and pectinolytic genes that were dependent on the GaaR transcriptional activator for induction by GA (Alazi *et al*. 2016). Eleven of the fifteen GA-induced genes identified by Martens-Uzunova and Schaap were up-regulated in the *gaaX* mutant (Table 3 and Figure 3). The three genes that are considered GA-inducible but not detected as

differentially expressed in the *gaaX* mutant are predicted to encode a transporter (NRRL3_04281), an exo-polygalacturonase (NRRL3_09810, *pgxA*) and a pectin lyase (NRRL3_00965, *pelA*). These three genes were not classified as differentially expressed according to the stringent statistical settings in our current study. The fourth gene induced on GA in the study of Martens-Uzunova and Schaap (2008) but missing in our study is *gaaX* itself.

In our recent study on the GaaR transcriptional activator, we identified 32 pectinolytic genes whose expression on GA was dependent on GaaR (Alazi *et al*. 2016). These genes overlap largely with the previously identified GA-responsive genes (Martens-Uzunova and Schaap 2008) (Table 3, Figure 3), but also include eighteen new potential GaaR target genes. Six of these genes (including NRRL3_02479 (*lacB*), NRRL3_05252 (*pmeC*), NRRL3_08325 (*pmeA*), NRRL3_07470 (*pmeB)*, NRRL3_10559 (*rgxC*) and NRRL3_01237 (*pelD*) were also found to be significantly up-regulated in *ΔgaaX* (Table 3 and Figure 3) and are therefore considered to be part of the core GA-regulon. The remaining twelve genes identified as being GaaR dependent for induction on GA (Alazi et al. 2016) were not identified as differentially expressed based on the stringent settings in this study. Whether these genes are indeed directly controlled by GaaR and GaaX, and therefore part of the core GA regulon, awaits further study.

The GaaX protein is predicted to be 697 amino acids long and displays significant similarity to the last three domains in the C-terminal half of the AROM protein. AROM is a large (1586 amino acids in *A. niger*) pentafunctional protein composed of five domains and the individual domains are involved in five different enzymatic steps representing the prechorismate shikimate pathway, which is required for aromatic amino acid biosynthesis (Duncan *et al*. 1987; Hawkins and Smith 1991). The last three domains of the AROM protein encode the shikimate kinase (SK), 3-dehydroquinate dehydratase (DQ) and shikimate dehydrogenase (SDH) and are homologous to the respective bacterial enzymes (aroL, aroD and aroE) (Lamb *et al,* 1996). The AROM protein is present in fungi, including yeasts, and Euglena. The evolutionary origin of AROM is likely to be bacterial and it has been suggested that the AROM protein is the result of gene fusion events (Richards *et al*. 2006). Sequence alignment and BLASTP searches showed that the GaaX protein has significant sequence homology with the last three domains of the AROM protein. The observation of a transcriptional activator (GaaR) located next to a possible repressor protein (GaaX) that displays significant homology to AROM is analogous to the clustered transcriptional

activator/repressor module regulating quinic acid utilization (Geever *et al*. 1989; Lamb *et al*. 1990). Like GaaX, the quinate repressor protein shows significant sequence similarities with the last three C-terminal domains of AROM (Lamb *et al*. 1996).

The regulation of metabolic enzymes required for quinic utilization has been a classical example of gene regulation both in *N. crassa* and *A. nidulans* (Geever *et al*. 1989; Leversley *et al*. 1996). In *A. nidulans* and *N. crassa*, the transcriptional activator and repressor are located in a gene cluster which consists of the activator and repressor and other genes involved in quinic acid catabolism and transport (Geever *et al*. 1989; Lamb *et al*. 1990). *A. niger* also has a quinic acid gene cluster that includes, besides the *qutA* gene (NRRL3_11038) and *qutR* gene (NRRL3_11039), a catabolic 3-dehydroquinase (NRRL3_11037) and a MFS transporter possibly involved in quinate uptake (NRRL3_11036). In contrast to the quinic acid gene cluster in which the regulatory genes (activator and repressor) are clustered with structural genes, no structural genes involved in GA ultilization were clustered with GaaR and GaaX. Deletion of the *qutA* transcription factor (NRRL3_11038) in *A. niger* results in a quinate non-utilizing mutant (M. Arentshorst and A.F.J. Ram, unpublished results). Both in *A. nidulans* and *N. crassa*, the regulation of genes involved in quinic acid metabolism has been studied in detail and is characterized by the presence of a transcriptional activator (named QutA in *A. nidulans*, and *qa-1F* in *N. crassa*) located next to a repressor protein (QutR in *A. nidulans,* and *qa-1S* in *N. crassa*). Loss of function of quinic acid repressor *qutR* or *qa-1S* in *A. nidulans* and *N. crassa,* respectively, leads to constitutive expression of quinic acid utilization genes (Lamb *et al*. 1996; Giles *et al*. 1985), very similar to the effect observed for the loss of function of GaaX, resulting in constitutive expression of GA utilization genes. Based on the phenotype of the *gaaX* mutant and the analogy to the organization of the quinic acid utilization gene cluster, our current working hypothesis is that *gaaX* encodes a repressor protein which is required to keep the transcriptional activator GaaR in an inactive form in the absence of the inducer molecule.

As noted earlier, *gaaX* (NRRL3_08194) was identified as a up-regulated gene when an *A. niger* culture pregrown for 18 h with 2 % fructose was transferred to a medium containing 1% GA as the sole carbon source (Martens-Uzunova and Schaap 2008). The expression of a functional GFP-tagged version of GaaX confirmed the induced expression and showed cytosolic localization of GaaX in the presence of GA (Figure 4). In the promoter region of *gaaX*, a GA-responsive element (GARE) was found, suggesting that activation of the transcription factor results in increased levels of repressor protein. Although this might seem

contradictory at the first glance, it could actually be an elegant mechanism to ensure that the expression of GA-induced genes is tightly controlled and quickly responds to the presence or absence of GA. The induction of the expression of the repressor is partially analogous the activation/repression system of the qa cluster in *N. crassa*. In *N. crassa* it has been shown that both the activator (qa-1F) and the repressor (qa-1S) are transcriptionally induced in the presence of quinic acid (Patel et al. 1981; Giles et al. 1991). In the GA regulation system of *A. niger*, only the repressor protein is induced and not the activator. It should be noted that in almost all of 17 species analysed, the *gaaX* and *gaaR* genes do not share the same promoter region (head to tail orientation; Figure 5), while the *qa-1S* and *qa-1F* genes of *N. crassa* share the same promoter region, which might function as a bidirectional promoter.

As a working model (Figure 6), we postulate that in the presence of GA, the inducer molecule, which could be GA or a derivative of GA, binds in the cytosol to repressor protein GaaX. Binding of the inducer to the GaaX repressor is posited to result in the activation of the transcription factor GaaR. Active GaaR is expected to induce the expression of GA-responsive genes involved in GA release, uptake and metabolism, but also induces the expression of repressor protein. As long as the inducer is present in sufficient amounts, the GaaX repressor is predicted to be inactive as a repressor and thereby the GaaR transcription factor remains active. When the concentration of inducer decreases, it is reasonable to suggest that repressor proteins lacking bound inducer could inactivate the GaaR transcriptional activator, thereby restraining the expression of GA-responsive genes. Thus, high expression of the repressor could serve as a sensitive system to ensure that, when intracellular GA levels decrease, the cell can tightly turn off expression of GA-responsive genes. Moreover, this mechanism also ensures the rapid response to the presence of GA as it does not require *de novo* synthesis of GaaR. Induction simply requires the binding of inducer to the repressor and subsequent activation of GaaR via post-translational mechanisms, as the expression of GaaR is not dramatically induced by GA (Alazi et al., 2016) or in the *gaaX* mutant (this study). The expression of GA-induced genes is also controlled via CreA mediated carbon catabolite repression (de Vries *et al*. 2002; Niu *et al*. 2015). The analysis of the P*pgaX-amdS* reporter strain (Figure 2) suggests that the expression of *pgaX* is carbon catabolite repressed even in the *ΔgaaX* strain. This suggests that CreA directly represses *pgaX* expression via CreA binding sites in the *pgaX* promoter, independent of GaaX repression (Figure 6).

The proposed model for the mechanism by which GaaR and GaaX regulate gene expression resembles in some aspects the Gal3/Gal4/Gal80 module of *S. cerevisiae*, but

shows at least two important differences. Whereas the Gal4 regulatory system consists of three proteins (Gal4 as the transcriptional activator, Gal80 as the repressor and Gal3 as possible galactose sensor), we have identified two genes/proteins involved in GA regulation and no evidence for a third member. Also in the regulation of quinate metabolism, no third regulatory gene has been identified even though saturating mutant screens have been performed. These observations do not exclude the possibility that a third factor is involved in the GA or quinic acid regulation, but it is unlikely with the available evidence. Whereas the sensor (Gal3)/repressor (Gal80) function is mediated via two different proteins in the Gal regulatory system in *S. cerevisiae*, in the GA and quinic acid regulatory systems, the sensor/repressor function might well be performed by a single protein, GaaX and QutR, respectively. Another important difference is that GaaX and QutR do not show homology to Gal80 or Gal3, nor do Gal80 or Gal3 display homology to AROM. Based on these observations, we suggest that the GAL repressor module has evolved independently from that of GaaX/QutR.

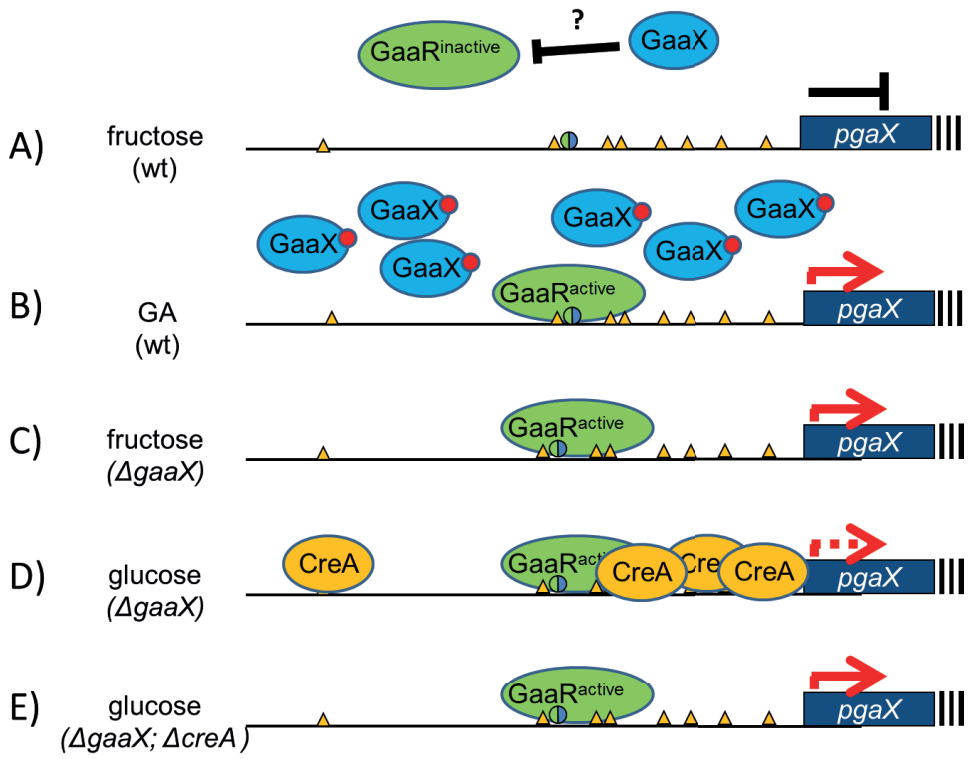


Figure 6. Model for the regulation of GA-induced gene expression in *A. niger.* A) GA-induced gene expression, with *pgaX* as an example, is controlled via interaction of the transcriptional activator (GaaR) and transcriptional repressor (GaaX) in combination with CreA-mediated carbon catabolite repression. A) In the presence of fructose (a non-repressible, non-inducing carbon source) *pgaX* expression is prevented because GaaX inhibits GaaR's activation . The question mark indicates that the mechanism by which GaaX controls GaaR activity is unknown. B) In the presence of GA, GA itself or a derivative of GA is predicted to bind to GaaX. The binding of the inducer to GaaX is expected to activate GaaR. GaaX is induced and remains cytosolic but the presence of the inducer keeps GaaX inactive. C) In the *ΔgaaX* strain, GaaR is no longer kept inactive by GaaX and therefore is constitutively active, resulting in constitutive expression of *pgaX*. D) In the *ΔgaaX* strain, the presence of glucose leads to CreA-mediated repression leading to reduced expression of *pgaX* and possibly other pectinolytic genes. E) Deletion of both *gaaX* and *creA* results in constitutive expression of *pgaX* even in the presence of glucose. The yellow triangles represent putative CreA binding sites. The green/blue circle represents a putative GaaR binding site. The red circle represents the postulated inducing sugar.

In addition to GaaX (NRRL3_08194) and QutR (NRRL3_11039), we identified two additional paralogues in the *A. niger* genome (NRRL3_08276 and NRRL3_07605). All four paralogues showed significant homology to the *A. niger* AROM protein as well as limited homology towards each other. Both NRRL3_08276 and NRRL3_07605 are also located next to predicted $Zn(II)_2Cys_6$ domain transcription factors, NRRL3_08275 and NRRL3_07604, respectively. Whereas the function of the GaaR/GaaX and QutA/QutR modules are related to GA and quinic acid metabolism, respectively, the function of the two other pairs that are present in *A. niger* remains to be elucidated. The sequence similarity of NRRL3_08276 and NRRL3_07605 to QutR and GaaX and their genome clustering with predicted transcription factors suggest that the proposed activator/repressor modules observed for GaaR-GaaX and QutA-QutR is an evolutionarily conserved mechanism to control gene expression in filamentous ascomycete fungi. The number of similar activator/repressor modules varies among Pezizomycotina species (Figure S4 and S5). Most Pezizomycotina species contain the galacturonic acid and quinic acid related transcriptional activator/repressor modules. It is interesting to note that some fungi, e.g. *Talaromyces stipitatus* and *B. cinerea,* seem to have lost the quinic acid specific repressor, which suggests they might have lost the capacity to utilize quinic acid. The GaaR/GaaX and QutA/QutR activator/repressor modules and their variants are specific for Pezizomycotina and missing in ascomycete yeasts, zycomycetes and basidiomycetes.

Acknowledgements

We thank Cees van den Hondel for helpful discussions and Tim Knetsch and Jos Reijngoud for assistance with bioreactor cultivations. Jing Niu is supported through a grant from the China Scholarship Council. Ebru Alazi is supported by a grant from BE-Basic within flagship 10 (AMBIC). This work was in part supported by Genome Canada and Génome Québec.

References

Alazi, E., J. Niu, J. E. Kowalczyk, M. Peng, M. V. Aguilar Pontes *et al.*, 2016 The transcriptional activator GaaR of *Aspergillus niger* is required for release and utilization of D-galacturonic acid from pectin. FEBS Lett 590: 1804-1815.

Arentshorst, M., A. F. Ram and V. Meyer, 2012 Using non-homologous end-joining-deficient strains for functional gene analyses in filamentous fungi. Methods Mol Biol 835**:** 133-150.

Arentshorst M., Niu J. and A. F. Ram, 2015 Efficient Generation of *Aspergillus niger* knock out Strains by combining NHEJ mutants and a split marker approach, pp. 263-272. In Genetic Transformation Systems in Fungi, Vol. 1, edited by van den Berg MA and Maruthachalam K (eds.). Springer International Publishing.

Battaglia, E., M. Zhou and R. P. de Vries, 2014 The transcriptional activators AraR and XlnR from *Aspergillus niger* regulate expression of pentose catabolic and pentose phosphate pathway genes. Res Microbiol 165**:** 531-540.

Bennett, J.W., Lasure, L., 1991. Growth media. In: Bennett J.W., Lasure L.L. (Eds.). More gene manipulations in fungi. p441-458. Academic Press, San Diego.

Bos C.J., A.J. Debets, K. Swart, A. Huybers, G. Kobus, S.M. Slakhorst, 1988 Genetic analysis and the construction of master strains for assignment of genes to six linkage groups in *Aspergillus niger*. Curr Genet 14: 437-443.

Damveld, R. A., A. Franken, M. Arentshorst, P. J. Punt, F. M. Klis *et al.*, 2008 A novel screening method for cell wall mutants in *Aspergillus niger* identifies UDP-galactopyranose mutase as an important protein in fungal cell wall biosynthesis. Genetics 178**:** 873-881.

de Vries, R. P., J. Visser and L. H. de Graaff, 1999 CreA modulates the XlnR-induced expression on xylose of *Aspergillus niger* genes involved in xylan degradation. Res Microbiol 150**:** 281-285.

de Vries, R. P., J. Jansen, G. Aguilar, L. Parenicova, V. Joosten *et al*., 2002 Expression profiling of pectinolytic genes from *Aspergillus niger*. FEBS Lett 530: 41-47.

Dowzer, C. E., and J. M. Kelly, 1991 Analysis of the *creA* gene, a regulator of carbon catabolite repression in *Aspergillus nidulans*. Mol Cell Biol 11**:** 5701-5709.

Duncan, K., R. M. Edwards and J. R. Coggins, 1987 The pentafunctional arom enzyme of *Saccharomyces cerevisiae* is a mosaic of monofunctional domains. Biochem J 246: 375-386.

Garrison, E. and G. Marth, 2012 Haplotype-based variant detection from short-read sequencing. ArXiv:1207.3907 [q-bio.GN].

Geever, R. F., L. Huiet, J. A. Baum, B. M. Tyler, V. B. Patel *et al.*, 1989 DNA sequence, organization and regulation of the qa gene cluster of *Neurospora crassa*. J Mol Biol 207**:** 15-34.

Giles, N. H., M. E. Case, J. Baum, R. Geever, L. Huiet *et al*., 1985 Gene organization and regulation in the qa (quinic acid) gene cluster of *Neurospora crassa*. Microbiol Rev 49: 338-358.

Giles N.H., R.F.Geever, D.K. Asch, J Avalos., M.E.Case 1991 The Wilhelmine E. Key 1989 invitational lecture. Organization and regulation of the qa (quinic acid) genes in *Neurospora crassa* and other fungi. J Hered 82: 1-7.

Grant, S., C.F. Roberts, H., Lamb, M. Stout, A.R. Hawkins 1988 Genetic regulation of the quinic acid utilization (QUT) gene cluster in *Aspergillus nidulans*. J Gen Microbiol 134:3 47-58.

Gruben, B. S., M. Zhou, A. Wiebenga, J. Ballering, K. M. Overkamp *et al.*, 2014 *Aspergillus niger* RhaR, a regulator involved in L-rhamnose release and catabolism. Appl Microbiol Biotechnol 98**:** 5531-5540.

Hawkins, A. R., and M. Smith, 1991 Domain structure and interaction within the pentafunctional arom polypeptide. Eur J Biochem 196: 717-724.

Jiang, H., R. Lei, S. Ding and S. Zhu, 2014 Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics 15:182.

Jørgensen, T. R., B. M. Nitsche, G. E. Lamers, M. Arentshorst, C. A. van den Hondel *et al*., 2010 Transcriptomic insights into the physiology of *Aspergillus niger* approaching a specific growth rate of zero. Appl Environ Microbiol 76: 5344-5355.

Kashyap, D. R., P. K. Vohra, S. Chopra and R. Tewari, 2001 Applications of pectinases in the commercial sector: a review. Bioresour Technol 77**:** 215-227.

Khan, M., E. Nakkeeran and S. Umesh-Kumar, 2013 Potential application of pectinase in developing functional foods. Annu Rev Food Sci Technol 4**:** 21-34.

Kopylova E., L. Noé and H. Touzet, 2012 SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28: 3211-3217.

Lamb, H. K., A. R. Hawkins, M. Smith, I. J. Harvey, J. Brown *et al.*, 1990 Spatial and biological characterisation of the complete quinic acid utilisation gene cluster in *Aspergillus nidulans*. Mol Gen Genet 223**:** 17-23.

Lamb, H. K., J. D. Moore, J. H. Lakey, L. J. Levett, K. A. Wheeler *et al.*, 1996 Comparative analysis of the QUTR transcription repressor protein and the three C-terminal domains of the pentafunctional AROM enzyme. Biochem J 313 (Pt 3)**:** 941-950.

Langmead, B. and S. Salzberg, 2012 Fast gapped-read alignment with Bowtie2. Nature Methods 9:357-359.

Levesley, I., G. H. Newton, H. K. Lamb, E. van Schothorst, R. W. Dalgleish *et al*., 1996 Domain structure and function within the QUTA protein of *Aspergillus nidulans*: implications for the control of transcription. Microbiology 142 (Pt 1): 87-98.

Love M. I., W. Huber and S. Anders, 2014 Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15:550.

Martens-Uzunova, E. S., and P. J. Schaap, 2008 An evolutionary conserved D-galacturonic acid metabolic pathway operates across filamentous fungi capable of pectin degradation. Fungal Genet Biol 45**:** 1449-1457.

Martens-Uzunova, E. S., and P. J. Schaap, 2009 Assessment of the pectin degrading enzyme network of *Aspergillus niger* by functional genomics. Fungal Genet Biol 46 Suppl 1**:** S170-S179.

Meyer V., M. Arentshorst , A. El-Ghezal A, A. C. Drews, R. Kooistra *et al*. 2007 Highly efficient gene targeting in the *Aspergillus niger kusA* mutant. J. Biotechnol 10:770-775.

Mohnen, D., 2008 Pectin structure and biosynthesis. Curr Opin Plant Biol 11**:** 266-277.

Nitsche, B. M., T. R. Jørgensen, M. Akeroyd, V. Meyer and A. F. Ram, 2012 The carbon starvation response of *Aspergillus niger* during submerged cultivation: insights from the transcriptome and secretome. BMC Genomics 13**:** 380.

Niu, J., T. G. Homan, M. Arentshorst, R. P. de Vries, J. Visser *et al.*, 2015 The interaction of induction and repression mechanisms in the regulation of galacturonic acid-induced genes in *Aspergillus niger*. Fungal Genet Biol 82**:** 32-42.

Niu, J., M. Arentshorst, F. Seelinger, A. F. Ram and J. P. Ouedraogo, 2016 A set of isogenic auxotrophic strains for constructing multiple gene deletion mutants and parasexual crossings in *Aspergillus niger*. Arch Microbiol 198**:** 861-868.

Patel V.B., M. Schweizer, C.C. Dykstra, S.R. Kushner, N.H. Giles. 1981. Genetic organization and transcriptional regulation in the qa gene cluster of *Neurospora crassa*. Proc Natl Acad Sci U S A 78: 5783-5787.

Patro, R., G. Duggal, M. I. Love, R. A. Irizarry, and C. Kingsford, 2016 Salmon provides accurate, fast, and bias-aware transcript expression estimates using dual-phase inference. bioRxiv 021592; doi:http://dx.doi.org/10.1101/021592

Petersen, K. L., J. Lehmbeck and T. Christensen, 1999 A new transcriptional activator for amylase genes in *Aspergillus*. Mol Gen Genet 262**:** 668-676.

Raulo, R., M. Kokolski and D. B. Archer, 2016 The roles of the zinc finger transcription factors XlnR, ClrA and ClrB in the breakdown of lignocellulose by *Aspergillus niger*. AMB Express 6**:** 5.

Richards, T. A., J. B. Dacks, J. M. Jenkinson, C. R. Thornton and N. J. Talbot, 2006 Evolution of filamentous plant pathogens: gene exchange across eukaryotic kingdoms. Curr Biol 16: 1857-1864.

Ruijter, G. J. G., and J. Visser, 1997 Carbon repression in *Aspergilli*. FEMS Microbiol Lett 151:103-114.

Sambrook J and Russell DW, 2001 *Molecular cloning: a laboratory manual*. Cold Spring Harbor Press, 3rd Edition.

Sloothaak, J., M. Schilders, P. J. Schaap and L. H. de Graaff, 2014 Overexpression of the *Aspergillus niger* GatA transporter leads to preferential use of D-galacturonic acid over D-xylose. AMB Express 4**:** 66.

Song, L., and L. Florea, 2015 Rcorrector: efficient and accurate error correction for Illumina RNA-seq reads. GigaScience 4:14.

Teather, R. M., and P. J. Wood, 1982 Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. Appl Environ Microbiol 43**:** 777-780.

van Hartingsveldt, W., I. E. Mattern, C. M. van Zeijl, P. H. Pouwels, and C. A. van den Hondel, 1987 Development of a homologous transformation system for *Aspergillus niger* based on the *pyrG* gene. Mol Gen Genet 206: 71-75.

van Peij, N. N., J. Visser and L. H. de Graaff, 1998 Isolation and analysis of *xlnR*, encoding a transcriptional activator co-ordinating xylanolytic expression in *Aspergillus niger*. Mol Microbiol 27**:** 131-142.

vanKuyk P.A., J.A. Benen, H.A. Wösten, J. Visser, R.P. de Vries, 2012 A broader role for AmyR in *Aspergillus niger*: regulation of the utilisation of D-glucose or D-galactose containing oligo- and polysaccharides. Appl Microbiol Biotechnol 93: 285-293.

Yuan, X. L., R. M. van der Kaaij, C. A. van den Hondel, P. J. Punt, M. J. van der Maarel *et al.*, 2008a *Aspergillus niger* genome-wide analysis reveals a large number of novel alpha-glucan acting enzymes with unexpected expression profiles. Mol Genet Genomics 279**:** 545-561.

Yuan, X. L., J. A. Roubos, C. A. van den Hondel and A. F. Ram, 2008b Identification of InuR, a new Zn(II)2Cys6 transcriptional activator involved in the regulation of inulinolytic genes in *Aspergillus niger*. Mol Genet Genomics 279**:** 11-26.

Zhang, L., R. J. Lubbers, A. Simon, J. H. Stassen, P. R. Vargas Ribera *et al.*, 2016 A novel Zn2Cys6 transcription factor BcGaaR regulates D-galacturonic acid utilization in *Botrytis cinerea*. Mol Microbiol 100**:** 247-262.

Supporting information

The supplementary material of this chapter are available via

http://www.genetics.org/content/205/1/169.figures-only and comprises the following:

Table S1. Primers used in this study.

Table S2. SNPs and indels detected in genomes of mutants.

Table S3. (A) TPM values of NRRL3 predicted ORFs (NRRL3_00001-NRRL3_11846) and (B) differentially expressed genes between WT and Δ*gaaX*.

Table S4. Expression data of genes proposed to be part of the GaaR-GaaX panregulon.

Table S5. Synteny of the GaaR-GaaX orthologs among 17 *Pezizomycotina*.

Figure S1. Verification of gaaX deletion strains in the JC1.5 and JN29.2 backgrounds.

Figure S2. Verification of the gaaX deletion strain in the MA323.1 background and targeted integration of the GFP fusion constructs at the pyrG locus.

Figure S3. Strategy and primer design for construction of GaaX-GFP fusion proteins.

Figure S4. Schematic overview of the presence or absence of GaaX-like repressor proteins in the genomes of 17 *Pezizomycotina* species, as well as S. cerevisiae, and representative Zygomycotina (*Mucor circinelloides*) and Basidiomycotina (*Ustilago maydis*).

Figure S5. Phylogenetic analysis of fungal AROM and AROM-related repressor proteins in 17 Pezizomycotina species, as well as S. cerevisiae, and representative Zygomycotina (*Mucor circinelloides*) and Basidiomycotina (*Ustilago maydis*).

Chapter 7

General discussion

General discussion

With the rapid advances of several high throughput technologies, a variety of functional genomics approaches have been developed and became available to study transcriptional regulation in *A. niger*. These approaches include: (i) those for genetic characterization of mutant genes from strains selected in forward genetic screens (e.g. via genetic linkage based methods in combination with next generation sequencing based methods), (ii) those for studying gene functions (e.g. efficient gene deletion approaches, CRISPR-Cas9 genome modification approaches), (iii) those for surveying transcriptomics (RNA-seq) and (iv) proteomics (LC-MS_MS) approaches data, and (v) those for identification of transcription factor binding sites (CHIP-seq). In the last twenty years these approaches have led to the identification of several transcription factors involved in the plant biomass degradation, both in *Neurospora crassa* (Sun *et al.* 2012; Xiong *et al.* 2017) and several *Aspergilli*. Examples of *Aspergillus* transcription factor involved in plant cell wall biomass deconstruction in include: XlnR responsible for xylan degradation and xylose utilization (Van Peij *et al.* 1998; Klaubauf *et al.* 2014); RhaR which is responsible for L- rhamnose utilization (Gruben et al., 2014); GalX/GalR which regulates the D-galactose utilization system in *A. nidulans* (Kowalczyk et al., 2015); and Clr-2/ClrB/ManR are involved in the cellulose utilization (Craig *et al.* 2015; Raulo *et al.* 2016). More information on additional transcription factors in *A. niger* as well as in other filamentous fungi can be found in a recent review by (Benocci *et al.* 2017).

As an important industrial enzyme producer, *A. niger* can secrete a variety of enzymes to degrade plant polysaccharides, such as starch, inulin, cellulose, hemicellulose (xylan and arabinan), galactomannan and pectin (De Vries and Visser 2001). The expression of the enzyme-encoding genes is subjected to tight regulatory mechanisms that involve an interplay between both inducing and repressing transcriptional factors (TFs) (Niu *et al.* 2015; Benocci *et al.* 2017). In the introductionary Chapter of my Thesis, I have focused attention to XlnR, which serves as a transcriptional activator of xylan and cellulose degrading enzymes (Van Peij *et al.* 1998), AmyR which is the transcriptional activator of starch degrading enzymes (Petersen *et al.* 1999), and CreA which acts as a carbon catabolite transcriptional repressor of enzymes involved in degradation of different carbon sources, including starch, xylan and pectin (De Vries *et al.* 2002; Tamayo *et al.* 2008; Ichinose *et al.* 2014). Despite all new and available tools and technologies, TFs related to polygalacturonic acids (PGA), the main substructure of pectin, were not identified when the research described in this thesis started.

To identify TFs involved in the regulation of pectin degradation enzymes, both targeted and non-targeted approaches were used for screening. For the targeted approach, a large set of 240 transcription factor mutants was screened for reduced growth on pectinolytic substrates such as D-galacturonic acid (GA), polygalacturonic acid (PGA), and pectins (apple pectin and sugar beet pectin). In the non-targeted approach, we designed a forward genetic screen to isolate mutants with constitutive expression of enzymes related to PGA degradation. In **Chapter 2**, we describe an approach which allows construction of gene deletion strains with high efficiency. This approach combines transformation of split marker fragments with non-homologous end joining (NHEJ) mutants. Typically, a gene deletion construct comprises a selection marker flanked by 5' and 3' sequences of the gene of interest (GOI). In the split marker approach, the gene deletion construct is split into two parts, each containing a flanking region and a truncated form of the selection marker (Fairhead *et al.* 1996; Nielsen *et al.* 2006; Goswami 2012). This approach involves only two rounds of PCR and does not require subcloning (Goswami 2012), thus allowing for rapid assembly of deletion constructs. Moreover, as only transformants in which the two overlapping marker fragments are successfully recombined can grow in selective medium, this approach increases the efficiency of obtaining gene deletion mutants.

After the gene deletion construct is introduced into the fungi by PEG-mediated protoplast transformation, targeted gene deletion is achieved by homologous recombination (HR) between the DNA sequences flanking the selection marker and target gene. However, as DNA recombination in filamentous fungi preferably happens via the NHEJ pathway, DNA will integrate randomly resulting in low HR frequencies and hence low efficiencies in getting targeted gene deletion mutants. Therefore, NHEJ mutants are used to improve the frequencies of HR. Combining the split marker approach with NHEJ mutants has proven to be able to greatly reduce time and effort for generation of gene deletion mutants. In principle, this approach is suitable to generate whole genome gene deletion collections. In our research, this method has been used in a pilot experiment to construct 240 TF mutants in *A. niger* (Mark Arentshorst, unpublished results). Whereas *A. niger* contains over 700 transcription factor genes, only a subset of TF mutants was selected for this pilot. The collection of 240 TF mutants was screened for mutants with reduced growth on pectinolytic substrate. Unfortunately this effort did not lead to successful identification of TFs involved in pectin degradation. However, we showed that the approach is in principle suitable to generate whole

genome knockout collections and has provided an interesting collection of TF mutants which can be screened to identify TF related to other regulatory processes.

As targeted deletion of the 240 transcription factors failed to identify a specific transcription factor involved in pectin degradation, one possibility is that the candidate TF is among the remaining 460 TFs present in the genome that have not been disrupted. Alternatively, since pectin is a complex polysaccharide composed of various different monosaccharides, several transcription factors with partially overlapping functions could be involved in pectin degradation and that single deletion of one TF was not sufficient to reduce growth sufficiently to be detected by plate growth assays. Thus, deletion of multiple transcription factors could be required to generate a pectin degradation deficient mutant. In **Chapter 3,** we construct a set of auxotrophic strains which allow an efficient way to create strains with multiple gene deletions. Four well–selectable auxotrophic markers (*pyrG*, *nicB*, *argB*, and *adeA*) were used to make a quadruple auxotrophic strain. Except for the *pyrG* marker, the other markers were introduced by targeted deletion. During this process, the *pyrG* marker was reused as a selection marker for deletion of other markers. Since *pyrG* marker is a bidirectional marker, mutations or loss of *pyrG* prevents the conversion of 5'-fluoroorotic acid (5'FOA) into a toxic compound, making a fungal strain resistant to 5'FOA-plates. So a $pyrG^{+}$ strain can be cured by growing on media containing 5'FOA. By introducing direct repeats around the *pyrG* selection marker, the *pyrG* gene is efficiently looped out during 5'FOA selection. In addition, two genes involved in colour pigmentation (*fwnA* and *olvA*) were combined with auxotrophic markers to construct isogenic strains which can be used to construct diploids for the isolation of haploid segregants from a diploid using the parasexual cycle.

Since identification of transcription factors involved in pectin degradation using the targeted approach by constructing gene deletion mutants in selected transcription factors was not successful, a non-targeted approach was chosen. To identify mutants affected in the regulation of pectinolytic genes, a forward genetic screen to isolate mutants with constitutive expression of pectinases was designed. In **Chapter 4**, we selected five genes that were specifically induced by galacturonic acid (GA) based on available genome-wide expression profiles, to construct promoter-reporter strains for studying gene expression related to PGA degradation. These genes consist of three exo-polygalacturonases (*pgaX*, *pgxB* and *pgxC*), a GA transporter (*gatA*), and an intracellular enzyme involved in GA metabolism (*gaaB*), each containing a putative conserved galacturonic acid-responsive element (GARE; 5'-

TCCNCCAAT-3') in their promoter regions. In **Chapter 4**, we show experimentally that the GARE is indeed required for GA-mediated induction by promoter deletion studies and site-directed mutagenesis.

To construct the promoter-reporter strains, the promoter regions of the five genes were selected and fused to the *amdS* reporter gene. Because expression of the *amdS* gene renders the strain the ability to utilize acetamide as the nitrogen source, the ability to grow on acetamide can be used as a direct measurement of promoter activity. Growth analysis of the reporter strains indicated that the promoters of four genes (*pgaX*, *pgxB*, *pgxC*, and *gatA*) were specifically induced by GA. It also showed the activation of these four promoters is under the control of carbon catabolite repression by glucose in a glucose concentration-dependent way. The major transcriptional repressor is CreA for carbon catabolite repression in *Aspergillus*. Except in the case of *pgxC*, deletion of *creA* in the other promoter-reporter strains abolished glucose repression. Deletion of *creA* in strain containing the P*pgxC_amdS* reporter was still repressed by glucose indicating an alternative repression mechanism for *pgxC*. For the other three reporter strains (P*pgaX*, P*pgxB*, and P*pgxC)* , in the *creA* mutant, GA was shown to be required as an inducer to support growth on medium containing acetamide reporter strains. Thus these promoter-reporter strains in *creA* null background can be used to perform forward genetic screens for inducer-independent mutants, which may contain a mutation responsible for constitutive activation in a GA-specific transcription factor. Such mutants can guide the search for transcription factors involved in the activation of GA-responsive genes and in the regulation of pectin-degrading enzymes in *A. niger*. The results in **Chapter 4** has led us to propose a model, in which a GA-specific transcription factor upon activation by GA or a GA-derivative could bind to the conserved motif GARE (possibly in combination with the HAP-complex) and drive GA-specific gene expression.

In Chapter 5, we identified a GA-specific transcription factor (GaaR*)* in *A. niger* by its homology to *Botrytis cinerea* GaaR (Zhang *et al.* 2016). GaaR in *B. cinerea* was identified via a yeast one hybrid (Y1H) screen using a GA-specific promoter containing the GA-motif and by expressing *B. cinerea* TF in *S. cerevisiae* (Zhang *et al*. 2016). For identification of transcription factors that can interact with a specific DNA sequence, the Y1H is widely recognized as a powerful and straightforward approach. The Y1H screening procedure relies on the interaction of a single protein (the prey) with a bait DNA sequence positioned upstream of a reporter gene. A cDNA library encoding prey proteins is needed for screening. The cDNA of the prey protein is fused to the sequence of a transcriptional activation domain.

Positive protein-DNA interactions result in recognition and binding of a transcription factor to the bait sequence, thereby activating downstream transcription of the reporter gene (Ouwerkerk and Meijer 2001; Ouwerkerk and Meijer 2011). In my study, we also used a different method and identified *gaaX* and *gaaR* as relevant regulatory factors by using a reporter strain to do forward genetic screening (**Chapter 4 and 6**).

A. niger GaaR is about 50% identical to *B. cinerea* GaaR on the amino acid level. Deletion of *gaaR* in *A. niger* showed a strongly reduced growth on GA and PGA, and a little reduced growth on sugar beet pectin (SBP) compared to parental strain. The growth phenotype indicates that GaaR is required for the utilization and release of GA from pectin. Further transcriptomic analysis of the *gaaR* deletion strain by RNA sequencing showed genes encoding 25 pectinolytic enzymes, are not induced on GA in the *ΔgaaR,* indicating that GaaR is required for the induction of these genes on GA. Other genes involved in the degradation pectin side chains and involved in catabolism of L-rhamnose, L-arabinose, and D-xylose were still expressed in *ΔgaaR*, indicating that the degradation and metabolism of pectin sugars other than GA support the growth of *ΔgaaR* on SBP.

GaaR is a member of the family of Zn(II)2Cys6 transcription factor. Both in *A. niger* and *B. cinerea* it controls expression of genes involved in GA utilization. The GaaR encoding gene in *A. niger gaaR* is 2476 bp long and contains 5 introns resulting in a 740-amino acid long protein. GaaR contains a fungal specific DNA-binding domain Zn(II)2Cys6 with the pattern of $CX_2CX_6CX_6CX_2CX_6C$ at the residues 26-56 and a fungal transcription factor regulatory middle homology region (fungal_TF_MHR) at residues 139-518. Orthologs of GaaR were found across all 20 *Aspergillus* except *Aspergillus glaucus* based on the information from the *Aspergillus* genome database (http://www.aspgd.org/). *A.glaucus* is not able to grow on medium with GA as single carbon source (http://www.fung-growth.org/) corresponding with the absence of *gaaR* in *A. glaucus*, in agreement with the fact that *gaaR* is responsible for the utilization of GA.

In **Chapter 6**, the P*pgaX-amdS* reporter strain mentioned in Chapter 4 was used to screen for inducer-independent mutants with constitutive expression of pectinases. For selection of mutants showing inducer-independent activation, we performed UV mutagenesis of the reporter strain. Mutants which can grow on a non-inducing carbon source with acetamide plates were selected. Including also spontaneous mutants obtained on selective acetamide plates, a total of 73 mutants were isolated that could grow on non-inducing conditions.

Supernatants of these mutants grown on liquid non-inducing conditions were harvested and assayed on polygalacturonic acid plates to test pectinolytic activity. Finally, 65 out of the 73 mutants showed constitutive pectinolytic activity and were considered to be trans-acting mutants. In **Chapter 5,** GaaR was shown to be the transcriptional factor which regulates the expression of genes involved in GA utilization. Considering many studies which showed that mutations in transcription factor can cause constitutive expression, we were wondering if the constitutive phenotype of these mutants could also be ascribed to mutations in *gaaR*. Therefore, the *gaaR* gene of 15 constitutive mutants was sequenced. Unfortunately, there were no mutations found in *gaaR* of these 15 constitutive mutants, meaning that mutations in other genes were responsible for the constitutive phenotype. In order to find these other genes, we selected five constitutive mutants, as well as parent strain to conduct whole genome sequencing. The sequencing results revealed allelic mutations in one particular gene encoding a previously uncharacterized protein (NRRL3_08194). The gene was named *gaaX*. Sequencing of the remaining constitutive mutants showed that all but two of the mutants contain a mutation in *gaaX*. One mutant contains a mutation in GaaR causing the constitutive expression of pectinolytic genes and one mutant did not show mutation in either GaaR or GaaX. Both mutants will be characterized and studied in more detail in future studies.

Subsequently, we also performed targeted deletion and transcriptomic analysis of the mutant strain to study the function of *gaaX*. Growth assay showed that the *ΔgaaX* mutant grow normally on a variety of C-sources including GA, PGA and pectin. Deletion of *gaaX* was also shown to result in constitutive pectinolytic activity in plates assay. RNAseq analysis revealed that 37 genes were upregulated in *ΔgaaX* mutant (FDR<0.001, FC>4.0). Gene Ontology (GO) enrichment analysis using FetGOat (Nitsche et al. 2012) and manual inspection of the genes upregulated in the *ΔgaaX* mutant indicated that genes involved in pectin catabolism were highly enriched. Of the 37 genes, 16 are predicted to encode extracellular enzymes acting on the GA backbone of pectin or acting on pectin side chains. Nine genes of the 37 upregulated genes are predicted to encode intracellular proteins. Four of these nine genes (*gaaA–gaaD*) are required for the conversion of GA into pyruvate and glycerol (Martens-Uzunova and Schaap 2008). The exact role of the other five genes and their possible role in pectin catabolism is currently unknown.

Genomic localization of *gaaX* (NRRL3_08194) showed the *gaaX* gene is next to the GA-specific transcriptional factor *gaaR* (NRRL3_08195). Among the 19 *Aspergillus* species mentioned in **Chapter 4,** *gaaR* and *gaaX* orthologs showed a strongly conserved genomic

clustering pattern: either next to each other or separated by only one or a few genes. Like *gaaR*, also *gaaX* are absent in *A. glaucus*. The *A. niger* GaaX protein is predicted to be 697 aa long. Sequence alignment and BLASTP searches displayed similarity of GaaX to the last three domains in the C-terminal half of the AROM protein, which is a large (1586 aa in *A. niger*) pentafunctional protein composed of five domains involved in different enzymatic steps of the prechorismate shikimate pathway (Duncan *et al.* 1987; Hawkins and Smith 1991). Like GaaX, the previously identified quinate repressor protein QutR is also highly similar to the last three C-terminal domains of AROM (Lamb *et al.* 1996a), Thus GaaX is homologous to the quinate repressor protein QutR (Grant *et al.* 1988). The genomic clustering of *gaaX* and *gaaR* is analogous to that of the quinic acid utilization transcriptional activator (QutA/Qa-1F) and repressor (QutR/ Qa-1S) in *A. nidulans* and *N. crassa*, respectively (Geever *et al.* 1989; Levesley *et al.* 1996). The amino acid homology to the quinate repressor, the similarity in chromosomal organization of *gaaR/gaaX* compared to the known activator/repressor genes *qutA*/*qutR*, and the constitutive phenotype of the isolated *gaaX* mutants point to the possibility that *gaaX* encodes a repressor protein that controls the activity of GaaR and keeps GaaR inactive under non-inducing conditions, similar as what was predicted for QutA/QutR (Lamb *et al.* 1996b; Levett *et al.* 2000; Watts *et al.* 2002).

Recent evidence from our group has shown that the GA-metabolic pathway intermediate 2-keto-3-deoxy-L-galactonate (2-KDG) is likely to act as the inducer (Alazi *et al.* 2017). We propose a model (Figure 1) in which it is postulated that under inducing conditions the inducer, 2-KDG, will bind to repressor protein GaaX. We speculate that the interaction of the inducer with the repressor results in the release of active GaaR from GaaX. Whether this occurs in the cytosol or nucleus remains to be determined, but preliminary experiments suggests that GaaR is present in the nucleus under inducing and non-inducing conditions (E. Alazi, personal communication) suggesting that the interaction takes place in the nucleus. Active GaaR induces not only the expression of GA-responsive genes involved in GA release, uptake and metabolism, but also induces the expression of *gaaX* (Martens-Uzunova and Schaap 2008; Niu *et al*. 2017). Under inducing conditions, a sufficient amount of the inducer is likely to keep GaaX dissociated from GaaR and consequently the GaaR transcription factor will be active. When the concentration of inducer decreases (non-inducing conditions), the amount of inducer is insufficient to bind to the repressor protein thereby allowing interaction between GaaR and the repressor GaaX to inactivate the GaaR transcriptional activator. The exact mechanism by which GaaX controls GaaR activity is unknown. However, such a

mechanism allows an elegant way which enables the cell to activate GA-induced expression when the inducer is present in the cell and allows to turn off expression of GA-responsible genes when inducer levels decrease. It also ensures the rapid response to the presence of GA as no *de novo* synthesis of GaaR is required.

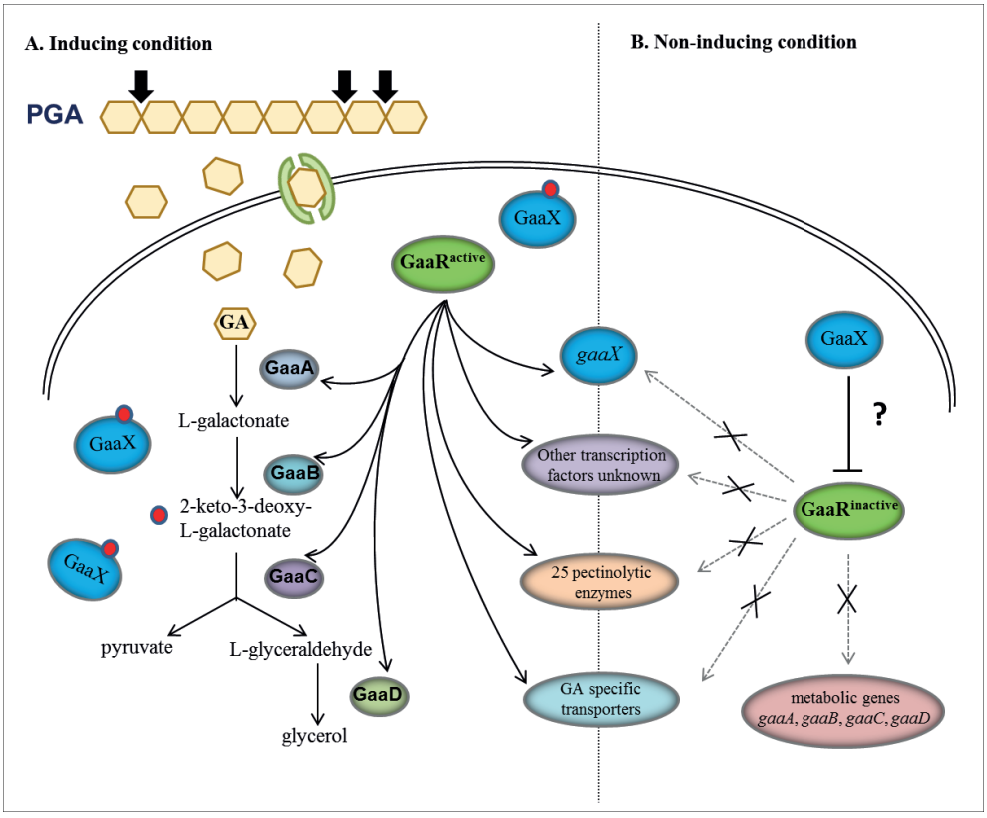


Figure 1. A) Model describing GaaR/GaaX controlled gene expression in response to galacturonic acid (GA). Under inducing conditions (presence of (poly) galacturonic acid), the inducer 2-keto-3-deoxy-L-galactonate (red dot) is predicted to bind to GaaX. The binding of the inducer to GaaX is expected to cause dissociation of the GaaX /GaaR complex resulting in non-GaaX-bound GaaR which can drive the expression of GA-responsive genes. These genes include: *gaaX* (Martens-Uzunova and Schaap 2008); other transcription factors unknown (Niu *et al.* 2017) (Supplementary Table 3, among the differentially expressed genes between WT and *ΔgaaX* , at least NRRL3_00899 and NRRL3_01451 are supposed to be fungal specific transcription factors); 25 pectinolytic genes which involved in GA release (Alazi *et al.* 2016; Niu *et al.* 2017) ; GA specific transporters which take up GA into the cell; and GA intracellular metabolism genes *gaaA*, *gaaB*, *gaaC*, *gaaD* (Alazi *et al.* 2016; Niu *et al.* 2017). Under inducing condition, *gaaX* is induced but the presence of the inducer keeps GaaX inactive. B) Under non-inducing conditions (absence of pectin or (poly)galacturonic acid), no inducer is present to bind to the repressor protein thereby allowing interaction between the repressor GaaX and the transcription activator GaaR to inactivate GaaR, preventing the expression of GA-responsible genes.

The proposed model on how GaaR/GaaX regulate gene expression is not only similar to what is shown for QutR/Qa-1S) but also shows similarities to the well-known galactose regulatory system from *S. cerevisiae*. In this system, three proteins, Gal3p/Gal4p/Gal80p, are involved in the regulation of galactose utilization. Gal4p works as a transcriptional activator, Gal80p works as a transcriptional repressor, and Gal3p possibly works as a galactose sensor. Gal80p can bind to both of Gal4p and Gal3p. In the absence of galactose, Gal80p binds to Gal4p preventing *GAL* gene expression. In the presence of galactose, galactose triggers an association between Gal3p and Gal80p, by binding of galactose to Gal3p relieving Gal4p from Gal80p (Platt *et al.* 2000; Timson *et al.* 2002; Diep *et al.* 2008; Jiang *et al.* 2009). The differences between the GA utilization system and Gal system are that in the GA case there are only two genes found to be involved in the GA utilization system so far. At this point we cannot exclude that there is a third gene involved in the GA utilization system. However, GaaX does not show homology to Gal80p or Gal3p making it unlikely that the two systems are evolutionary related. To further elucidate the regulation mechanism of the GA utilization system, future research should be aimed at understanding the biophysical and biochemical interactions between GaaR, GaaX and the inducer. Protein-protein interaction between GaaX and GaaR could be studies by co-immunoprecipitation experiments under inducing and non-inducing condition or via biomolecular fluorescence complementation (BiFC). As we have shown that GaaX can be labeled with GFP at either the N- or the C-terminus and likewise GaaR can be labeled with GFP either at the N- or C-terminus resulting in functional proteins (Chapter 6, and E. Alazi, personal communication). Constructs to perform BiFC can be designed accordingly.

Another intriguing remaining question is the binding of GaaR and/or GaaX to DNA. Although GaaX does not contain known DNA binding motifs, DNA binding can not be excluded. GaaR is expected to bind to the GARE motive, but for *A. niger* this has not been experimentally confirmed. Chip-seq experiments could be performed to identify the promoter sequences that are actually directly bound by GaaR and/or GaaX.

Finally, based on the knowledge available for the QutR/QutA regulatory system, it would be of interest to dissect the functional domains in both GaaX and GaaR involved in activation, repression, protein-protein interactions, protein-inducer interactions, and DNA-protein interactions.

References

Alazi, E., C. Khosravi, T. G. Homan, S. du Pre, M. Arentshorst *et al.*, 2017 The pathway intermediate 2-keto-3-deoxy-L-galactonate mediates the induction of genes involved in D-galacturonic acid utilization in *Aspergillus niger*. Febs Letters 591**:** 1408-1418.

Alazi, E., J. Niu, J. E. Kowalczyk, M. Peng, M. V. Aguilar Pontes *et al.*, 2016 The transcriptional activator GaaR of *Aspergillus niger* is required for release and utilization of d-galacturonic acid from pectin. FEBS Lett 590**:** 1804-1815.

Benocci, T., M. V. Aguilar-Pontes, M. Zhou, B. Seiboth and R. P. de Vries, 2017 Regulators of plant biomass degradation in ascomycetous fungi. Biotechnol Biofuels 10**:** 152.

Craig, J. P., S. T. Coradetti, T. L. Starr and N. L. Glass, 2015 Direct target network of the *Neurospora crassa* plant cell wall deconstruction regulators CLR-1, CLR-2, and XLR-1. MBio 6**:** e01452-01415.

de Vries, R. P., P. J. van de Vondervoort, L. Hendriks, M. van de Belt and J. Visser, 2002 Regulation of the alpha-glucuronidase-encoding gene (aguA) from *Aspergillus niger*. Mol Genet Genomics 268**:** 96-102.

de Vries, R. P., and J. Visser, 2001 *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. Microbiol Mol Biol Rev 65**:** 497-522, table of contents.

Diep, C. Q., X. Tao, V. Pilauri, M. Losiewicz, T. E. Blank *et al.*, 2008 Genetic evidence for sites of interaction between the Gal3 and Gal80 proteins of the *Saccharomyces cerevisiae* GAL gene switch. Genetics 178**:** 725-736.

Duncan, K., R. M. Edwards and J. R. Coggins, 1987 The pentafunctional arom enzyme of *Saccharomyces cerevisiae* is a mosaic of monofunctional domains. Biochem J 246**:** 375-386.

Fairhead, C., B. Llorente, F. Denis, M. Soler and B. Dujon, 1996 New vectors for combinatorial deletions in yeast chromosomes and for gap-repair cloning using 'split-marker' recombination. Yeast 12**:** 1439-1457.

Geever, R. F., L. Huiet, J. A. Baum, B. M. Tyler, V. B. Patel *et al.*, 1989 DNA sequence, organization and regulation of the qa gene cluster of *Neurospora crassa*. J Mol Biol 207**:** 15-34.

Goswami, R. S., 2012 Targeted gene replacement in fungi using a split-marker approach. Methods Mol Biol 835**:** 255-269.

Grant, S., C. F. Roberts, H. Lamb, M. Stout and A. R. Hawkins, 1988 Genetic regulation of the quinic acid utilization (QUT) gene cluster in *Aspergillus nidulans*. J Gen Microbiol 134**:** 347-358.

Hawkins, A. R., and M. Smith, 1991 Domain structure and interaction within the pentafunctional arom polypeptide. Eur J Biochem 196**:** 717-724.

Ichinose, S., M. Tanaka, T. Shintani and K. Gomi, 2014 Improved alpha-amylase production by *Aspergillus oryzae* after a double deletion of genes involved in carbon catabolite repression. Appl Microbiol Biotechnol 98**:** 335-343.

Jiang, F., B. R. Frey, M. L. Evans, J. C. Friel and J. E. Hopper, 2009 Gene activation by dissociation of an inhibitor from a transcriptional activation domain. Mol Cell Biol 29**:** 5604-5610.

Klaubauf, S., H. M. Narang, H. Post, M. Zhou, K. Brunner *et al.*, 2014 Similar is not the same: differences in the function of the (hemi-)cellulolytic regulator XlnR (Xlr1/Xyr1) in filamentous fungi. Fungal Genet Biol 72**:** 73-81.

Lamb, H. K., J. D. Moore, J. H. Lakey, L. J. Levett, K. A. Wheeler *et al.*, 1996a Comparative analysis of the QUTR transcription repressor protein and the three C-terminal domains of the pentafunctional AROM enzyme. Biochem J 313 (Pt 3)**:** 941-950.

Lamb, H. K., G. H. Newton, L. J. Levett, E. Cairns, C. F. Roberts *et al.*, 1996b The QUTA activator and QUTR repressor proteins of *Aspergillus nidulans* interact to regulate transcription of the quinate utilization pathway genes (vol 142, pg 1477, 1996). Microbiology-Uk 142**:** 2983-2983.

Levesley, I., G. H. Newton, H. K. Lamb, E. van Schothorst, R. W. Dalgleish *et al.*, 1996 Domain structure and function within the QUTA protein of *Aspergillus nidulans*: implications for the control of transcription. Microbiology 142 (Pt 1)**:** 87-98.

Levett, L. J., S. M. Si-Hoe, S. Liddle, K. Wheeler, D. Smith *et al.*, 2000 Identification of domains responsible for signal recognition and transduction within the QUTR transcription repressor protein. Biochemical Journal 350**:** 189-197.

Martens-Uzunova, E. S., and P. J. Schaap, 2008 An evolutionary conserved d-galacturonic acid metabolic pathway operates across filamentous fungi capable of pectin degradation. Fungal Genet Biol 45**:** 1449-1457.

Nielsen, M. L., L. Albertsen, G. Lettier, J. B. Nielsen and U. H. Mortensen, 2006 Efficient PCR-based gene targeting with a recyclable marker for *Aspergillus nidulans*. Fungal Genetics and Biology 43**:** 54-64.

Niu, J., E. Alazi, I. D. Reid, M. Arentshorst, P. J. Punt *et al.*, 2017 An Evolutionarily Conserved Transcriptional Activator-Repressor Module Controls Expression of Genes for D-Galacturonic Acid Utilization in *Aspergillus niger*. Genetics 205**:** 169-183.

Niu, J., T. G. Homan, M. Arentshorst, R. P. de Vries, J. Visser *et al.*, 2015 The interaction of induction and repression mechanisms in the regulation of galacturonic acid-induced genes in *Aspergillus niger*. Fungal Genet Biol 82**:** 32-42.

Ouwerkerk, P. B., and A. H. Meijer, 2001 Yeast one-hybrid screening for DNA-protein interactions. Curr Protoc Mol Biol Chapter 12**:** Unit 12 12.

Ouwerkerk, P. B., and A. H. Meijer, 2011 Yeast one-hybrid screens for detection of transcription factor DNA interactions. Methods Mol Biol 678**:** 211-227.

Petersen, K. L., J. Lehmbeck and T. Christensen, 1999 A new transcriptional activator for amylase genes in *Aspergillus*. Mol Gen Genet 262**:** 668-676.

Platt, A., H. C. Ross, S. Hankin and R. J. Reece, 2000 The insertion of two amino acids into a transcriptional inducer converts it into a galactokinase. Proceedings of the National Academy of Sciences of the United States of America 97**:** 3154-3159.

Raulo, R., M. Kokolski and D. B. Archer, 2016 The roles of the zinc finger transcription factors XlnR, ClrA and ClrB in the breakdown of lignocellulose by *Aspergillus niger*. AMB Express 6**:** 5.

Sun, J., C. Tian, S. Diamond and N. L. Glass, 2012 Deciphering transcriptional regulatory mechanisms associated with hemicellulose degradation in *Neurospora crassa*. Eukaryot Cell 11**:** 482-493.

Tamayo, E. N., A. Villanueva, A. A. Hasper, L. H. de Graaff, D. Ramon *et al.*, 2008 CreA mediates repression of the regulatory gene xlnR which controls the production of xylanolytic enzymes in *Aspergillus nidulans*. Fungal Genet Biol 45**:** 984-993.

Timson, D. J., H. C. Ross and R. J. Reece, 2002 Gal3p and Gal1p interact with the transcriptional repressor Gal80p to form a complex of 1 : 1 stoichiometry. Biochemical Journal 363**:** 515-520.

van Peij, N. N., J. Visser and L. H. de Graaff, 1998 Isolation and analysis of xlnR, encoding a transcriptional activator co-ordinating xylanolytic expression in *Aspergillus niger*. Mol Microbiol 27**:** 131-142.

Watts, C., S. M. Si-Hoe, H. K. Lamb, L. J. Levett, J. R. Coggins *et al.*, 2002 Kinetic analysis of the interaction between the QutA and QutR transcription-regulating proteins. Proteins-Structure Function and Genetics 48**:** 161-168.

Xiong, Y., V. W. Wu, A. Lubbe, L. Qin, S. Deng *et al.*, 2017 A fungal transcription factor essential for starch degradation affects integration of carbon and nitrogen metabolism. PLoS Genet 13**:** e1006737.

Zhang, L., R. J. Lubbers, A. Simon, J. H. Stassen, P. R. Vargas Ribera *et al.*, 2016 A novel Zn2 Cys6 transcription factor BcGaaR regulates D-galacturonic acid utilization in Botrytis cinerea. Mol Microbiol 100**:** 247-262.

Summary

&

Samenvatting

Summary

Aspergillus niger is a saprotrophic filamentous fungus feeding on dead plant material in Nature. To be able to utilize the carbon from plant material, the first determined step is to break down the plant polysaccharides, which apart from storage carbohydrates like starch and inulin, are mainly present in the cell wall of plants. Plant cell walls consist of complex polysaccharides, such as cellulose, hemicellulose and pectin. *A. niger* can produce hydrolytic enzymes, such as cellulases, beta-glucosidases, and cellobiohydrolases to degrade cellulose. Similarly, xylanases, xylosidases, and arabinofuranosidases are produced to degrade hemicellulose, and pectin degrading enzymes, such as exo-polygalacturonases, pectin lyase and pectin esterases are produced to breakdown pectin. After degradation of these polymers into monosaccharides, these monosaccharides can then be taken up by the cell for intracellular utilization. Because of the attractive applications of the various hydrolytic enzymes, *A. niger* is frequently used as host for the industrial production of these enzymes. These enzymes find their application in various fields of the food, medical and textile industries. Based on its long history of safe use, *A. niger* enzymes have also acquired the socalled GRAS (generally regarded as safe) status. For these reasons high level production of enzymes in *A. niger* has obtained considerable scientific attention. Scientists have made much progress in improving enzyme production using a wide variety of approaches: such as screening for hyperproductive mutants, optimizing of production medium and optimizing culture conditions. From research related to the transcriptional regulation of plant cell wall degrading enzymes in the last 20 years it has become clear that the expression of the genes encoding these enzymes are tightly regulated in filamentous fungi, including *A. niger*. These genes are under the control of substrate specific transcriptional activators and only highly express in response to the presence of a particular substrate. The starch regulator AmyR and the xylan regulator XlnR, were the first regulators identified using classical techniques. With the rapid development of modern functional genomics approaches, several new transcription factors involved in plant biomass degradation were identified in Aspergilli in the last few years. These include InuR (inulin utilization), Clr-2/ClrB/ManR (cellulose and mannan) utilization; GalX/R (galactose utilization); AraR (arabinose utilization), RhaR (rhamnose utilization). However, at the beginning of the research described in this thesis, the transcription factors involved in pectin degradation were not identified.

Therefore, in this thesis I have focused on the identification of transcription factors responsible for pectin degradation and on subsequent studies regarding its specific regulatory

mechanisms. The main constituent of pectin is the monosaccharide D-galacturonic acid (GA), and GA is the predominant product released from pectin degradation. Elucidation of the mechanism underlying the regulation of genes encoding pectin degrading enzymes is expected to contribute to improve industrial production of pectinases.

In **Chapter 1,** I give an introduction of filamentous fungi with a focus on *A. niger*, as well as their industrial application. It describes the impact of advances in next-generation sequencing technologies for fungal genetics and genomics. A variety of functional genomics approaches have been developed with the rapid advances of several high throughput technologies. These approaches include (i) genetic linkage based methods in combination with next generation sequencing to characterize mutant genes from strains selected in forward genetic screens; (ii) efficient gene deletion approaches like split marker approaches combined with NHEJ mutants and CRISPR-Cas9-based genome modification approaches and overexpression method for studying gene function; (iii) Microarray and RNA-seq for surveying transcriptomics data; and (iv) CHIP-seq for identification of transcription factor binding sites. All these methods together provide useful perspectives to understand gene regulation. This chapter also describes the major polysaccharides starch, xylan and polygalacturonic acid of plant biomass and their substrate specific transcription factors and corresponding regulation mechanisms for the expression of genes involved in the degradation of these polysaccharides. **Chapter 2** describes a split marker approach in combination with NHEJ mutants for efficient targeted gene deletion. In **Chapter 3,** we constructed a set of isogenic auxotrophic strains by recycling *pyrG* marker. These auxotrophic strains can be used for constructing multiple gene deletion mutants to study more complex multi-gene families. In **Chapter 4,** we selected five galacturonic acid (GA) induced genes based on available genome-wide expression profiles to analyze the regulation of promoter activity of these genes *in vivo* by constructing promoter_*amdS* reporter strains. Using these strains, we show that induction and repression of GA-induced genes is differentially fine-tuned in response to inducing and repressing conditions. In **Chapter 5**, we identified the GA responsive transcription activator GaaR of *A. niger* by homology to BcGaaR of *Botrytis cinerea*. Targeted deletion of *gaaR* and transcriptomic profiling studies of *ΔgaaR* showed that the *A. niger* GaaR ortholog is required for releasing and utilization of GA from pectin. In **Chapter 6**, we describe the use of one of the reporter strains containing P*pgax_amdS* construct to screen inducer independent mutants that constitutively express pectinolytic enzymes. In **Chapter 6** full genome sequencing of five mutants out of totally 65 mutants showing constitutive pectinolytic activity revealed allelic mutations in one particular gene, we subsequently named

gaaX. Targeted deletion of *gaaX* in combination with RNA-seq analysis of *ΔgaaX* showed that deletion of *gaaX* or mutations in *gaaX* cause constitutive expression of a large number of genes involved in releasing and utilization of GA. Chromosomal localization shows that *gaaX* is syntenous to *gaaR*, which encodes the transcription activator for GA utilization identified in **Chapter 5**. In **Chapter 6**, evidence is provided that GaaX is likely to act as a transcription repressor which inhibits GaaR activity under non-inducing conditions. The newly discovered GA-specific activator/repressor module involved in pectin degradation represents a new and unexpected, yet conserved mechanism for controlling transcription in filamentous fungi. In the last Chapter (**Chapter 7**) we summarize and discuss the major conclusions of this thesis and propose some future directions to study the proposed interactions between GaaX, GaaR, and the inducer to decipher further the regulatory mechanism of pectin degradation system in *A. niger*.

Samenvatting

Aspergillus niger is een saprotrofe filamenteuze schimmel die zich van nature voedt met dood plantenmateriaal. Om de koolhydraten van het plantenmateriaal, welke, naast opslag-koolhydraten zoals zetmeel en inuline, vooral bestaan uit celwand koolhydraten, te kunnen gebruiken moeten deze polysacchariden allereerst worden afgebroken tot monomeren. De celwanden bestaan uit complexe polysacchariden zoals cellulose, hemicellulose en pectine. *A. niger* kan hydrolytische enzymen zoals cellulases, beta-glucosidases en cellobiohydrolases produceren voor de afbraak van cellulose. Op dezelfde manier worden xylanases, xylosidases en arabinofuranosidases geproduceerd om hemicellulose af te breken, en pectine afbrekende enzymen zoals exo-polygalacturonases, pectine lyase en pectine esterases worden geproduceerd als *A. niger* op pectine wordt gekweekt. Wanneer deze polysacchariden zijn afgebroken tot monomeren kunnen ze door de cel worden opgenomen en intracellulair gebruikt. Vanwege de aantrekkelijke toepassingen van de verschillende hydrolytische enzymen wordt *A. niger* vaak gebruikt als gastheer voor de industriële productie van deze enzymen. De toepassingen van deze enzymen zijn te vinden in de medische, voedings- en textielindustrie. Gebaseerd op zijn lange geschiedenis van veilig gebruik heeft *A. niger* de GRAS (Generally Regarded As Safe) status verkregen.

De bovengenoemde aspecten hebben gezorgd voor veel wetenschappelijke aandacht voor het bereiken van hoge productieniveaus van enzymen in *A. niger*. Wetenschappers hebben veel vooruitgang geboekt met het verbeteren van enzymproductie, waarbij een breed scala aan methoden zijn gebruikt: het screenen van hyperproductieve mutanten, optimalisatie van productiemedium en van kweekcondities. Uit onderzoek van de laatste twintig jaar naar de transcriptionele regulatie van plantencelwand afbrekende enzymen is gebleken dat de expressie van de genen die coderen voor deze enzymen strikt wordt gereguleerd in filamenteuze schimmels, zoals *A. niger*. De genen staan onder controle van substraat specifieke transcriptionele activatoren en vertonen alleen hoge expressie in aanwezigheid van een specifiek substraat. De zetmeel regulator AmyR en de xylaan regulator XlnR zijn de eerste regulatoren die zijn geïdentificeerd met behulp van klassieke methoden. Door de snelle ontwikkeling van moderne “functional genomics” methoden zijn de laatste jaren verschillende nieuwe, bij de afbraak van biomassa van planten betrokken, transcriptiefactoren geïdentificeerd in *Aspergillus* soorten. Dit zijn onder andere InuR (inuline afbraak), Clr-2/ClrB/ManR (cellulose en mannan), GalX/R (galactan), AraR (arabinan) en RhaR

(rhamnan). Echter, aan het begin van het onderzoek beschreven in dit proefschrift waren de transcriptiefactoren die betrokken zijn bij de afbraak van pectine nog niet geïdentificeerd.

In dit proefschrift heb ik me derhalve gericht op de identificatie van transcriptiefactoren die verantwoordelijk zijn voor de afbraak van pectine om vervolgens de specifieke regulerende mechanismen hiervan te bestuderen. Het hoofdbestanddeel van pectine is het suikerzuur D-galacturonzuur (GA), en na afbraak van pectine is GA dan ook het voornaamste afbraakproduct. De verwachting is dat het ophelderen van het mechanisme van regulatie van genen coderend voor pectine afbrekende enzymen zal bijdragen aan het verbeteren van de industriële productie van pectinases.

In **Hoofdstuk 1** geef ik een introductie over filamenteuze schimmels, met nadruk op de regulatie van enzym productie in *A. niger*, en hun industriële toepassingen. Het hoofdstuk beschrijft het effect van de ontwikkelingen in next-generation sequencing technieken op schimmel-genetica en genomics. Een verscheidenheid aan functional genomics methoden is ontwikkeld door de snelle vooruitgang van verschillende high throughput technologieën. Dit zijn o.a. (i) methoden gebaseerd op klassieke genetica gecombineerd met next-generation sequencing om mutante genen te karakteriseren van stammen die geselecteerd zijn in een zogn. forward genetic screen; (ii) efficiënte gen-deletie methoden, zoals de Split Marker methode gecombineerd met NHEJ mutanten en CRISPR-Cas9 gebaseerde genoommodificaties, en over expressie methoden om daarmee de functie van genen te bestuderen; (iii) Micro-array en RNA-seq methoden om transcriptome analyse uit tevoren en (iv) CHIP-seq methodes voor het identificeren van de bindingsplaats van transcriptiefactoren op het DNA. Deze methoden samen zorgen voor een bruikbare aanpak om genregulatie te begrijpen. Dit hoofdstuk beschrijft ook de voornaamste polysacchariden van biomassa van planten, nl. zetmeel, xylan, en polygalacturonzuur, en hun substraat specifieke transcriptiefactoren en de corresponderende regulatie mechanismen voor de expressie van genen die betrokken zijn bij de afbraak van deze polysacchariden. **Hoofstuk 2** beschrijft een Split Marker methode gecombineerd met NHEJ mutanten, voor het efficiënt maken van gerichte gen-deleties. In **Hoofdstuk 3** hebben we een set van isogene, auxotrofe stammen gemaakt door middel van het gebruik van de recycleerbare *pyrG* marker. Deze auxotrofe stammen kunnen gebruikt worden voor het maken van meervoudige gen-deleties voor het bestuderen van meer complexe multi-gen families. In **Hoofdstuk 4** worden, gebaseerd op beschikbare genoom-brede gen-expressieprofielen, vijf genen geselecteerd die worden geïnduceerd op medium met galacturonzuur (GA), om de regulatie van de promoter activiteit

van deze genen *in vivo* te bestuderen d.m.v. het maken van stammen met de promoter_*amdS* reporter. Met behulp van deze stammen tonen we aan dat de inductie en repressie van GA-geïnduceerde genen op verschillende wijze wordt verfijnd in reactie op inducerende en represserende condities. In **Hoofdstuk 5** hebben we de GA-reactieve transcriptionele activator GaaR geïdentificeerd d.m.v. homologie met de GA-transcriptie-activator BcGaaR uit *Botrytis cinerea*. Gerichte deletie van *gaaR* en transcriptomische profiling van een gaaR deletie stam (*ΔgaaR*) heeft aangetoond dat de GaaR ortholoog van *A. niger* noodzakelijk is voor het vrijkomen van GA vanuit pectine en het gebruik ervan. In **Hoofdstuk 6** beschrijven we het gebruik van de reporter stam die het *PpgaX-amdS* construct bevat, om te screenen voor inductie onafhankelijke mutanten die pectinolytische enzymen constitutief tot expressie brengen. Sequentie-analyse van het volledige genoom van vijf van totaal 65 van deze mutanten, onthulde mutante allelen van één enkel gen, dat we *gaaX* hebben genoemd. Gerichte deletie van *gaaX*, gecombineerd met RNA-seq analyse van een *gaaX* deletie stam, toonde aan dat deletie van *gaaX (*of mutaties in *gaaX*) constitutieve expressie veroorzaakt van een groot aantal genen die betrokken zijn bij het vrijmaken van GA uit pectine en het gebruiken van GA. Chromosomale localisatie toont aan dat *gaaX* gelegen is direct naast *gaaR*, dat codeert voor de transcriptionele activator voor het gebruik van GA, zoals beschreven in **Hoofdstuk 5**. In **Hoofdstuk 6** wordt experimenteel aannemelijk gemaakt dat GaaX werkt als een transcriptonele repressor, die GaaR activiteit onderdrukt tijdens niet-inducerende condities. Deze nieuw ontdekte GA-specifieke activator/repressor module, betrokken bij afbraak van pectine, vertegenwoordigt een nieuw en onverwacht, maar duidelijk geconserveerd, mechanisme om transcriptie te controleren in filamenteuze schimmels. In het laatste Hoofdstuk (**Hoofdstuk 7**) vatten we de belangrijkste conclusies van dit proefschrift samen en bediscussiëren deze. Ook stellen we enkele toekomstige onderzoeksrichtingen voor om de voorgestelde interacties tussen GaaX, GaaR en de inducer te bestuderen om zo de regulatie mechanismen van het pectine afbrekende systeem in *A. niger* verder te ontcijferen.

Curriculum Vitae

Jing Niu was born in Qujing, Yunan Province, China during the hot summer of 1986, on the 20th of June. In 2002, she attended senior middle school at Qujing NO. 1 Middle School. From 2005 to 2009 she studied Biological Sciences at the Minzu University of China. After graduation with a Bachelor degree, she went to the Chinese Academy of Agricultural Sciences (CAAS) to study Microbiology with Prof. Dr. Xiliang Jiang. In 2012, she graduated with a Master degree and was awarded a scholarship by the Chinese Scholarship Council (CSC) to study for her PhD at Leiden University, where she joined the Molecular Microbiology and Biotechnology research group and started her PhD under the supervision of Dr. A.F.J. Ram and Prof. Dr. P.J. Punt. The topic of her project was "Transcriptional control of pectin degrading enzymes in *Aspergillus niger*", and the results are described in this thesis.

Publications

1. **Niu, J.**, Alazi, E., Reid, I.D., Arentshorst, M., Punt P.J., Visser, J., Tsang, A., and Ram, A.F.J., 2017 An evolutionarily conserved transcriptional activator-repressor module controls expression of genes for D-galacturonic acid utilization in *Aspergillus niger*. Genetics 205: 169-183.

2. Alazi, E.*, **Niu, J.***, Kowalczyk, J.E.*, Peng, M., Pontes, M.V.A., van Kan, J.A.L., Visser, J., de Vries, R.P., and Ram, A.F.J., 2016 The transcriptional activator GaaR of *Aspergillus niger* is required for release and utilization of D-galacturonic acid from pectin. FEBS Letters 590: 1804-1815. (*co-first author)

3. **Niu, J.**, Arentshorst, M., Seelinger, F., Ram, A.F.J, and Ouedraogo, J.P., 2016 A set of isogenic auxotrophic strains for constructing multiple gene deletion mutants and parasexual crossings in *Aspergillus niger*. Arch Microbiol 198: 861-868.

4. **Niu, J.**, Arentshorst, M., Nair, P.D.S., Dai, Z.Y., Baker, S.E., Frisvad, J.C., Nielsen, K.F., Punt, P.J., and Ram, A.F.J., 2016 Identification of a classical mutant in the industrial host *Aspergillus niger* by systems genetics: LaeA is required for citric acid production and regulates the formation of some secondary metabolites. G3-Genes Genomes Genetics 6: 193-204.

5. **Niu, J.**, Homan, T.G., Arentshorst, M., de Vries, R.P., Visser, J., and Ram, A.F.J., 2015 The interaction of induction and repression mechanisms in the regulation of galacturonic acid-induced genes in *Aspergillus niger*. Fungal Genetics and Biology 82: 32-42.

6. Arentshorst, M., **Niu, J.**, and Ram, A.F.J., 2015 Efficient generation of *Aspergillus niger* knock out strains by combining NHEJ mutants and a split marker approach, pp. 263-272 in Genetic Transformation Systems in Fungi, Volume 1, edited by M. A. van den Berg and K. Maruthachalam. Springer International Publishing, Cham.