Transcriptional control of pectin degrading enzymes in Aspergillus niger

Niu, J.; Niu J.

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*

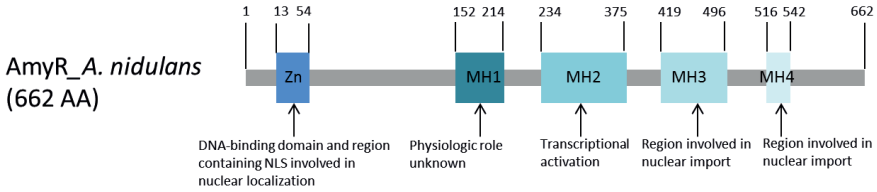

Figure1. Schematic diagram of the AmyR protein domains in *A. nidulans*. This figure was adapted from (Suzuki *et al.* 2015)

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).

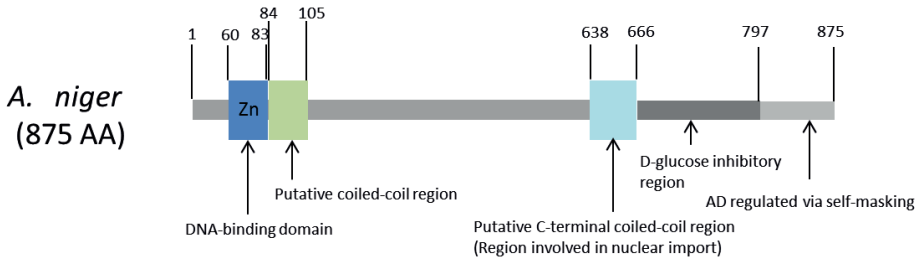

Figure 2. Schematic diagram of the XlnR protein domains in *A. niger*. This figure was adapted from Hasper et al., 2004. Zn domain was analyzed manually. The two coiled-coil regions were predicted on line by website: http://www.bioinformatics.nl/cgi-bin/emboss/pepcoil.

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).

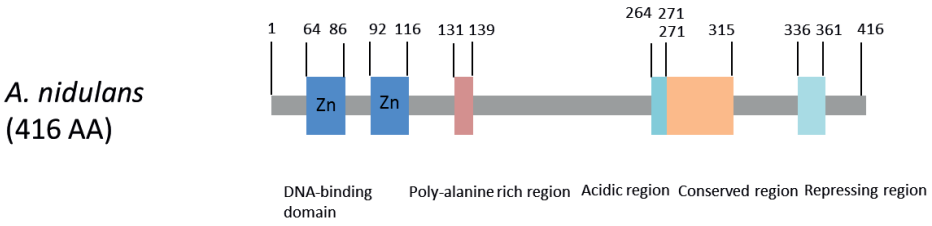

Figure 3. Schematic diagram of different regions in CreA (Roy *et al.* 2008).

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.