Discovery and exploitation of the transcriptional regulatory system of pectinases in Aspergillus niger

Alazi, E.

Discovery and exploitation of the transcriptional regulatory system of pectinases in *Aspergillus niger*

Ebru Alazi

Cover & layout design: Ebru Alazi

Printed by: Ridderprint BV, Ridderkerk, The Netherlands

ISBN: 978-94-6299-971-8

Funding statement: This research was funded by BE-Basic (Flagship 10).

Discovery and exploitation of the transcriptional regulatory system of pectinases in *Aspergillus niger*

Proefschrift

ter verkrijging van

de graad van Doctor aan de Universiteit Leiden,

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

volgens besluit van het College voor Promoties

te verdedigen op woensdag 6 juni 2018

klokke 16.15 uur

door

Ebru Alazi

geboren te Üsküdar, Turkije

in 1985

Promoter:	Prof. dr. P.J. Punt
Co-promoter:	Dr. A.F.J. Ram
Promotiecommissie:	Prof.dr. H.P. Spaink (Leiden University)
	Prof.dr. G.P. van Wezel (Leiden University)
	Prof.dr. H.A.B. Wösten (Utrecht University)
	Prof.dr. J.H. de Winde (Leiden University)
	Prof.dr. N.L. Glass (The University of California)
	Prof.dr. R.P. de Vries (Utrecht University)

Contents

CHAPTER 1

Aim and outline of the thesis

Besides animals and plants, fungi represent one of the eukaryotic kingdoms of life. The fungal kingdom is thought to consist of over 1.5 million species and harbors a very diverse group of microorganisms varying from single celled yeasts to multicellular filamentous fungi (molds) and mushroom fungi (Stajich et al., 2009). Example species to each of these groups include Baker's Yeast (*Saccharomyces cerevisiae*), *Aspergillus* and *Penicillium* species, and edible mushrooms like the *Agaricus bisporus*, respectively. *Aspergillus niger* is a filamentous fungus that forms colonies on solid substrates. Depletion of the carbon source causes differentiation of the vegetative mycelium and induces the formation of spore-containing structures called conidiophores. The asexually derived spores, called conidia, can germinate and extend to form hyphae (Figure 1). In liquid medium, *A. niger* shows either dispersed or pelleted growth (Krijgsheld et al., 2013). In our daily life, we often encounter the black conidia of *A. niger* on the outside skin of onions (Gherbawy et al., 2015).

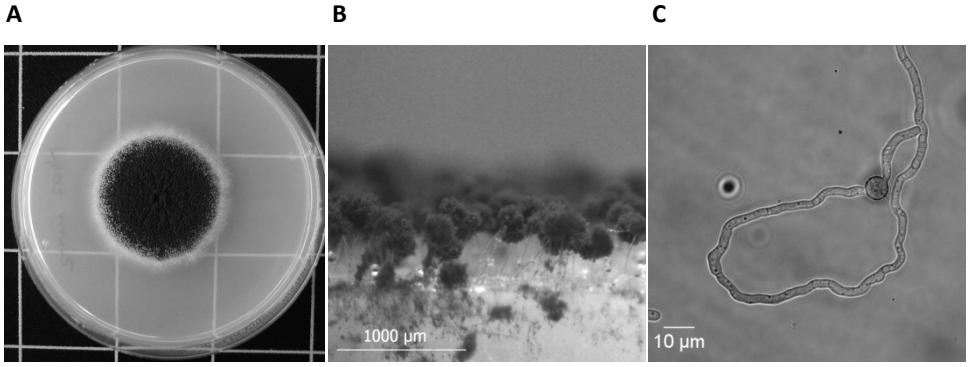

Figure 1. Pictures of *A. niger* strain N402 (or its derivatives). **(A)** Top view of a sporulating colony formed on solid agar plate. **(B)** Side view of conidiophores formed on solid agar plate. Chains of conidia are seen on top of the stalk. **(C)** Hyphae emerging from a single conidium in liquid medium.

A. niger is a saprobe, growing on decaying plant matter in nature. It secretes enzymes to degrade the large plant polysaccharides, and takes in and catabolizes the resulting di- or monomeric sugars to generate energy for life. These di- or monomeric sugars (or derivatives thereof) act as inducer molecules resulting in the activated and abundant expression the genes encoding plant polysaccharide degrading enzymes via substrate-specific transcription factors (TFs) (Culleton et al., 2013). These genes are often transcriptionally down-regulated via carbon catabolite repression (CCR) when a more energetically favorable carbon source, such as glucose, compared to plant polysaccharides is available as substrate (Adnan et al., 2018).

Being a generally regarded as a safe (GRAS) production organism, *A. niger* is also utilized industrially to produce large amounts of plant biomass degrading enzymes. These enzymes are subsequently used for food applications or for the hydrolysis of plant biomass to produce bioethanol and high-value biochemicals (Gupta et al., 2016; Kowalczyk et al., 2014). The requirement for the presence of specific inducers for the activated expression of particular enzyme networks hampers the use of cheap and sustainable feedstock for fungal fermentation. Approaches to modulate transcriptional regulation in order to rationally design fungal strains with increased or constitutive (i.e. inducer-independent) production of plant biomass degrading enzymes are reviewed in **Chapter 2**. Overexpression and constitutive activation of transcriptional activators, accumulation of inducers, and deletion of substrate-specific repressors or wide domain repressors involved in CCR are the main approaches discussed.

Plant biomass degrading enzymes are named based on the plant polysaccharide they act on. For example, enzymes that degrade the complex heteropolysaccharide pectin are called pectinases. Pectin is mainly made of polygalacturonic acid which is a linear polymer of D-galacturonic acid (GA) (Caffall and Mohnen, 2009). *A. niger* can secrete a large number of enzymes acting on pectin (Martens-Uzunova and Schaap, 2009; de Vries et al., 2017). The objective of the research described in this thesis is to understand the transcriptional regulation of pectin degrading enzymes in *A. niger* and to use this knowledge to construct strains constitutively expressing the genes encoding pectinases. The transcription of the genes encoding pectinases is induced when grown on GA and requires a common GA-responsive element in the promoter regions of these genes (Martens-Uzunova and Schaap, 2008; Niu et al., 2015). These observations indicated that there is a GA-responsive TF coordinating the expression of the genes encoding pectinases. In **Chapter 3**, the GA-responsive transcriptional activator GaaR in *A. niger* is identified by homology to BcGaaR in *Botrytis cinerea* (Zhang et al., 2016). In **Chapter 5**, the physiological inducer of GA-responsive genes is found to be the GA catabolic pathway intermediate 2-keto-3-deoxy-L-galactonate.

A forward genetic screen was conducted in search for mutants that constitutively express pectinases. As described in **Chapter 4**, this search led to the identification of the repressor protein GaaX. The same screen also revealed that the W361R mutation in GaaR results in constitutive expression of the genes encoding pectinases (**Chapter 7**). In **Chapter 7**, we also show the use of the CRISPR-Cas9 technology to construct strains expressing point mutated or eGFP-tagged versions of *gaaR* from the endogenous locus. Furthermore, **Chapter 6** describes that overexpression of *gaaR* also brings about inducer-independent expression of the genes encoding pectinases in *A. niger*.

In this thesis, the key players in the transcriptional regulatory system of pectinases in *A. niger*, such as the transcriptional activator GaaR, the repressor GaaX and the physiological inducer 2-keto-3-deoxy-L-galactonate, are discovered. Several approaches to exploit this system for increased or constitutive expression of the genes encoding pectinases are also presented, such as overexpression and constitutive activation of *gaaR*, accumulation of 2-keto-3-deoxy-L-galactonate, and deletion of *gaaX* or *creA*, the main transcriptional repressor involved in CCR. An overall conclusion and outlook to the future is given in **Chapter 8**.

REFERENCES

Adnan M, Zheng W, Islam W, Arif M, Abubakar Y, Wang Z, Lu G (2017) Carbon catabolite repression in filamentous fungi. *International Journal of Molecular Sc*iences 19(48)

Caffall KH, Mohnen D (2009) The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydrate Research* 344:1879-1900

Culleton H, Mckie V, de Vries RP (2013) Physiological and molecular aspects of degradation of plant polysaccharides by fungi: What have we learned from *Aspergillus*?. *Biotechnology Journal* 8(8):884–894

de Vries RP, Riley R, Wiebenga A, Aguilar-Osorio G, Amillis S, Uchima CA, Anderluh G, Asadollahi M, Askin M, Barry K, Battaglia E, Bayram Ö, Benocci T, Braus-Stromeyer SA, Caldana C, Cánovas D, Cerqueira GC, Chen F, Chen W, Choi C, Clum A, Dos Santos RA, Damásio AR, Diallinas G, Emri T, Fekete E, Flipphi M, Freyberg S, Gallo A, Gournas C, Habgood R, Hainaut M, Harispe ML, Henrissat B, Hildén KS, Hope R, Hossain A, Karabika E, Karaffa L, Karányi Z, Kraševec N, Kuo A, Kusch H, LaButti K, Lagendijk EL, Lapidus A, Levasseur A, Lindquist E, Lipzen A, Logrieco AF, MacCabe A, Mäkelä MR, Malavazi I, Melin P, Meyer V, Mielnichuk N, Miskei M, Molnár ÁP, Mulé G, Ngan CY, Orejas M, Orosz E, Ouedraogo JP, Overkamp KM, Park HS, Perrone G, Piumi F, Punt PJ, Ram AF, Ramón A, Rauscher S, Record E, Riaño-Pachón DM, Robert V, Röhrig J, Ruller R, Salamov A, Salih NS, Samson RA, Sándor E, Sanguinetti M, Schütze T, Sepčić K, Shelest E, Sherlock G, Sophianopoulou V, Squina FM, Sun H, Susca A, Todd RB, Tsang A, Unkles SE, van de Wiele N, van Rossen-Uffink D, Oliveira JV, Vesth TC, Visser J, Yu JH, Zhou M, Andersen MR, Archer DB, Baker SE, Benoit I, Brakhage AA, Braus GH, Fischer R, Frisvad JC, Goldman GH, Appl Microbiol Biotechnol Houbraken J, Oakley B, Pócsi I, Scazzocchio C, Seiboth B, vanKuyk PA, Wortman J, Dyer PS, Grigoriev IV (2017) Comparative genomics reveals high biological diversity and specific adaptations in the industrially and medically important fungal genus *Aspergillus*. *Genome Biology* 18(28)

Gherbawy Y, Elhariry H, Kocsubé S, Bahobial A, Deeb BE, Altalhi A, Varga J, Vágvölgyi C (2015) Molecular characterization of black *Aspergillus* species from onion and their potential for ochratoxin A and fumonisin B2 production. *Foodborne Pathogens and Disease* 12(5):414-423

Gupta VK, Kubicek CP, Berrin JG, Wilson DW, Couturier M, Berlin A, Filho EXF, Ezeji T (2016) Fungal Enzymes for Bio-Products from Sustainable and Waste Biomass. *Trends Biochemical Sciences* 41(7):633-645

Kowalczyk JE, Benoit I, de Vries RP (2014) Regulation of plant biomass utilization in *Aspergillus*. *Advances in Applied Microbiology* 88:31-56

Krijgsheld P, Bleichrodt R, van Veluw GJ, Wang F, Müller WH, Dijksterhuis J, Wösten HAB (2013) Development in *Aspergillus*. *Studies in Mycology* 74(1):1-29

Martens-Uzunova ES, Schaap PJ (2008) An evolutionary conserved D-galacturonic acid metabolic pathway operates across filamentous fungi capable of pectin degradation. *Fungal Genetics and Biology* 45:1449–1457

Martens-Uzunova ES, Schaap PJ (2009) Assessment of the pectin degrading enzyme network of *Aspergillus niger* by functional genomics. *Fungal Genetics and Biology* 46:S170–S179

Niu J, Homan TG, Arentshorst M, de Vries RP, Visser J, Ram AF (2015) The interaction of induction and repression mechanisms in the regulation of galacturonic acid-induced genes in *Aspergillus niger*. *Fungal Genetics and Biology* 82:32–42

Stajich JE, Berbee ML, Blackwell M, Hibbett DS, James TY, Spatafora JW, Taylor JW (2009) The fungi. *Current Biology* 19(18):R840-R845

Zhang L, Lubbers RJM, Simon A, Stassen JHM, Ribera PRV, Viaud M, van Kan JAL (2016) A novel Zn_2Cys_6 transcription factor BcGaaR regulates D-galacturonic acid utilization in *Botrytis cinerea*. *Molecular Microbiology* 100:247–262

CHAPTER 2

Modulating transcriptional regulation of plant biomass degrading enzyme networks for rational design of industrial fungal strains

Alazi E, Ram AFJ

Submitted to *Frontiers in Bioengineering and Biotechnology*

ABSTRACT

Filamentous fungi are the most important microorganisms for the industrial production of plant polysaccharide degrading enzymes due to their unique ability to secrete these proteins very efficiently. These carbohydrate active enzymes (CAZymes) are utilized industrially for the hydrolysis of plant biomass for the subsequent production of biofuels and high-value biochemicals. The expression of the genes encoding plant biomass degrading enzymes is tightly controlled.

Naturally, large amounts of CAZymes are produced and secreted only in the presence of the plant polysaccharide they specifically act on. The signal to produce is conveyed via so-called inducer molecules which are di- or mono-saccharides (or derivatives thereof) released from the specific plant polysaccharides. The presence of the inducer results in the activation of a substrate-specific transcription factor (TF), which is required for the controlled expression of the genes encoding not only the CAZymes, but also the transporters and catabolic pathway enzymes needed to utilize the released monosaccharide. Over the years, several substrate-specific TFs involved in the degradation of cellulose, hemicellulose, pectin, starch and inulin have been identified in several fungal species and systems biology approaches have made it possible to uncover the enzyme networks controlled by these TFs.

The requirement for specific inducers for TF activation and subsequently the expression of particular enzyme networks determines the choice of feedstock to produce enzyme cocktails for industrial use. It also results in batch-to-batch variation in the composition and amounts of enzymes due to variations in sugar composition and polysaccharide decorations of the feedstock which hampers the use of cheap feedstocks or waste streams for constant quality of enzyme cocktails. It is therefore of industrial interest to produce specific enzyme cocktails constitutively and independently of inducers.

In this review, we focus on the methods to modulate TF activities for inducer-independent production of CAZymes and highlight various approaches that have been used to construct strains displaying constitutive expression of plant biomass degrading enzyme networks. These approaches and combinations thereof have been also used to construct strains

displaying increased expression of CAZymes under inducing conditions, and make it possible to design strains in which different enzyme mixtures are simultaneously produced independently of the carbon source.

INTRODUCTION

Plant biomass is the most abundant renewable carbon source in the world and represents the major natural substrate for fungi (Kowalczyk et al., 2014). Plant biomass mainly consists of plant cell wall material, which contains the polysaccharides, i.e. cellulose, hemicellulose and pectin, lignin, and structural proteins (Loqué et al., 2015). Cellulose, the most abundant plant cell wall polysaccharide, is a linear chain of glucose (Kolpak and Blackwell, 1976). Hemicelluloses are complex heteropolysaccharides with xylan- (linear chain of xylose), glucan- (linear chain of glucose), or mannan- (linear chain of mannose) backbones, with several different types of monomers attached to the backbone to give e.g. arabinoxylan or xyloglucan (Scheller and Ulvskov, 2010). Pectins are other complex polysaccharides containing D-galacturonic acid (GA) as the main sugar acid in their backbones. Polygalacturonic acid (PGA) is a linear chain of GA and the most abundant pectin substructure. Other pectin substructures are rhamnogalacturonan I (linear chain of alternating GA and rhamnose residues), rhamnogalacturonan II and xylogalacturonan, and contain several different types of monomers or polymers attached to their backbones (Caffall and Mohnen, 2009). Plant biomass also contains the plant cell storage polysaccharides starch, a linear (amylose) or branched (amylopectin) polymer of glucose, and inulin, a polymer of fructose residues with a terminal glucose residue (Gidley, 2001; Ritsema and Smeekens, 2003).

Polysaccharides present in plant biomass are the target of the degrading enzymes secreted abundantly by filamentous fungi. Fungal carbohydrate active enzymes (CAZymes) are also utilized industrially for the hydrolysis of plant biomass for the subsequent production of mainly bioethanol, and high-value biochemicals (Gupta et al., 2016; Kowalczyk et al., 2014; Benocci et al., 2017). Plant biomass degrading enzymes are classified into families based on their sequence, such as glycoside hydrolases, polysaccharide lyases and carbohydrate esterases, abbreviated as GH, PL and CE, respectively (Carbohydrate Active Enzymes database, http://www.cazy.org/) (Lombard et al., 2014). Filamentous fungi secrete large amounts of CAZymes only in the presence of the plant polysaccharide they specifically act on. For a review on the substrate specificity of CAZymes we refer to Kowalczyk et al., 2014 and de Vries et al., 2017. The signal for tailor-made production of CAZymes is conveyed via so-called inducer molecules which are di- or mono-saccharides (or derivatives thereof) released from the specific plant polysaccharides. The presence of the inducer results in the activation of a substrate-specific transcription factor (TF), which is required for the controlled expression of the genes encoding not only the CAZymes, but usually also the transporters and catabolic pathway enzymes needed to utilize the released monosaccharide (Culleton et al., 2013; Benocci et al., 2017).

Several substrate-specific TFs involved in plant biomass degradation have been identified in filamentous fungal species. For a review on the TFs involved in plant biomass

degradation we refer to Huberman et al., 2016 and Benocci et al., 2017. For example, it has been shown that induction of expression of the genes encoding cellulases when cellulose is present requires the presence of the transcriptional activators Clr-1 and Clr-2 in *Neurospora crassa*, ClrB in *Aspergillus nidulans* and *Penicillium oxalicum*, ManR in *A. oryzae* and Xyr1 in *Trichoderma reesei* (Coradetti et al., 2012; Li et al., 2015; Ogawa et al., 2013; Stricker et al., 2006). Clr-2, ClrB and ManR are orthologs (Kunitake and Kobayashi, 2017), while Xyr1 shows significant sequence homology to the TFs involved in xylan degradation, XlnR in *A. niger* and Xlr-1 in *N. crassa* (Rauscher et al., 2006). The enzyme families controlled by orthologous transcription factors can vary. Although Xyr1 in *T. reesei* is essential for the expression of cellulases and hemicellulases, Xyr1 orthologs (named XlnR or Xlr-1 in other species) are essential only for the expression of hemicellulases (e.g. xylanases) in filamentous fungal species, such as *A. oryzae* (on xylose), *A. niger* (on xylan), *A. nidulans* (on xylan), *Talaromyces cellulolyticus* (on xylan), *P. oxalicum* (on cellulose), *N. crassa* (on xylan), *Fusarium oxysporum* (on xylan or wheat cell walls) and *Myceliophthora thermophila* (Stricker et al., 2006; Marui et al., 2002b; van Peij et al., 1998; Tamayo et al., 2008; Fujii et al., 2014; Li et al., 2015; Sun et al., 2012; Brunner et al., 2007; Wang et al., 2015). In *A. oryzae, A. niger*, *T. cellulolyticus* and *P. oxalicum*, contribution of XlnR to the expression of cellulases during growth on cellulose and/or xylose/xylan has also been reported (Marui et al., 2002; Gielkens et al., 1999; Okuda et al., 2016; Li et al., 2015). Arabinan, present in the side chains of some form of hemicellulose, such as arabinogalactan, or some pectins, is degraded via CAZymes that are transcriptionally regulated by AraR in *A. niger* (Battaglia et al., 2011; Kowalczyk et al., 2017). The TF GaaR, required for the expression of CAZymes degrading pectin, especially PGA, has been identified in both *Botrytis cinerea* and *A. niger* (Zhang et al., 2016; Alazi et al., 2016). Recently, Pdr-1 in *N. cr*assa was described as a TF required for the expression of the genes encoding CAZymes that degrade several pectin substructures including PGA and rhamnogalacturonan I (Thieme et al., 2017). Transcriptional control of the genes involved in starch degradation is conducted by AmyR in *A. niger* and *A. nidulans* (vanKuyk et al., 2012; Tani et al., 2001).

The transcriptional activators involved in plant biomass degradation mentioned above were identified either via classical approaches, such as mutant complementation (*xlnR* in *A. niger*) and gene cloning (*xlnR* and *amyR* in *A. nidulans* and *amyR A. niger*), or via post-genomic approaches, such as yeast one-hybrid screening (*gaaR* in *B. cinerea*), screening TF deletion mutant collections (*pdr-1* and *clr-2* in *N. crassa*, *manR* in *A. oryzae*, *clrB* and *xlnR* in *P. oxalicum*), and also based on sequence homology (*clrB* in *A. nidulans*, *araR* and *gaaR* in *A. niger*, *xyr1* in *T. reesei and M. thermophila*, *xlnR* in *F. oxysporum*, *T. cellulolyticus* and *A. oryzae*) or their expression levels in the transcriptomics data (*xlr-1* in *N. crassa*) (van Peij et al., 1998; Tani et al., 2001; vanKuyk et al., 2012; Zhang et al., 2016; Thieme et al., 2017; Coradetti et al., 2012; Ogawa et al., 2012; Li et al., 2015; Battaglia et al., 2011; Alazi et al., 2016; Rauscher et al., 2006; Wang et al., 2015; Brunner et al., 2007; Fujii et al., 2014; Marui et al., 2002b; Sun et al., 2012). Positively acting TFs involved in controlling expression of plant cell wall degrading enzymes belong to the Zn_2Cys_6 type family of TFs. TFs belonging to this family are found specifically in fungi (both yeasts and filamentous fungi) and contain a DNA-binding domain with six cysteine residues bound to two zinc atoms, usually close to their NH_2-terminal end. Most of the Zn_2Cys_6 type TFs also contain a

fungal-specific TF domain, known as the middle homology region, with a proposed role in regulating the activity of the TF (MacPherson et al., 2006).

Diversity in the control of transcription factor activities

The mechanisms by which the activity of TFs is regulated in response to stimuli can be diverse. Firstly, the protein level of the TFs might be controlled. This can be realized by regulation at the level of transcription, mRNA stability or protein stability resulting in low protein levels under non-inducing conditions and higher levels under inducing conditions. Secondly, the subcellular localization (nuclear import/export), DNA binding activity, and/or transcriptional activity of the TFs might be regulated by post-transcriptional modifications, such as phosphorylation, and/or via protein-protein or protein-metabolite interactions (MacPherson et al., 2006; Chang and Ehrlich, 2013; Tani et al., 2014). Finally, DNA site occupancy of TFs might depend on their cooperation or competition with other proteins and the chromatin accessibility (Biggin, 2011; Granek and Clarke, 2005).

Our knowledge about regulation of the activity of the TFs involved in biomass degradation is limited, yet growing. The amount of Clr-2 in *N. crassa* is regulated at the level of transcription by Clr-1, which gets activated and positively regulates *clr-2* expression only under inducing conditions. The expression of the *clr-2* ortholog in *A. nidulans*, *clrB*, on the other hand, is not drastically induced under inducing conditions and the Clr-1 ortholog ClrA is not essential but only contributes to cellulase gene expression (Coradetti et al., 2012). This comparison indicates that activation mechanisms of even orthologous transcription factors differ between fungal species.

The expression of *xyr1* in *T. reesei* is subject to Cre1-dependent carbon catabolite repression (CCR) (see below) and induced on carbon sources inducing cellulase expression (e.g. lactose, sophorose or cellulose), but not on carbon sources inducing hemicellulase expression (e.g. xylose) (Mach-Aigner et al., 2008; Portnoy et al., 2011; Lichius et al., 2014). Xyr1 was shown to accumulate in the nucleus during growth on an inducing carbon source (i.e. sophorose or low concentration of xylose), whereas it is degraded in the nucleus during growth on a repressing carbon source (i.e. glucose or a high concentration of xylose). The increased amount of nuclear Xyr1 correlates with the increased expression level of cellulase gene *cbh1* on sophorose, and that of xylanase gene *xyn2* on a low concentration of xylose (Lichius et al., 2014). Similar to *T. reesei xyr1*, the expression of *xlnR* in *A. nidulans* and *P. oxalicum* is also subject to CreA-dependent CCR (see below), whereas *xlnR* in *A. niger* is not regulated at the level of transcription, but constitutively transcribed at low levels (Tamayo et al., 2008; Li et al., 2015; Mach-Aigner et al., 2012). Moreover, XlnR was found to be localized in the nucleus regardless of the presence of inducer in *A. niger* (Hasper, 2004). In *F. oxysporum*, *xlnR* is transcriptionally regulated by both CCR and induction on xylose/xylan (Calero-Nieto et al., 2007). XlnR in *A. oryzae* is reversibly phosphorylated in response to xylose, which does not affect its protein stability and correlates with the expression of XlnR target genes (Noguchi et al., 2011). Recently, Kunitake and Kobayashi proposed that a conserved sequence in XlnR is involved in the xylose-mediated phosphorylation of XlnR in *A. oryzae*, implying a conserved mechanism regulating XlnR activity among Ascomycete fungi. Moreover, the observation that XlnR in

A. oryzae is not phosphorylated in response to cellobiose, but required for cellulase expression, indicates that its activity is regulated by a different mechanism on cellobiose (Kunitake and Kobayashi, 2017; Noguchi et al., 2011).

The activity of the GA-responsive transcriptional activator GaaR in *A. niger* is regulated by a different mechanism. It is suggested to be inhibited by the repressor protein GaaX via protein-protein interaction under non-inducing conditions. Under inducing conditions, the inducer is proposed to bind to GaaX, resulting in a free and active form of GaaR (Niu et al., 2017). The amount of GaaR is not significantly regulated at the level of transcription (Alazi et al., 2016; Alazi et al., 2018). In addition, GaaR-eGFP was shown to be constitutively localized in the nucleus (Alazi et al., 2018). Transcription of *pdr-1* in *N. cr*assa is induced under inducing conditions and repressed under repressing conditions. Moreover, the activity of Pdr-1 is suggested to be post-transcriptionally regulated, as its nuclear accumulation and transcriptional activity requires the presence of rhamnose in a strain overexpressing *pdr-1* (Thieme et al., 2017).

The transcription of *amyR* is upregulated on inducing carbon sources (e.g. starch or maltose) and is subject to CreA-dependent CCR (see below) in *A. nidulans* (Tani et al., 2001). Further, AmyR was shown to localize in the nucleus in an inducer (i.e. isomaltose)-dependent manner and activate the expression of its target genes (Makita et al., 2009).

The role of CreA-dependent carbon catabolite repression in the production of plant biomass degrading enzymes

Apart from being upregulated under inducing conditions, the expression of CAZymes might also be controlled via CCR when a more energetically favorable carbon source, such as glucose, compared to plant biomass polysaccharides is available for fungi. The Cys_2His_2 type TF CreA/Cre1/Cre-1 (in *Aspergillus* species, *T. reesei*, and *N. crassa*, respectively) is the main transcriptional repressor contributing to CCR. Regulation of CreA activity has not been fully understood, but several studies have indicated that post-transcriptional modifications, such as ubiquitination and phosphorylation, affect CreA protein stability, subcellular localization and/or DNA binding activity (Please see Adnan et al., 2018 for a recent review; Ries et al., 2016; Cziferszky et al., 2002). CreA not only represses the expression of genes encoding CAZymes, but it might also repress the expression of some transcriptional activators, such as *clrB* and *xlnR* in *P. oxalicum*, *xyr1/xlnR* in *T. reesei*, *A. nidulans* and *F. oxysporum*, and *amyR* in *A. nidulans*, that are required for the expression of CAZymes (Li et al., 2015; Mach-Aigner et al., 2008; Tamayo et al., 2008; Calero-Nieto et al., 2007; Tani et al., 2001). Furthermore, CreA might repress the expression of CAZymes under high xylose concentrations, too, as was shown to be the case for Xyr1/XlnR target genes in *T. reesei* and *A. nidulans*, and for XlnR itself in *A. oryzae* (Nakari-Setälä et al., 2009; Tamayo et al., 2008; Ichinose et al., 2017).

Elimination of CCR due to a lack of function of CreA and/or CreB, a ubiquitin-specific protease involved in CCR, or their orthologs has been reported to result in an increased expression of CAZymes degrading cellulose, hemicellulose, pectin or starch, under inducing and/or non-inducing conditions in filamentous fungi (Prathumpai et al., 2004; Mach-Aigner et al., 2008; Nakari-Setälä et al., 2009; Denton and Kelly, 2011; Fujii et al.,

2013; Ichinose et al., 2014; Niu et al., 2015; Jang et al., 2015; Ichinose et al., 2017). However, even when CreA-dependent CCR is circumvented, the expression of CAZymes might require the presence of active transcriptional activators, which is normally achieved in the presence of inducing metabolites. For example, in a Cre1-negative *T. reesei* strain (Rut-C30), the full expression of Xyr1 target genes requires the presence of the inducer (i.e. xylose) (Mach-Aigner et al., 2008). Another study revealed that deletion of *cre1* in *T. reesei* results in elevated production of cellulases and hemicellulases under inducing and, to a lesser extent, under non-inducing conditions (Nakari-Setälä et al., 2009). Although the expression of pectinases is subject to CreA-dependent CCR in *A. niger*, deletion of *creA* is not sufficient for an increased production of polygalacturonases under non-inducing conditions. Polygalacturonases are only produced in the presence of the inducer (i.e. GA) or when *gaaX* is deleted, showing that GA-responsive gene expression requires the presence of active GaaR relieved from GaaX inhibition even in a *ΔcreA* strain (Niu et al., 2015; Niu et al., 2017).

APPROACHES FOR INCREASED OR CONSTITUTIVE EXPRESSION OF PLANT BIOMASS DEGRADING ENZYMES BY MODULATING THEIR TRANSCRIPTIONAL REGULATION

Plant biomass with varying sugar composition present in waste streams from agriculture, forestry and food industries represent a cheap, sustainable and renewable feedstock for the production of the CAZymes, and subsequently, valuable chemicals (Meyer et al., 2015; Sweeney and Xu, 2012). However, the composition of the enzyme cocktail will vary because of variation in the composition of the plant biomass. It is therefore of industrial interest to produce specific fungal enzyme cocktails constitutively, independently of inducers and the substrate used. Fungal production strains, such as *A. niger* CBS513.88, *T. reesei*, Rut-C30 and *P. oxalicum* JU-A10-T, have been for a long time improved via multiple rounds of classical mutagenesis and screening approaches that can be time-consuming (Montenecourt and Eveleigh, 1979; Pel et al., 2007; Fang et al., 2010; van Hanh et al., 2010; Ho and Ho, 2015; Yao et al., 2015). However, emergence of the -omics era (genomics, transcriptomics etc.) and advances in recombinant technologies allow nowadays efficient strain improvement via genetic engineering approaches with minimal alterations in the genome (Meyer et al., 2010; Liu et al., 2013). In the remaining part of this review, we discuss the approaches to modulate transcriptional regulation in order to rationally design fungal strains with increased or constitutive production of plant biomass degrading enzymes, examples of which are given in Table 1. As will become clear in the following sections, the success of the approach used highly depends on the mechanism regulating the activity of the targeted TF.

Overexpression of TFs

Increasing the amount of a TF at the level of transcription can be achieved by expressing multiple copies of the TF via its endogenous promoter, or (multiple copies of) the TF via a strong inducible/constitutive promoter. Overexpression of several TFs in *Saccharomyces cerevisiae* has been shown to result in an increased expression of the TF target genes, even under non-inducing conditions (Chua et al., 2006).

Table 1. Examples of rational design of industrial fungal strains with increased or constitutive production of CAZymes.

Target - Organism	Approach	CAZyme	Carbon source tested (Inducing/non-inducing)	Reference
Ace1 - *T. reesei*	Deletion of *ace1*	Cellulases, hemicellulases	Inducing	Aro et al., 2003
AmyR - *A. nidulans*	Constitutive activation in $AmyR_{1-514}$	Amylases	Inducing and non-inducing	Makita e al., 2009
AmyR - *A. niger*	Overexpression of *amyR*	Cellulases, hemicellulases, amylases	Inducing	vanKuyk et al., 2012
Clr-2 - *N. crassa*	Overexpression of *clr-2*	Cellulases	Inducing and non-inducing	Coradetti et al., 2013
ClrB - *A. nidulans*	Overexpression of *clrB*	Cellulases	Inducing	Coradetti et al., 2013
ClrB (and CreA)(and β-glucosidase) - *P. oxalicum*	Overexpression of *clrB* (and deletion/a lack of function of *creA*) (and deletion of *bgl2*)	Cellulases (and hemicellulases)	Inducing (and non-inducing)	Li et al., 2015; Yao et al., 2015; Chen et al., 2013
CreA/CreB or their orthologs - please see the text	Elimination of CCR by deletion/a lack of function of *creA*/*creB*	Cellulases, hemicellulases, pectinases, amylases	Inducing and/or non-inducing	Please see the text
GA catabolic pathway - *A. niger*	Deletion of *gaaC*	Pectinases	Inducing	Alazi et al., 2017
GaaR (and CreA) - *A. niger*	Overexpression of *gaaR* (and deletion/a lack of function of *creA*)	Pectinases	Inducing and non-inducing	Alazi et al., 2018
	Constitutive activation in $GaaR^{W361R}$	Pectinases	non-inducing	Alazi and Ram, manuscript in prep.
GaaX - *A. niger*	Deletion of *gaaX*	Pectinases	Non-inducing	Niu et al., 2017
GCN5-related N-acetyltransferase- *T. reesei*	Overexpression of gene 123668	Cellulases	Inducing	Häkkinen et al., 2014
Hcr-1 - *N. crassa*	Deletion of *hcr-1*	Hemicellulases	Inducing	Li et al., 2014
Lae1 - *T. reesei*	Overexpression of *lae1*	Cellulases	Inducing	Seiboth et al., 2012
ManR - *A. oryzae*	Overexpression of *manR*	Cellulases, hemicellulases	Inducing	Ogawa et al., 2013; Ogawa et al., 2012
Mhr1 - *M. thermophila*	Gene silencing of *mhr1*	Cellulases, hemicellulases	Inducing	Wang et al., 2018
Pdr-1 - *N. crassa*	Overexpression of *pdr-1*	Pectinases	Inducing	Thieme et al., 2017

Table 1. Examples of rational design of industrial fungal strains with increased or constitutive production of CAZymes. (Continued)

Target - Organism	Approach	CAZyme	Carbon source tested (Inducing/non-inducing)	Reference
Rce1 - *T. reesei*	Deletion of *rce1*	Cellulases	Inducing	Cao et al, 2017
Synthetic TF - *P. oxalicum*	Overexpression of cx^{c}-*s*	Cellulases, hemicellulases	Inducing and non-inducing	Gao et al., 2017
Synthetic TF - *T. reesei*	Overexpression of *xyr1-cre1*$_b$	Cellulases, hemicellulases	Non-inducing	Zhang et al., 2017
SxlR - *T. reesei*	Deletion of *sxlR*	Hemicellulases	Inducing	Liu et al., 2017
XlnR - *A. nidulans*	Overexpression of *xlnR*	Hemicellulases	Inducing	Tamayo et al, 2008
XlnR - *A. niger*	Overexpression of *xlnR*	Cellulases, hemicellulases	Inducing	Gielkens et al., 1999
	Constitutive activation in $XlnR^{V756F}$/$XlnR^{L668Stop}$	Hemicellulases	Inducing and non-inducing	Hasper, 2004; Hasper et al., 2004
XlnR - *A. oryzae*	Overexpression of *xlnR*	Cellulases, hemicellulases	Inducing	Marui et al., 2002; Noguchi et al., 2009
XlnR - *F. oxysporum*	Overexpression of *xlnR*	Hemicellulases	Inducing	Calero-Nieto et al., 2007
XlnR - *P. oxalicum*	Constitutive activation in $XlnR^{A871V}$	Cellulases, hemicellulases	Inducing	Gao et al., 2017
XlnR - *T. cellulolyticus*	Overexpression of *xlnR*	Cellulases	Inducing	Okuda et al., 2016
Xlr-1 - *N. crassa*	Constitutive activation in Xlr-1^{A828V}	Hemicellulases	Inducing and non-inducing	Craig et al., 2015
Xpp1 - *T. reesei*	Deletion of *xpp1*	Hemicellulases	Inducing	Derntl et al., 2015
Xylose/arabinose catabolic pathway - *T. reesei*	Deletion of xylose/arabinose catabolic pathway genes	Hemicellulases	Inducing	Herold et al., 2013
Xyr1 - *M. thermophila*	Overexpression of *xyr1*	Hemicellulases	Inducing and non-inducing	Wang et al., 2015
Xyr1 - *T. reesei*	Overexpression of *xyr1* (and gene silencing of *ace1* in RUT-C30)	Cellulases (and hemicellulases)	Inducing and non-inducing	Lv et al., 2015; Wang et al., 2013
	Constitutive activation in $Xyr1^{A824V}$	Cellulases, hemicellulases	Inducing and non-inducing	Derntl et al., 2013
β-glucosidases - *N. crassa*	Deletion of *gh1-1*, *gh3-3* and *gh3-4* (and deletion/a lack of function of *cre-1)*	Cellulases	Inducing	Znameroski et al., 2012

Inducer-independent production of cellulases in *N. crassa* was achieved by constitutively overexpressing *clr-2* via the *ccg-1* promoter (Coradetti et al., 2013). While deletion of *clr-2* resulted in a drastic decrease in the expression of most of the genes encoding cellulases under inducing conditions (i.e. on insoluble crystalline cellulose (Avicel)), overexpression of *clr-2* resulted in an increased expression of cellulases under inducing conditions and constitutive expression under non-inducing conditions. Expression of the genes encoding cellulases was higher under starvation (no carbon) conditions than on sucrose highlighting the effect of CCR on cellulase encoding genes under repressing conditions (i.e. on sucrose) even when *clr-2* was overexpressed (Coradetti et al., 2013).

Unlike overexpression of *clr-2* in *N. crassa*, overexpression of the *clr-2* ortholog *clrB* in *A. nidulans* or in *P. oxalicum*, both via the constitutive *A. nidulans gpdA* promoter, did not result in an increased expression of CAZymes encoding genes under non-inducing conditions, but only under inducing conditions (Coradetti et al., 2013; Li et al., 2015). It was shown that deletion of *creA* in combination with *clrB* overexpression allows increased expression of cellulases under non-inducing conditions in *P. oxalicum*, indicating that the strong CreA-dependent repression on cellulase genes overrules ClrB-dependent induction (Li et al., 2015).

ManR in *A. oryzae* is the *N. crassa clr-2* ortholog, regulating the expression of genes encoding cellulose and hemicellulose (including mannan, but not xylan) degrading enzymes in *A. oryzae* (Ogawa et al., 2012; 2013). Overexpression of ManR via the *tef1* promoter resulted in an increased expression of cellulases and hemicellulases involved in mannan degradation under inducing conditions (i.e. on Avicel and mannan, respectively) (Ogawa et al., 2013; Ogawa et al., 2012). The effect of overexpression of ManR under non-inducing conditions has not been reported.

As mentioned before, in *T. reesei*, Xyr1 controls the transcriptional regulation of both cellulase encoding and hemicellulase encoding genes. Expression of *xyr1* is subject to CCR and *xyr1* expression is upregulated on carbon sources inducing cellulase production. Overexpression of *xyr1* via the *tcu1* promoter resulted in an inducer-independent production of cellulases, even under repressing conditions (i.e. on glucose) (Lv et al., 2015). All reported *T. reesei* Xyr1 orthologs in other filamentous fungal species regulate mainly the transcription of hemicellulase encoding genes and contribute less to the transcriptional regulation of cellulase encoding genes. Overexpression of *xlnR* via *gpdA* promoter in *T. cellulolyticus* resulted in an increased production of cellulases on cellulose, but not that of xylanases (Okuda et al., 2016). In *A. niger*, increased expression of cellulases and xylanases on xylose was observed in a strain carrying multiple copies of *xlnR* (Gielkens et al., 1999). Similarly, an *A. oryzae* strain overexpressing *xlnR* via the *tef1* promoter showed an increased production of both cellulases and hemicellulases when grown on carbon sources known to induce cellulase (i.e. Avicel or cellobiose) or hemicellulase (i.e. xylose or xylan) production (Marui et al., 2002; Noguchi et al., 2009).

The studies mentioned above reported on the increased expression of hemicellulases in fungal strains overexpressing *xlnR* or its orthologs under inducing conditions. Only a few studies have reported the effect of overexpression of *xlnR* on the expression of hemicellulases under non-inducing conditions. For instance, *xlnR* overexpression via the constitutive *gpdA* promoter in *A. nidulans* or in *F. oxysporum* yields to increased

production of hemicellulase encoding genes only in the presence of inducing carbon sources (e.g. xylose or xylan) (Tamayo et al, 2008; Calero-Nieto et al., 2007). It was also reported that when GFP-XlnR is overproduced in *A. oryzae*, it localizes constitutively in the nucleus, but its target genes are expressed only in the presence of xylose (Noguchi et al., 2011). These results indicate that XlnR in these organisms gets activated by an inducer-mediated post-transcriptional mechanism and that XlnR activation under non-inducing conditions cannot simply be achieved by overexpression. In *M. thermophila*, overexpression of *xyr1* via the *pdc* promoter resulted in an increased expression of hemicellulases (xylanases) under both inducing (i.e. corncob) and non-inducing conditions (i.e. glucose) (Wang et al., 2015). This indicates that the mechanism of Xyr1 activation in *M. thermophila* differ from other filamentous fungi.

Another example of inducer-independent production of CAZymes by overexpression of TFs was given by a recent study in *A. niger*. Here it was shown that overexpression *gaaR* in *A. niger* via the *A. nidulans gpdA* promoter leads to constitutive expression of the genes encoding pectinases, as well as GA transporters and GA catabolic pathway enzymes (Alazi et al., 2018). Deletion of *creA* further enhanced pectinase production under mildly repressing conditions (i.e. on fructose) indicating competing roles of GaaR and CreA to control the expression of GA-induced genes (Alazi et al., 2018). The effect of overexpression of *gaaR* on the expression of its target genes under non-inducing conditions is likely caused by its specific activation mechanism. Regulation of the activity of GaaR includes a specific repressor protein (GaaX) (Niu et al., 2017). It has been proposed that under non-inducing conditions the activity of GaaR is controlled through direct interaction by GaaX which prevents GaaR to be active (Niu et al, 2017). Modulating the amount of GaaR by overexpression affects the balance of GaaR-GaaX and results in the presence of uncomplexed active GaaR even under non-inducing conditions (Alazi et al., 2018).

The regulation of GA-induced gene expression shows some striking similarities with the regulation of genes involved in quinic acid catabolism. Quinic acid is present as an aromatic compound in the plant cell and can be released from tannins, which are water-soluble polyphenols, by the action of tannases known to be secreted by filamentous fungi (Wagh, 2010). Similar to regulation of GA utilization in *A. niger*, regulation of quinic acid utilization in *A. nidulans* involves a Zn_2Cys_6 type transcriptional activator (QutA) and a repressor (QutR), and strains expressing multiple copies of *qutA* displayed constitutive expression of the genes encoding quinic acid catabolic pathway enzymes (Lamb et al., 1996). The QutR and GaaX repressor proteins are both multidomain proteins with sequence similarity to the three C-terminal domains of AROM, indicating that these repressors share a common evolutionary origin (Niu et al., 2017).

On the other hand, overexpression of *pdr-1*, the pectin degradation regulator in *N. crassa*, via the *gpd* promoter of *M. thermophila* resulted in elevated expression of its target genes only under inducing conditions (i.e. on rhamnose). Therefore, it was proposed that Pdr-1 activity is regulated post-transcriptionally in a manner depending on the presence of the inducer but not on the amount of Pdr-1 (Thieme et al., 2017).

An *A. niger* strain carrying multiple copies of *amyR* displayed increased expression of AmyR target genes, such as CAZymes acting on starch as well as on cellulose and

hemicellulose, under inducing conditions (i.e. on maltose, starch, and low concentration of glucose), but not under non-inducing condition (vanKuyk et al., 2012). This result is in line with the observation that in *A. nidulans*, nuclear localization of AmyR and thereby activation of its target genes is inducer-dependent (Makita et al., 2009).

In conclusion, we have seen that overexpression of TFs involved in plant biomass degradation can result in increased expression of their target genes under inducing conditions. However, in most cases overexpression of TFs, such as XlnR, Pdr-1 and AmyR, does not result in inducer-independent expression of their target genes. In these cases it is likely that an inducer molecule is required to activate the TF. The exact activation mechanism of XlnR, Pdr-1 and AmyR are currently unknown, and could involve direct interaction of the TF with its inducer, or could be related to post-translational modifications connected to the presence of inducers. Overexpression of TFs like GaaR and QutA results in inducer-independent expression of GaaR and QutA target genes most likely by titrating away the corresponding repressor proteins (GaaX and QutR, respectively). This illustrates that different mechanisms controlling TF activities result in different outcomes of overexpression of TFs under non-inducing conditions.

Constitutive Activation of TFs

The activity of a TF can be controlled in different ways, including post-transcriptional modifications affecting its localization, DNA binding or interaction with repressor protein(s). Over the past fifteen years, mutations resulting in changes in amino acid sequences and thereby in constitutively active TFs have been identified via classical mutagenesis and screening approaches, or via pre-designed amino acid substitutions, domain removal or protein truncation analyses.

The first reported constitutively active form of XlnR (XlnR^{V756F}) was identified in *A. niger* via a forward genetic screen. Expression of *xlnR^{V756F}* resulted in a constitutive expression of hemicellulases (i.e. xylanases) even under repressing conditions (i.e. on fructose or glucose) (Hasper, 2004; Hasper et al., 2004). Later, in *T. reesei*, a point mutation in Xyr1 (Xyr1^{A824V}) introduced via UV mutagenesis was found to result in a constitutively active Xyr1 and constitutive expression of cellulases and hemicellulases (i.e. xylanases) even in the presence of a repressing carbon source (Derntl et al., 2013). In addition, overexpression of *xyr1^{A824V}* in *T. reesei* was shown to be more effective in increasing cellulase production than overexpression of *xyr1* under inducing conditions (e.g. Avicel) (Jiang et al., 2016). Both amino acids changes (V756F in XlnR and A824V in Xyr1) are located within the same predicted α-helix in the fungal-specific TF domain of XlnR/Xyr1 (Derntl et al., 2013). Although overexpression of *xlr-1* in *N. crassa* via the *ccg-1* promoter did not yield to constitutive expression of hemicellulases (i.e. xylanases), overexpression of *xlr-1^{A828V}*, which carries the homologous mutation as in *xyr1^{A824V}*, resulted in a constitutive and increased production of hemicellulases under both inducing (i.e. on xylan) and non-inducing conditions (Craig et al., 2015). Similarly, overexpression of *xlnR* carrying the homologous point mutation (*xlnR^{A871V}*) using the *PDE_02864* promoter in a *P. oxalicum* strain that lacks *creA* and overexpresses *clrB* (see above), enabled even more increased

production of cellulases and hemicellulases under inducing conditions (i.e. on wheat bran) (Gao et al., 2017).

Recently, via a forward genetic screen, several different point mutations throughout AraR were found to give rise to a constitutively active AraR and constitutive expression of AraR target genes (*abfA*, *abfB* and *abfC*). Unlike the mutations in XlnR that gave rise to inducer-independent and constitutive expression of its target genes, the mutations in AraR lead to constitutive expression of AraR target genes only under de-repressed conditions. Deletion of *creA* improved the production of CAZymes that degrade arabinan to a large extent, indicating the strong CCR on the genes encoding these CAZymes under repressing conditions on fructose (Reijngoud, Deseke and Ram, unpublished results).

We recently conducted a large forward genetic screen for mutants with constitutive expression of pectinases (Niu et al., 2017). Apart from identifying GaaX as a repressor for GA-induced genes expression, it also brought about the identification of a constitutive allele of GaaR ($GaaR^{W361R}$) resulting in constitutive expression of pectinases under even repressing conditions (i.e. on glucose or fructose). Within GaaR, W361 is situated in the fungal-specific TF domain and is highly conserved among *Aspergillus* species (Alazi and Ram, manuscript in preparation).

Deletion of the C-terminal regions of several TFs involved in plant biomass degradation was shown to result in constitutive activation of the TFs indicating that these C-terminal parts contain an inhibitory domain. For example, in *A. niger*, truncation of XlnR from L668 resulted in a constitutive expression of hemicellulases (Hasper et al., 2004), and truncation of AraR from P646 was reported to cause constitutive activation of AraR (Jiang et al., 2016). Expression of the C-terminally truncated $AmyR_{1-514}$ and $AmyR_{1-511}$ resulted in constitutive localization of AmyR in the nucleus both in *A. nidulans* and *A. oryzae*, respectively. However, while constitutive amylase production was observed in *A. nidulans* (Makita e al., 2009), loss of expression was observed in *A. oryzae* (Suzuki et al., 2015), indicating that the effect of the truncation can also be species-specific.

Deletion or Down-regulation of Specific Repressors

It has been shown that besides the general CCR, specific repressors might play a role in the transcriptional regulation of the genes encoding CAZymes. For instance, the Cys_2His_2 type TF Ace1 represses the expression of both cellulase and hemicellulase (i.e. xylanase) encoding genes, as well as the expression of *xyr1* in *T. reesei* (Aro et al., 2003; Wang et al., 2013). Furthermore, Ace1 was shown to compete with Xyr1 to bind to the same sequence in the promoter of the xylanase gene *xyn1* (Rauscher et al., 2006). Deletion of *ace1* in the wild type background resulted in an increased expression of the genes encoding cellulases and hemicellulases only under inducing conditions, and gene silencing of *ace1* led to an increase in constitutive expression of these genes when combined with the overexpression of *xyr1* via the *pdc* promoter in a Cre1-negative background (Rut-C30) (Aro et al., 2003; Wang et al., 2013). Recently, another repressor, the Zn_2Cys_6 type TF Rce1, was identified in *T. reesei* that is involved in the repression of the genes encoding cellulases, but not hemicellulases (i.e. xylanases). It was shown that Rce1 is constitutively localized in the nucleus and competes with Xyr1 to bind to the same sequence in the promoter of the

cellulase gene *cbh1* (Cao et al, 2017). Deletion of *rce1* resulted in an increased production of cellulases under inducing conditions (i.e. on cellulose) (Cao et al, 2017). Deletion of a basic helix-loop-helix TF *xpp1*, xylanase promoter-binding protein 1, in *T. reesei* resulted in an increased expression of hemicellulase (i.e. xylanase) encoding genes, but not cellulase encoding genes, at late stages of cultivation under inducing conditions (i.e. on xylan) (Derntl et al., 2015). However, Xpp1 was later described as a general regulator of both primary and secondary metabolism and therefore not a factor specifically controlling xylanases expression (Derntl et al., 2017). In addition, the Zn_2Cys_6 type transcriptional repressor SxlR was found to bind to the promoters of specific (GH11 family) xylanase genes, and deletion of *sxlR* resulted in an increased expression of these genes under inducing conditions (Liu et al., 2017).

Similarly, the Cys_2His_2 type TF Hcr-1 in *N. crassa* was shown to repress the expression of the genes encoding hemicellulases (i.e. xylanases), but not the ones encoding cellulases. A *Δhcr-1* strain exhibited an increased hemicellulase production under inducing conditions (i.e. on xylan or Avicel) (Li et al., 2014).

Recently a new TF, MhR1, was identified in *M. thermophila* as the regulator of cellulase and xylanase genes. Gene silencing of *mhr1* resulted in an increased production of cellulases and xylanases, as well as increased expression of *xyr1* and the genes encoding cellulases under inducing conditions (Wang et al., 2018).

As mentioned before, the activity of GaaR in *A. niger* is inhibited by the repressor protein GaaX, the amount of which is regulated at the level of transcription by induction on GA. GaaX was proposed to bind to- and inhibit GaaR under non-inducing conditions, and bind to the inducer molecule and release GaaR under inducing conditions. Deletion of *gaaX* resulted in a constitutive expression of pectinases, providing additional evidence for the proposed model of regulation of GA-responsive gene expression in *A. niger* (Niu et al., 2017). Similarly, deletion of QutR, the repressor protein involved in controlling the expression of quinic acid-responsive genes, also results in constitutive expression of at least eight genes involved in quinic acid uptake and metabolism (Levett et al., 2000).

Accumulation of Inducers

The intracellular accumulation of inducers is another effective method to boost the production of CAZymes by fungi. Cellulase production by *N. crassa* is greatly induced on insoluble, crystalline cellulose (Avicel). However, cellulase production is not observed on cellobiose, which is the soluble degradation product of cellulose, possibly due to rapid action of β-glucosidases converting cellobiose to glucose and subsequent glucose-mediated CCR. As first shown in *N. crassa*, deletion of three genes encoding the major β-glucosidases (intracellular enzyme Gh1-1, and extracellular enzymes Gh3-3 and Gh3-4) disrupted the hydrolysis of the inducer, cellobiose, and resulted in higher levels of expression of cellulases on cellobiose compared to the wild type strain grown on cellobiose, and similar to the wild type strain grown on Avicel. Moreover, deletion of *cre-1* further increased cellulase production on cellobiose (Znameroski et al., 2012). Similarly, deletion of the major intracellular β-glucosidase *bgl2* in a carbon catabolite de-repressed (*ΔcreA*) *P. oxalicum* strain that overexpresses *clrB* using the *A. nidulans gpdA* promoter,

gave rise to higher levels of cellulase and hemicellulase (i.e. xylanase) production on cellulose compared to the wild type strain. These were similar to the levels of the industrial strain JU-A10-T grown on wheat bran (Yao et al., 2015). A similar effect of deleting *bgl2* on cellulase and hemicellulase production was observed in *P. oxalicum* grown on cellulose (Chen et al., 2013).

Expression of hemicellulases, specifically xylanases, in *T. reesei* is induced both on xylose and arabinose. Although the catabolic pathways assimilating xylose and arabinose are interconnected, the physiological inducers triggering hemicellulase production appear to be different and were suggested to be xylose and L-arabitol, respectively. Expression of xylanases increased dramatically in single or double xylose/arabinose catabolic pathway enzyme deletion mutants, indicating the effect of accumulation of physiological inducers in these mutants (Seiboth et al., 2012; Herold et al., 2013).

One of the GA catabolic pathway intermediates, 2-keto-3-deoxy-L-galactonate, was recently shown to be the physiological inducer of the genes encoding pectinases in *A. niger*. It was demonstrated that deletion of *gaaC*, the gene encoding 2-keto-3-deoxy-L-galactonate aldolase, results in accumulation of 2-keto-3-deoxy-L-galactonate and thereby elevated expression levels of pectinase encoding genes (Alazi et al., 2017). Inducer molecules that are responsible for the activation of TFs represent therefore another interesting target to enhance expression of CAZymes by constructing strains in which the inducer accumulates either due to prevention of rapid hydrolysis of the inducer or due to inactivation of the metabolic genes that function downstream of the enzyme that forms the inducer.

FUTURE PERSPECTIVES AND CONCLUDING REMARKS

Increased expression of plant cell wall degrading enzymes can be achieved by overexpression of specific TFs or by identifying mutations in TFs leading to constitutive activation, and combinations thereof. It is also well established that CreA-dependent CCR has in many cases a negative effect on the production of enzymes and therefore CreA is an important target for improving enzyme production under both inducing and non-inducing conditions. Apart from the approaches described in the previous paragraphs, modulating chromatin accessibility is yet an under-utilized approach to increase the expression of CAZyme encoding genes in filamentous fungi. Chromatin remodeling of the promoters of CAZyme encoding genes through the actions of the histone acetyltransferase Gcn5, the CCAAT-binding complex, and possibly the putative protein methyltransferase Lae1 is required for the full expression of these CAZymes in *T. reesei* (Xin et al., 2013; Zeilinger et al., 1998; Aghcheh and Kubicek, 2015; Li et al., 2016). Overexpression of a putative GCN5-related N-acetyltransferase (gene ID 123668) via the *A. nidulans gpdA* promoter or *lae1* via the *tef1* promoter resulted in an increased expression of cellulase encoding genes under inducing conditions (i.e. on lactose) (Häkkinen et al., 2014; Seiboth et al., 2012). More recently, Cre1 was shown to be involved in chromatin accessibility of *xyr1* promoter (Mello-de-Sousa et al., 2016).

As our knowledge about regulation of TF activities accumulates, various possibilities for rational design of industrial fungal strains emerge, such as combinations of different

approaches as already mentioned above, or the use of synthetic TFs. Recently, inducer-independent production of CAZymes by filamentous fungi was achieved via synthetic TFs. For instance, overexpression via *pdc1* promoter of a synthetic TF (*xyr1-cre1*$_b$) consisting of Xyr1 DNA-binding and effector domains, and Cre1 DNA-binding domain, resulted in inducer-independent production of cellulases and hemicellulases (i.e. xylanases) in the Cre1-negative *T. reesei* strain Rut-C30 (Zhang et al., 2017). Another example of synthetic TFs (CX^C-S) was shown in *P. oxalicum*, where the DNA-binding domain of the constitutively active $XlnR^{A871V}$ was replaced with that of ClrB. This synthetic TF was overexpressed using the *A. nidulans gpdA* promoter, yielding to inducer-independent, but still glucose-repressed, expression of cellulases and hemicellulases (i.e. xylanases) (Gao et al., 2017b).

To conclude, fungal strains with increased or constitutive production of plant biomass degrading enzymes can be rationally designed by tuning the transcriptional regulatory systems. Mutations that lead to constitutive expression of enzymes are difficult to predict and so far only identified via genetic screens. Once identified, these mutations can be successfully transferred to industrial strains or related species. In this review, we also highlighted that regulation of activities of orthologues TFs, or the set of genes regulated by orthologues TFs might be species-specific. It is therefore important that detailed studies on how TFs are activated are performed not only in a single species, which is subsequently used as a blue print to predict the activation mechanism in other fungal species, but performed in several representative fungal species across the filamentous fungi.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

EA and AR designed the content of the manuscript and collected literature. EA wrote the manuscript and AR critically commented on the manuscript.

ACKNOWLEDGEMENTS

We acknowledge the helpful comments of Dr. Jaap Visser and Prof. dr. Peter J. Punt on the manuscript.

REFERENCES

Adnan, M., Zheng, W., Islam, W., Arif, M., Abubakar, Y., Wang, Z. et al. (2017). Carbon catabolite repression in filamentous fungi. Int. J. Mol. Sci., 19, 48. doi: 10.3390/ijms19010048

Aghcheh, R.K., and Kubicek, C.P. (2015). Epigenetics as an emerging tool for improvement of fungal strains used in biotechnology. Appl. Microbiol. Biotechnol., 99, 6167–6181. doi: 10.1007/s00253-015-6763-2

Alazi, E., Khosravi, C., Homan, T.G., du Pré, S., Arentshorst, M., Di Falco, M. et al. (2017). The pathway intermediate 2-keto-3-deoxy-L-galactonate mediates the induction of genes involved in D-galacturonic acid utilization in *Aspergillus niger*. FEBS Lett., 591:10. doi: 10.1002/1873-3468.12654

Alazi, E., Knetsch, T., Di Falco, M., D. Reid, I., Arentshorst, M., Visser, J. et al. (2018). Inducer-independent production of pectinases in *Aspergillus niger* by overexpression of the D-galacturonic acid responsive transcription factor *gaaR*. Appl. Microbiol. Biotechnol. doi: 10.1007/s00253-018-8753-7

Alazi, E., Niu, J., Kowalczyk, J.E., Peng, M., Aguilar Pontes, M.V., van Kan, J.A.L. et al. (2016). The transcriptional activator GaaR of *Aspergillus niger* is required for release and utilization of D-galacturonic acid from pectin. FEBS Lett., 590:12, 1804-1815. doi: 10.1002/1873-3468.12211

Aro, N., Ilmén, M., Saloheimo, A., and Penttilä, M. (2003). ACEI of *Trichoderma reesei* is a repressor of cellulase and xylanase expression. Appl. Environ. Microbiol., 69:1, 56–65. doi: 10.1128/AEM.69.1.56-65.2003

Battaglia, E., Visser, L., Nijssen, A., van Veluw, G.J., Wösten, H.A.B., and de Vries, R.P. (2011). Analysis of regulation of pentose utilisation in *Aspergillus niger* reveals evolutionary adaptations in *Eurotiales*. Stud. Mycol., 69:1, 31–38. doi: 10.3114/sim.2011.69.03

Benocci, T., Aguilar-Pontes, M.V., Zhou, M., Seiboth, B., and de Vries, R.P. (2017). Regulators of plant biomass degradation in ascomycetous fungi. Biotechnol. Biofuels, 10, 1–25. doi: 10.1186/s13068-017-0841-x

Biggin, M.D. (2011). Animal transcription networks as highly connected, quantitative continua. Dev. Cell, 21, 611-626.

Brunner, K., Lichtenauer, A.M., Kratochwill, K., Delic, M., and Mach, R.L. (2007). Xyr1 regulates xylanase but not cellulase formation in the head blight fungus *Fusarium graminearum*. Curr. Genet., 52:5–6, 213–220. doi: 10.1007/s00294-007-0154-x

Caffall, K.H., and Mohnen, D. (2009). The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. Carbohydr. Res., 344:14, 1879–900. doi: 10.1016/j.carres.2009.05.021

Calero-Nieto, F., Di Pietro, A., Roncero, M.I.G., and Hera, C. (2007). Role of the transcriptional activator XlnR of *Fusarium oxysporum* in regulation of xylanase genes and virulence. Mol. Plant-Microbe Interact., 20:8, 977–985. doi: 10.1094/MPMI-20-8-0977

Cao, Y., Zheng, F., Wang, L., Zhao. G., Chen, G., Zhang, W. et al. (2017). Rce1, a novel transcriptional repressor, regulates cellulase gene expression by antagonizing the trans-activator Xyr1 in *Trichoderma reesei*. Mol. Microbiol.,105:1, 65-83. doi: 10.1111/mmi.13685

Chang, P.K., and Ehrlich, K.C. (2013). Genome-wide analysis of the $Zn(II)_2Cys_6$ zinc cluster-encoding gene family in *Aspergillus flavus*. Appl. Microbiol. Biotechnol., 97:10, 4289-4300. doi: 10.1007/s00253-013-4865-2

Chen, M., Qin, Y., Cao, Q., Liu, G., Li, J., Li, Z. et al. (2013). Promotion of extracellular lignocellulolytic enzymes production by restraining the intracellular β-glucosidase in *Penicillium decumbens*. Bioresour. Technol., 137, 33–40. doi: 10.1016/j.biortech.2013.03.099

Chua, G., Morris, Q.D., Sopko, R., Robinson, M.D., Ryan, O., Chan, E.T. et al. (2006). Identifying transcription factor functions and targets by phenotypic activation. Proc. Natl. Acad. Sci. U. S. A., 103:32, 12045–50. doi: 10.1073/pnas.0605140103

Coradetti, S.T., Craig, J.P., Xiong, Y., Shock, T., Tian, C., and Glass, N.L. (2012). Conserved and essential transcription factors for cellulase gene expression in ascomycete fungi. Proc. Natl. Acad. Sci., 109:19, 7397–7402. doi: 10.1073/pnas.1200785109

Coradetti, S.T., Xiong, Y., and Glass, N.L. (2013). Analysis of a conserved cellulase transcriptional regulator reveals inducer-independent production of cellulolytic enzymes in *Neurospora crassa*. MicrobiologyOpen, 2:4, 595–609. doi: 10.1002/mbo3.94

Craig, J.P., Coradetti, S.T., Starr, T.L., and Glass, N.L. (2015). Direct target network of the *Neurospora crassa* plant cell wall deconstruction regulators CLR-1, CLR-2, and XLR-1. mBio, 6:5, 1–11. doi: 10.1128/mBio.01452-15.Editor

Culleton, H., Mckie, V., and de Vries, R.P. (2013). Physiological and molecular aspects of degradation of plant polysaccharides by fungi: What have we learned from *Aspergillus*?. Biotechnol. J., 8:8, 884–894. doi: 10.1002/biot.201200382

Cziferszky, A., Mach, R.L., and Kubicek, C.P. (2002). Phosphorylation positively regulates DNA binding of the carbon catabolite repressor Cre1 of *Hypocrea jecorina* (*Trichoderma reesei*). J. Biol. Chem., 277:17, 14688-14694. doi: 10.1074/jbc.M200744200

de Vries, R.P., Riley, R., Wiebenga, A., Aguilar-Osorio, G., Amillis, S., Uchima, C.A. 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. doi: 10.1186/s13059-017-1151-0

Denton, J.A., and Kelly, J.M. (2011). Disruption of *Trichoderma reesei cre2*, encoding an ubiquitin C-terminal hydrolase, results in increased cellulase activity. BMC Biotechnol., 11, 1–10. doi: 10.1186/1472-6750-11-103

Derntl, C., Gudynaite-Savitch, L., Calixte, S., White, T., Mach, R.L., and Mach-Aigner, A.R. (2013). Mutation of the Xylanase regulator 1 causes a glucose blind hydrolase expressing phenotype in industrially used *Trichoderm*a strains. Biotechnol. Biofuels, 6, 1–11. doi: 10.1186/1754-6834-6-62

Derntl, C., Kluger, B., Bueschl, C., Schuhmacher, R., Mach, R.L., and Mach-Aigner, A.R. (2017). Transcription factor Xpp1 is a switch between primary and secondary fungal metabolism. Proc. Natl. Acad. Sci., 114:4, E560–E569. doi: 10.1073/pnas.1609348114

Derntl, C., Rassinger, A., Srebotnik, E., Mach, R.L., and Mach-Aigner, A.R. (2015). Xpp1 regulates the expression of xylanases, but not of cellulases in *Trichoderma reesei*. Biotechnol. Biofuels, 8:1, 1–11. doi: 10.1186/s13068-015-0298-8

Fang, X., Shen, Y., Zhao, J., Bao, X., and Qu, Y. (2010). Status and prospect of lignocellulosic bioethanol production in China. Bioresour. Technol., 101:13, 4814–4819. doi: 10.1016/j.biortech.2009.11.050

Fujii, T., Inoue, H., and Ishikawa, K. (2013). Enhancing cellulase and hemicellulase production by genetic modification of the carbon catabolite repressor gene, *cre*A, in *Acremonium cellulolyticus*. AMB Express, 3, 1–9. doi: 10.1186/2191-0855-3-73

Fujii, T., Inoue, H., and Ishikawa, K. (2014). Characterization of the xylanase regulator protein gene, *xlnR*, in *Talaromyces cellulolyticus* (formerly known as *Acremonium cellulolyticus*). Biosci. Biotechnol. Biochem., 78:9, 1564–1567. doi: 10.1080/09168451.2014.923298

Gao, L., Li, Z., Xia, C., Qu, Y., Liu, M., Yang, P. et al. (2017). Combining manipulation of transcription factors and overexpression of the target genes to enhance lignocellulolytic enzyme

production in *Penicillium oxalicum*. Biotechnol. Biofuels, 10, 1–16. doi: 10.1186/s13068-017-0783-3

Gao, L., Xia, C., Xu, J., Li, Z., Yu, L., Liu, G. et al. (2017b). Constitutive expression of chimeric transcription factors enables cellulase synthesis under non-inducing conditions in *Penicillium oxalicum*. Biotechnol. J., 12, 1–5. doi: 10.1002/biot.201700119

Gidley, M.J. (2001). "Starch structure/ function relationships: Achievements and challenges," in Starch: Advances in structure and function. ed. T.L. Barsby, A.M. Donald, and P.J. Frazier. (Cambridge, UK: The Royal Society of Chemistry), 1-7.

Gielkens, M.M.C., Dekkers, E., Visser, J., and de Graaff, L.H. (1999). Two cellobiohydrolase-encoding genes from *Aspergillus niger* require D -Xylose and the xylanolytic transcriptional activator XlnR for their expression. Appl. Environ. Microbiol., 65:10, 4340-4345.

Granek, J.A., and Clarke, N.D. (2005). Explicit equilibrium modeling of transcription-factor binding and gene regulation. Genome. Biol., 6: R87.

Gupta, V.K., Kubicek, C.P., Berrin, J.G., Wilson, D.W., Couturier, M., Berlin, A. et al. (2016) Fungal Enzymes for Bio-Products from Sustainable and Waste Biomass. Trends Biochem. Sci. 41:7, 633-645. doi: 10.1016/j.tibs.2016.04.006

Häkkinen, M., Valkonen, M.J., Westerholm-Parvinen, A., Aro, N., Arvas, M., Vitikainen, M. et al. (2014). Screening of candidate regulators for cellulase and hemicellulase production in *Trichoderma reesei* and identification of a factor essential for cellulase production. Biotechnol. Biofuels, 7:1. doi: 10.1186/1754-6834-7-14

Hasper, A.A. (2004) Function and mode of regulation of the transcriptional activator XlnR from *Aspergillus*. [dissertation]. [Wageningen (The Netherlands)]: Wageningen University

Hasper, A.A., Trindade, L.M., van der Veen, D., van Ooyen, A.J.J., and de Graaff L.H. (2004). Functional analysis of the transcriptional activator XlnR from *Aspergillus niger*. Microbiology, 150:2004, 1367–1375. doi: 10.1099/mic.0.26557-0

Herold, S., Bischof, R., Metz, B., Seiboth, B., and Kubicek, C.P. (2013). Xylanase gene transcription in *Trichoderma reesei* is triggered by different inducers representing different hemicellulosic pentose polymers. Eukaryot. Cell, 12:3, 390–398. doi: 10.1128/EC.00182-12

Ho, H., and Ho, K. (2015). Fungal strain improvement of *Aspergillus brasiliensis* for overproduction of xylanase in submerged fermentation through UV irradiation and chemicals mutagenesis. J. Adv. Biol. Biotechnol., 3:3, 117–131. doi: 10.9734/JABB/2015/17274

Huberman, L.B., Liu, J., Qin, L., and Glass, N.L. (2016). Regulation of the lignocellulolytic response in filamentous fungi. Fungal Biol. Rev., 30, 101–111. doi: 10.1016/j.fbr.2016.06.001

Ichinose, S., Tanaka, M., Shintani, T., and Gomi, K. (2014). Improved *α*-amylase production by *Aspergillus oryzae* after a double deletion of genes involved in carbon catabolite repression. Appl. Microbiol. Biotechnol., 98:1, 335–343. doi: 10.1007/s00253-013-5353-4

Ichinose, S., Tanaka, M., Shintani, T., and Gomi, K. (2017). Increased production of biomass-degrading enzymes by double deletion of *creA* and *creB* genes involved in carbon catabolite repression in *Aspergillus oryzae*. J. Biosci. Bioeng., 125:2, 141-147. doi: 10.1016/j.jbiosc.2017.08.019

Jiang, Y., Duarte, A.V., van den Brink, J., Wiebenga, A., Zou, G., Wang, C. et al. (2016). Enhancing saccharification of wheat straw by mixing enzymes from genetically-modified *Trichoderma reesei* and *Aspergillus niger*. Biotechnol. Lett., 38:1, 65–70. doi: 10.1007/s10529-015-1951-9

Kolpak, F.J., and Blackwell, J. (1976). Determination of the structure of cellulose II. Macromolecules, 9:2, 273–278. doi: 10.1021/ma60050a019

Kowalczyk, J.E., Benoit, I., and de Vries, R.P. (2014). Regulation of plant biomass utilization in *Aspergillus*. Adv. Appl. Microbiol., 88, 31-56. doi: 10.1016/B978-0-12-800260-5.00002-4

Kowalczyk, J.E., Lubbers, R.J.M., Peng, M., Battaglia, E., Visser, J., and de Vries, R.P. (2017). Combinatorial control of gene expression in *Aspergillus niger* grown on sugar beet pectin. Sci. Rep., 7:1, 1–12. doi: 10.1038/s41598-017-12362-y

Kunitake, E., and Kobayashi, T. (2017). Conservation and diversity of the regulators of cellulolytic enzyme genes in Ascomycete fungi. Curr. Genet., 63:6, 951–958. doi: 10.1007/s00294-017-0695-6

Lamb, H.K., Newton, G.H., Levett, L.J., Cairns, E., Roberts, C.F., and Hawkins, A.R. (1996). The QUTA activator and QUTR repressor proteins of *Aspergillus nidulans* interact to regulate transcription of the quinate utilization pathway genes. Microbiology, 142:1 996, 1477–1490.

Levett, L.J., Si-Hoe, S.M., Liddle, S., Wheeler, K., Smith, D., Lamb, H.K. et al. (2000) Identification of domains responsible for signal recognition and transduction within the QUTR transcription repressor protein. Biochem. J., 350, 189-197. doi: 10.1042/bj3500189

Li, J., Lin, L., Li, H., Tian, C., and Ma, Y. (2014). Transcriptional comparison of the filamentous fungus *Neurospora crassa* growing on three major monosaccharides D-glucose, D-xylose and L-arabinose. Biotechnol. Biofuels, 7:1, 1–15. doi: 10.1186/1754-6834-7-31

Li, Y., Zheng, X., Zhang, X., Bao, L., Zhu, Y., Qu, Y. et al. (2016). The different roles of *Penicillium oxalicum* LaeA in the production of extracellular cellulase and ß-xylosidase. Front. Microbiol., 7:DEC. doi: 10.3389/fmicb.2016.02091

Li, Z., Yao, G., Wu, R., Gao, L., Kan, Q., Liu, M. et al. (2015). Synergistic and dose-controlled regulation of cellulase gene expression in *Penicillium oxalicum*. PLoS Genetics, 11, 1–45. doi: 10.1371/journal.pgen.1005509

Lichius, A., Seidl-Seiboth, V., Seiboth, B., and Kubicek, C.P. (2014). Nucleo-cytoplasmic shuttling dynamics of the transcriptional regulators XYR1 and CRE1 under conditions of cellulase and xylanase gene expression in *Trichoderma reesei*. Mol. Microbiol., 94, 1162–1178. doi: 10.1111/mmi.12824

Liu, G., Zhang, L., Qin, Y., Zou, G., Li, Z., Yan, X. et al. (2013). Long-term strain improvements accumulate mutations in regulatory elements responsible for hyper-production of cellulolytic enzymes. Sci. Rep., 3, 1–7. doi: 10.1038/srep01569

Liu, R., Chen, L., Jiang, Y., Zou, G., and Zhou, Z. (2017). A novel transcription factor specifically regulates GH11 xylanase genes in *Trichoderma reesei*. Biotechnol. Biofuels, 10:194. doi: 10.1186/s13068-017-0878-x

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res., 42:D1, 490–495. doi: 10.1093/nar/gkt1178

Loqué, D., Scheller, H.V., and Pauly, M. (2015). Engineering of plant cell walls for enhanced biofuel production. Curr. Opin. Plant Biol., 25, 151–161. doi: 10.1016/j.pbi.2015.05.018

Lv, X., Zheng, F., Li, C., Zhang, W., Chen, G., and Liu, W. (2015). Characterization of a copper responsive promoter and its mediated overexpression of the xylanase regulator 1 results in an

induction-independent production of cellulases in *Trichoderma reesei*. Biotechnol. Biofuels, 8, 1–14. doi: 10.1186/s13068-015-0249-4

Mach-Aigner, A.R., Omony, J., Jovanovic, B., van Boxtel, A.J.B., and de Graaff, L.H. (2012). D-Xylose concentration-dependent hydrolase expression profiles and the function of CreA and XlnR in *Aspergillus niger*. Appl. Environ. Microbiol., 78:9, 3145-3155. doi: 10.1128/AEM.07772-11

Mach-Aigner, A.R., Pucher, M.E., Steiger, M.G., Bauer, G.E., Preis, S.J., and Mach, R.L. (2008). Transcriptional regulation of *xyr1*, encoding the main regulator of the xylanolytic and cellulolytic enzyme system in *Hypocrea jecorina*. Appl. Environ. Microbiol., 74:21, 6554-6562. doi: 10.1128/AEM.01143-08

MacPherson, S., Larochelle, M., and Turcotte, B. (2006). A fungal family of transcriptional regulators: the zinc cluster proteins. Microbiol. Mol. Biol. Rev., 70:3, 583–604. doi: 10.1128/MMBR.00015-06

Makita, T., Katsuyama, Y., Tani, S., Suzuki, H., Kato, N., Todd, R.B. et al. (2009). Inducer-dependent nuclear localization of a $Zn(II)_2Cys_6$ transcriptional activator, AmyR, in *Aspergillus nidulans*. Biosci. Biotechnol. Biochem., 73:2, 391–9. doi: 10.1271/bbb.80654

Marui, J., Kitamoto, N., Kato, M., Kobayashi, T., and Tsukagoshi, N. (2002). Transcriptional activator, AoXlnR, mediates cellulose-inductive expression of the xylanolytic and cellulolytic genes in *Aspergillus oryzae*. FEBS Lett., 528, 279–282. doi: 10.1016/S0014-5793(02)03328-8

Marui, J., Tanaka, A., Mimura, S., de Graaff, L.H., Visser, J., Kitamoto N. et al. (2002b) A transcriptional activator, AoXlnR, controls the expression of genes encoding xylanolytic enzymes in *Aspergillus oryzae*. Fungal Genet. Biol., 35, 157-169. doi:10.1006/fgbi.2001.1321

Mello-de-Sousa, T.M., Rassinger, A., Derntl, C., Poças-Fonseca, M.J., Mach, R.L., and Mach-Aigner, A.R. (2016). The relation between promoter chromatin status, Xyr1 and cellulase ex-pression in *Trichoderma reesei*. Curr. Genomics, 17:2, 145–52. doi: 10.2174/1389202917666151116211812

Meyer, V., Fiedler, M., Nitsche, B., and King, R. (2015). The cell factory *Aspergillus* enters the big data era : Opportunities and challenges for optimising product formation. Adv. Biochem. Eng. Biotechnol., 149, 91-132. doi: 10.1007/10_2014_297

Meyer, V., Ram, A.F.J., and Punt, P.J. (2010). "Genetics, genetic manipulation, and approaches to strain improvement of filamentous fungi ," in Manual of industrial microbiology and biotechnology. ed. R. Baltz, A. Demain, J. Davies, A. Bull, B. Junker, L. Katz et al. (Washington, DC: ASM Press), 318-329.

Montenecourt, B., and Eveleigh, D. (1979) "Selective screening methods for the isolation of high yielding cellulase mutants of *Trichoderma reesei*," in Hydrolysis of cellulose: mechanisms of enzymatic and acid catalysis. ed. R. Brown, and L. Jurasek. (Washington DC: American Chemical Society), 289-301.

Nakari-Setälä, T., Paloheimo, M., Kallio, J., Vehmaanperä, J., Penttilä, M., and Saloheimo, M. (2009). Genetic modification of carbon catabolite repression in *Trichoderma reesei* for improved protein production. Appl. Environ. Microbiol., 75:14, 4853–4860. doi: 10.1128/AEM.00282-09

Niu, J., Alazi, E., Reid, I.D., Arentshorst, M., Punt, P.J., Visser, J. et al. (2017). An evolutionarily conserved transcriptional activator-repressor module controls expression of genes for D-galacturonic acid utilization in *Aspergillus niger*. Genetics, 205:1. doi: 10.1534/genetics.116.194050

Niu, J., Homan, T.G., Arentshorst, M., de Vries, R.P., Visser, J., and Ram, A.F.J. (2015). The interaction of induction and repression mechanisms in the regulation of galacturonic acid-induced genes in *Aspergillus niger*. Fungal Genet. Biol., 82, 32–42. doi: 10.1016/j.fgb.2015.06.006

Noguchi, Y., Sano, M., Kanamaru, K., Ko, T., Takeuchi, M., Kato, M. et al. (2009). Genes regulated by AoXlnR, the xylanolytic and cellulolytic transcriptional regulator, in *Aspergillus oryzae*. Appl. Microbiol. Biotechnol., 85, 141–154. doi: 10.1007/s00253-009-2236-9

Noguchi, Y., Tanaka, H., Kanamaru, K., Kato, M., and Kobayashi, T. (2011). Xylose triggers reversible phosphorylation of XlnR, the fungal transcriptional activator of xylanolytic and cellulolytic genes in *Aspergillus oryzae*. Biosci. Biotechnol. Biochem., 75:5, 953-959. doi: 10.1271/bbb.100923

Ogawa, M., Kobayashi, T., and Koyama, Y. (2012). ManR, a novel $Zn(II)_2Cys_6$ transcriptional activator, controls the β-mannan utilization system in *Aspergillus oryzae*. Fungal Genet. Biol., 49:12, 987–95. doi: 10.1016/j.fgb.2012.09.006

Ogawa, M., Kobayashi, T., and Koyama, Y. (2013). ManR, a transcriptional regulator of the β-mannan utilization system, controls the cellulose utilization system in *Aspergillus oryzae*. Biosci. Biotechnol. Biochem., 77:2. 426–429. doi: 10.1271/bbb.120795

Okuda, N., Fujii, T., Inoue, H., Ishikawa, K., and Hoshino, T. (2016). Enhancing cellulase production by overexpression of xylanase regulator protein gene, *xlnR*, in *Talaromyces cellulolyticus* cellulase hyperproducing mutant strain. Biosci. Biotechnol. Biochem., 80:10, 2065–2068. doi: 10.1080/09168451.2016.1189315

Pel, H.J., de Winde, J.H., Archer, D.B., Dyer, P.S., Hofmann, G., Schaap, P.J. et al. (2007) Genome sequencing and analysis of the versatile cell factory Aspergillus niger CBS 513.88. Nat.

Portnoy, T., Margeot, A., Seidl-Seiboth, V., Le Crom, S., Chaabane, F.B., Linke, R. et al. (2011). Differential regulation of the cellulase transcription factors XYR1, ACE2, and ACE1 in *Trichoderma reesei* strains producing high and low levels of cellulase. Eukaryot. Cell, 10:2, 262–271. doi: 10.1128/EC.00208-10

Prathumpai, W., McIntyre, M., and Nielsen, J. (2004). The effect of CreA in glucose and xylose catabolism in *Aspergillus nidulans*. Appl. Microbiol. Biotechnol., 63:6, 748–753. doi: 10.1007/s00253-003-1409-1

Rauscher, R., Wurleitner, E., Wacenovsky, C., Aro, N., Stricker, A.R., Zeilinger, S., et al. (2006). Transcriptional regulation of *xyn1*, encoding xylanase I, in *Hypocrea jecorina*. Eukaryot. Cell, 5:447-456.

Ries, L.N.A., Beattie, S.R., Espeso, E.A., Cramer, R.A., and Goldman, G.H. (2016). Diverse regulation of the CreA carbon catabolite repressor in *Aspergillus nidulans*. Genetics, 203:1, 335-352. doi: 10.1534/genetics.116.187872

Ritsema, T., and Smeekens, S. (2003). Fructans: Beneficial for plants and humans. Curr. Opin. Plant Biol., 6:3, 223–230. doi: 10.1016/S1369-5266(03)00034-7

Scheller, H.V., and Ulvskov, P. (2010). Hemicelluloses. Annu. Rev. Plant Biol., 61:1, 263–289. doi: 10.1146/annurev-arplant-042809-112315

Seiboth, B., Herold, S., and Kubicek, C.P. (2012). "Metabolic engineering of inducer formation for cellulase and hemicellulase gene expression in *Trichoderma reesei*," in Reprogramming microbial metabolic pathways, Volume 64 of Subcellular Biochemistry. ed. X. Wang, J. Chen, P.J. Quinn. (Dordrecht : Springer Netherlands), 367-390.

Seiboth, B., Karimi, R.A., Phatale, P.A., Linke, R., Hartl, L., Sauer, D.G. et al. (2012). The putative protein methyltransferase LAE1 controls cellulase gene expression in *Trichoderma reesei*. Mol. Microbiol., 84:6, 1150–1164. doi: 10.1111/j.1365-2958.2012.08083.x

Stricker, A.R., Grosstessner-Hain, K., Würleitner, E., and Mach, R.L. (2006). Xyr1 (Xylanase Regulator 1) regulates both the hydrolytic enzyme system and D-xylose metabolism in *Hypocrea jecorina*. Eukaryot. Cell, 5:12, 2128–2137. doi: 10.1128/EC.00211-06

Sun, J., and Glass, N.L. (2011). Identification of the CRE-1 cellulolytic regulon in *Neurospora crassa*. PLoS ONE, 6:9. doi: 10.1371/journal.pone.0025654

Sun, J., Tian, C., Diamond, S., and Glass, N.L. (2012). Deciphering transcriptional regulatory mechanisms associated with hemicellulose degradation in *Neurospora crassa*. Eukaryot. Cell, 11:4, 482–493. doi: 10.1128/EC.05327-11

Suzuki, K., Tanaka, M., Konno, Y., Ichikawa, T., Ichinose, S., Hasegawa-Shiro, S. 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.

Sweeney, M.D., and Xu, F. (2012). Biomass Converting Enzymes as Industrial Biocatalysts for Fuels and Chemicals: Recent Developments. Catalysts, 2:2, 244–263. doi: 10.3390/catal2020244

Tamayo, E.N., Villanueva, A., Hasper, A.A., de Graaff, L.H., Ramón, D., and Orejas, M. (2008). CreA mediates repression of the regulatory gene *xlnR* which controls the production of xylanolytic enzymes in *Aspergillus nidulans*. Fungal Genet. Biol., 45:6, 984–993. doi: 10.1016/j.fgb.2008.03.002

Tani, S., Katsuyama, Y., Hayashi, T., Suzuki, H., Kato, M., Gomi, K. et al. (2001). Characterization of the *amyR* gene encoding a transcriptional activator for the amylase genes in *Aspergillus nidulans*. Curr. Genet., 39:1, 10–15. doi: 10.1007/s002940000175

Tani, S., Kawaguchi, T., and Kobayashi, T. (2014). Complex regulation of hydrolytic enzyme genes for cellulosic biomass degradation in filamentous fungi. Appl. Microbiol. Biotechnol., 98:11, 4829-4837. doi: 10.1007/s00253-014-5707-6

Thieme, N., Wu, V.W., Dietschmann, A., Salamov, A.A., Wang, M., Johnson, J. et al. (2017). The transcription factor PDR-1 is a multi-functional regulator and key component of pectin deconstruction and catabolism in *Neurospora crassa*. Biotechnol. Biofuels, 10:1, 1–21. doi: 10.1186/s13068-017-0807-z

van Hanh, V., Pham, T.A., and Kim, K. (2010). Improvement of a fungal strain by repeated and sequential mutagenesis and optimization of solid-state fermentation for the hyper-production of raw-starch-digesting enzyme. J. Microbiol. Biotechnol., 20, 718–726. doi: 10.4014/jmb.0908.08016

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

vanKuyk, P.A., Benen, J.A.E., Wösten, H.A.B., Visser, J., and de Vries, R.P. (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., 93:1, 285–93. doi: 10.1007/s00253-011-3550-6

Wagh, S.A. (2010). Bioconversion of tannic acid to gallic acid by using fungal tannase. [dissertation]. [Pune (India)]: National Chemical Laboratory

Wang J., Wu, Y., Gong Y., Yu S., and Liu G. (2015). Enhancing xylanase production in the thermophilic fungus *Myceliophthora thermophila* by homologous overexpression of *Mtxyr1*. J. Ind. Microbiol. Biotechnol., 42, 1233–1241. doi: 10.1007/s10295-015-1628-3

Wang, J., Gong, Y., Zhao, S., and Liu, G. (2018). A new regulator of cellulase and xylanase in the thermophilic fungus *Myceliophthora thermophila* strain ATCC 42464. 3 Biotech, 8:160. doi: 10.1007/s13205-017-1069-y

Wang, S., Liu, G., Wang, J., Yu, J., Huang, B., and Xing, M. (2013). Enhancing cellulase production in *Trichoderma reesei* RUT C30 through combined manipulation of activating and repressing genes. J. Ind. Microbiol. Biotechnol., 40, 633–641. doi: 10.1007/s10295-013-1253-y

Xin, Q., Gong, Y., Lv, X., Chen, G., and Liu, W. (2013). *Trichoderma reesei* histone acetyltransferase Gcn5 regulates fungal growth, conidiation, and cellulase gene expression. Curr. Microbiol., 67:5, 580–589. doi: 10.1007/s00284-013-0396-4

Yang, F., Gong, Y.F., Liu, G., Zhao, S.M., and Wang, J. (2015) Enhancing cellulase production in thermophilic fungus *Myceliophthora thermophila* ATCC42464 by RNA interference of *cre1* gene expression. J. Microbiol. Biotechnol., 25:7, 1101–1107. doi: 10.4014/jmb.1501.01049

Yao, G., Li, Z., Gao, L., Wu, R., Kan, Q., Liu, G. et al. (2015). Redesigning the regulatory pathway to enhance cellulase production in *Penicillium oxalicum*. Biotechnol. Biofuels, 8:71. doi: 10.1186/s13068-015-0253-8

Zeilinger, S., Mach, R.L., and Kubicek, C.P. (1998). Two adjacent protein binding motifs in the *cbh2* (Cellobiohydrolase II-encoding) promoter of the fungus *Hypocrea jecorina* (*Trichoderma reesei*) cooperate in the induction by cellulose. J. Biol. Chem., 273, 34463-34471. doi: 10.1074/jbc.273.51.34463

Zhang, L., Lubbers, R.J.M., Simon, A., Stassen, J.H.M., Vargas Ribera, P.R., Viaud, M. et al. (2016). A novel Zn_2Cys_6 transcription factor BcGaaR regulates D-galacturonic acid utilization in *Botrytis cinerea*. Mol. Microbiol., 100:2, 247-262. doi: 10.1002/acr.22212

Zhang, X., Li, Y., Zhao, X., and Bai, F. (2017). Constitutive cellulase production from glucose using the recombinant *Trichoderma reesei* strain overexpressing an artificial transcription activator. Bioresour. Technol., 223, 317–322. doi: 10.1016/j.biortech.2016.10.083

Znameroski, E.A., Coradetti, S.T., Roche, C.M., Tsai, J.C., and Iavarone, A.T. (2012). Induction of lignocellulose-degrading enzymes in *Neurospora crassa* by cellodextrins. PNAS, 109:16, 6012-6017. doi: 10.1073/pnas.1118440109

CHAPTER 3

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

Alazi E, Niu J, Kowalczyk JE, Peng M, Aguilar Pontes MV, van Kan JAL, Visser J, de Vries RP, Ram AFJ

Published in *FEBS Letters* (2016) 590:1804-1815

ABSTRACT

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

INTRODUCTION

Pectins are complex heterogeneous polysaccharides found in plant cell walls. Four substructures of pectin have been identified, and include polygalacturonic acid (PGA) also known as homogalacturonan, xylogalacturonan (XGA), rhamnogalacturonan I (RG-I), and rhamnogalacturonan II (RG-II) [1]. The backbones of PGA, XGA, and RG-II are made up of *α*-1,4-linked D-galacturonic acid (GA) residues. PGA, a linear polymer of GA, is the most abundant polysaccharide present in pectin [1]. 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 [1,2]. Side chains of RG-II contain at least 12 different types of monosaccharides, whereas the side chains of RG-I are mainly arabinan and arabinogalactan comprising of L-arabinose and D-galactose residues [1].

In nature, pectin is an important carbon source for many saprotrophic fungi such as *Aspergillus niger*. Previous studies demonstrated that *A. niger* can produce more pectin-degrading enzymes than other more specialized fungi such as *Podospora anserina* or *Neurospora crassa* [3–5]. GA is the main product of pectin degradation. In *A. niger*, GA is transported into the cell by a GA-induced sugar transporter named GatA [6]. GA is then catabolized into pyruvate and glycerol [7], through a pathway consisting of four enzymes: GaaA, D-galacturonate reductase; GaaB, L-galactonate dehydratase; GaaC, 2-keto-3-deoxy-L-galactonate aldolase; and GaaD, L-glyceraldehyde reductase [7]. Deletion of *gaaA*, *gaaB*, or *gaaC* abolished growth on GA as the sole carbon source [8–10]. *gaaD*, also known as the L-arabinose reductase gene, *larA*, is involved in the L-arabinose catabolic pathway, and the *ΔlarA* strain showed a reduced growth on L-arabinose as the sole carbon source [11].

The production of extra- and intracellular enzymes in *A. niger* is regulated by a network of transcription factors (TFs) [12]. Small sugar molecules (mono- and disaccharides) act as inducers and stimulate TFs which can bind to conserved motifs in the promoters of their target genes and activate or repress their expression. Expression of pectinase genes is highly controlled and depends on both induction and carbon catabolite repression [13,14]. Induction of the genes required for pectin degradation, GA transport, and GA catabolism requires the presence of GA, and it has been shown that GA or a derivative of GA induces the expression of pectinase genes [9,10,13].

Coordination of the induction of genes encoding extracellular enzymes and sugar uptake systems in fungi are often mediated by Zn_2Cys_6 TFs that bind to conserved promoter elements in the coregulated genes [12,15,16]. TFs inducing the genes required for the utilization of L-rhamnose (RhaR), arabinan/L-arabinose (AraR), xylan/D-xylose (XlnR), D-galactose (GalX), and cellulose (XlnR, ClrA, and ClrB) have been identified in *A. niger* [17–21]. Although L-rhamnose, L-arabinose, D-xylose, and D-galactose are also present in complex pectins, knock out mutants in these TFs display no signs of reduced growth on pectin [17,18,20], suggesting that the utilization of GA, the main component of this substrate, is not affected.

Martens-Uzunova and Schaap [7] have previously identified a set of GA-induced genes in *A. niger*, containing several pectinases (*pgaX*, *pgxA*, *pgxB*, *pgxC*, *paeA*, *pelA*, and *abfC*), sugar transporter-encoding genes (*gatA*, An03g01620, and An07g00780) and the GA

catabolic pathway genes (*gaaA-D*). These genes were suggested as the GA-regulon and contain a common GA-responsive element (GARE) in their promoter regions. The consensus element was defined as CCNCCAA [7]. Deletion and mutational analysis of GARE showed that the element is required for GA-induced gene expression in both *A. niger* and *Botrytis cinerea* [14,22]. A yeast one-hybrid study using a GA-responsive promoter in *B. cinerea* recently identified a novel Zn_2Cys_6 TF (BcGaaR) required for GA utilization [22]. In this study, the GA-responsive transcriptional activator, GaaR, of *A. niger* was identified by homology to BcGaaR. Deletion analysis and transcriptomic profiling studies performed in this study showed that the *A. niger* GaaR ortholog is required for growth on GA and PGA and for the induction of the GA-regulon when grown on sugar beet pectin (SBP).

MATERIALS AND METHODS

Strains, media, and growth conditions

A. niger strains MA234.1 (*cspA1*, *kusA::DR-amdS-DR*) and N593.20 (*cspA1*, *pyrG*$^-$, *kusA::amdS*) were used to create the *ΔgaaR* strains. N593.20 was made by transformation of N593 [23] with a deletion construct (*kusA::amdS*) [24] resulting in the deletion of *kusA*. Strain FP-1132.1 (*cspA1*, *pyrG::AOpyrG*, *kusA::amdS*) was obtained by transformation of N593.20 with *pyrG* from *Aspergillus oryzae*. MA234.1 was obtained by transformation of MA169.4 (*kusA*$^-$, *pyrG*$^-$) [25] with a 3.8 kb *Xba*I fragment containing the *A. niger pyrG* gene, resulting in the full restoration of the *pyrG* locus.

Complementation studies were performed with JN35.1 (*cspA1*, *kusA::DR-amdS-DR*, *gaaR::hygB*). To restore functionality of the *kusA* gene to allow ectopic integration of the complementing fragment, the *amdS* marker was looped out of JN35.1 by FAA counter selection as described [26] to give JN36.1. The *gaaR*-complemented strain, JN37.4, was created using JN36.1, by transformation of the *gaaR* gene including promoter and terminator regions (see below). All strains used are listed in Table S1.

Media were prepared as described [26]. For growth phenotype analyses, strains were grown on minimal medium (MM) with 1.5% (w/v) agar and various sole carbon sources: 25 or 50 mM glucose (VWR International, Amsterdam, the Netherlands), GA (Chemodex, St Gallen, Switzerland), L-rhamnose (Fluka, Zwijndrecht, the Netherlands), L-arabinose (Sigma-Aldrich, Zwijndrecht, the Netherlands), or D-xylose (Merck, Amsterdam, the Netherlands), and 1% (w/v) PGA (Sigma, Zwijndrecht, the Netherlands), SBP (Pectin Betapec RU301 Herbstreith & Fox KG, Neuenbürg, Germany), citrus pectin (CP) (Acros Organics, Geel, Belgium), or apple pectin (AP) (Pectin Classic AU2022 Herbstreith & Fox KG). pH was adjusted to 5.8 with NaOH or HCl buffer. The plates were inoculated with 2 µL 0.9 % (w/v) NaCl solution containing 1000 freshly harvested spores and cultivated at 30 °C for 4 days. For gene expression analyses, freshly harvested spores were inoculated with a final concentration of 10^6 spores mL^{-1} in 100 mL complete medium (CM) (pH 5.8) with 2% (w/v) D-fructose (Sigma-Aldrich) and were pregrown for 16 h. For northern blot analysis, mycelium was harvested by filtration through sterile myracloth, washed twice with MM with no carbon sources (pH 4.5), and 1.5 g (wet weight) mycelium was transferred and grown in 50 mL MM (pH 4.5) with 50 mM GA or 50 mM D-fructose for 2, 4, and 6 h. For

RNA-seq analysis, 2.5 g of pregrown mycelia were transferred to 50 mL MM (pH 4.5) with 25 mM GA and incubated for 2 h or to 50 mL MM with 1% SBP and incubated for 2, 8, or 24 h. All incubations were performed in rotary shaker at 30 °C and 250 r.p.m.

Construction of gene deletion and complementation strains

Protoplast-mediated transformation of *A. niger*, purification of the transformants, and genomic DNA extraction were performed as described [26]. To construct the deletion cassettes, 5' and 3' flanks of the *gaaR* gene were PCR-amplified using the primer pairs listed in Table S2 and N402 genomic DNA as template. To create JN35.1 strain, the split marker fragments with *hygB* selection were created using fusion PCR [27] and transformed to MA234.1. To create FP-1126.1 strain, the flanking regions were fused with a fragment containing the *A. oryzae pyrG* gene using GoTaq Long polymerase (Promega, Leiden, the Netherlands) and transformed into N593.20 strain. Parental strains and *gaaR* deletion mutants were deposited at the Centraal Bureau Schimmelcultures (CBS) under accession numbers indicated in Table S1. To complement the *gaaR* gene, the *gaaR* gene together with its 5' and 3' flanks was PCR-amplified using the primer pairs listed in Table S2, ligated into pJET1.2/ blunt cloning vector (Fermentas, Landsmeer, the Netherlands), amplified in the *E. coli* strain DH5*α* and transformed in to strain JN36.1 together with plasmid pMA357. pMA357 contains the *A. nidulans amdS* gene, cloned behind the *A. nidulans gdpA* promoter (Mark Arentshorst, unpublished vector). Deletion and complementation of *gaaR* were confirmed via Southern blot analysis or diagnostic PCR.

Gene expression analysis

For northern blot analysis, strains MA234.1 (reference strain) and JN35.1 (*ΔgaaR*) were pregrown in CM with D-fructose. At the time of transfer (t = 0) and 2, 4, and 6 h after the transfer to MM with GA or D-fructose, mycelium was harvested from cultures by filtration through sterile myracloth and frozen immediately in liquid nitrogen. Mycelium samples were stored at 80 °C. Total RNA was extracted from frozen mycelium samples after grinding in liquid nitrogen, using NucleoSpin RNA Kit (Macherey-Nagel, Düren, Germany) following the protocol provided by the supplier, including the rDNase treatment. Total RNA samples were stored at 80 °C. Quantitation and purity assessment of total RNA was done by spectrophotometric method (Nano-Drop 2000; Thermo Scientific, Breda, the Netherlands). Standard molecular techniques were applied as described [28]. About 3.5 µg RNA was loaded per sample and hybridized with [*α*-32P]-dCTP labeled probes after blotting (Deca-Label DNA Labelling Kit; Thermo Scientific). Probes were PCR-amplified using the N402 genomic DNA and the primer pairs are listed in Table S2. For RNA-seq analysis, the mycelium of FP-1132.1 (reference strain) and FP-1126.1 (*ΔgaaR*) was ground in Tissue Lyser II (Qiagen, Venlo, the Netherlands) and RNA was extracted using TRIzol reagent (Invitrogen, Breda, the Netherlands) and purified with NucleoSpin RNA Clean-up kit (Macherey-Nagel) with rDNase treatment. RNA quantity of the samples was checked with a NanoDrop-1000 spectrophotometer and the quality by RNA gel electrophoresis. Single-read samples were sequenced using Illumina HiSeqTM 2000 platform (http://illumina.com). Purification of mRNA, synthesis of cDNA library, and sequencing

reactions were conducted in the BGI Tech Solutions Co., Ltd. (Hong Kong, China). Transfer experiments and subsequent RNA-sequencing were performed in duplicates.

Bioinformatics

Raw reads were produced from the original image data by base calling. On average, ~ 13 million read of 51 bp per sample were obtained. After data filtering, the adaptor sequences, highly 'N' containing reads (> 10% of unknown bases), and low-quality reads (more than 50% bases with quality value of < 5%) were removed. After data filtering, in average, ~ 97.5% clean reads remained in each sample. Clean reads were then mapped to the genome of *Aspergillus niger* NRRL3 (http://genome.jgi.doe.gov/Aspni_NRRL3_1) using BOWTIE2 [29] and BWA software [30]. In average, 63.8% total mapped reads to the genome was achieved. The gene expression level was measured in 'fragments per kilobase of exon model per million mapped reads' (FPKM) [31] using RSEM tool [32]. Genes with expression value lower than 14 were considered low-expressed (approximately bottom 50%) and differential expression was identified by Student's t-test with a *P*-value cutoff 0.05. The RNAseq data have been submitted to Gene Expression Omnibus (GEO) [33] with accession number: GSE80227. Homology searches were performed using the BLASTP algorithm from NCBI against the nonredundant database and proteins with an E-value ≤ 1E-50 were defined as homologous [34]. Hierarchical clusters using the average expression values of genes were made via GENESIS 1.7.7 [35] with Pearson correlation and complete linkage. Low-expressed pectinases in all conditions were not included.

RESULTS AND DISCUSSION

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

A putative *A. niger* GA-responsive transcriptional activator was identified by homology to the recently identified *B. cinerea* Zn_2Cys_6 TF (BcGaaR) [22]. The *A. niger* ortholog (named GaaR) is a 740-amino acid-long protein encoded by *gaaR* (An04g00780/NRRL3_08195) and the bidirectional best BLAST hit of the 817-amino acid-long BcGaaR (Bcin09g00170). Analysis of the presence of GaaR among 20 *Aspergillus* species using the *Aspergillus* genome database (http://www.aspgd.org/) revealed that all Aspergilli, except *Aspergillus glaucus* contain a GaaR ortholog in their genome (data not shown). Interestingly, *A. glaucus* is not able to grow on GA as the sole carbon source (http://www.funggrowth.org), indicating the requirement of GaaR for GA utilization. *A. niger* GaaR and BcGaaR show 50.3% identity on the amino acid level throughout the entire protein sequence (Fig. S1). GaaR contains a typical Zn_2Cys_6 DNA-binding domain with the pattern of CX2CX6CX6CX2CX6C close to its NH_2-terminal end (residues 26–56) and a fungal-specific TF domain (residues 139–518). Amino acid alignment and phylogenetic analysis of GaaR revealed no significant similarity (an E-value cutoff < 1E-50) of GaaR to other TFs involved in plant cell wall utilization such as XlnR, AraR, RhaR, GalX, ClrA, and ClrB or to any other TF in *A. niger* (data not shown).

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

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

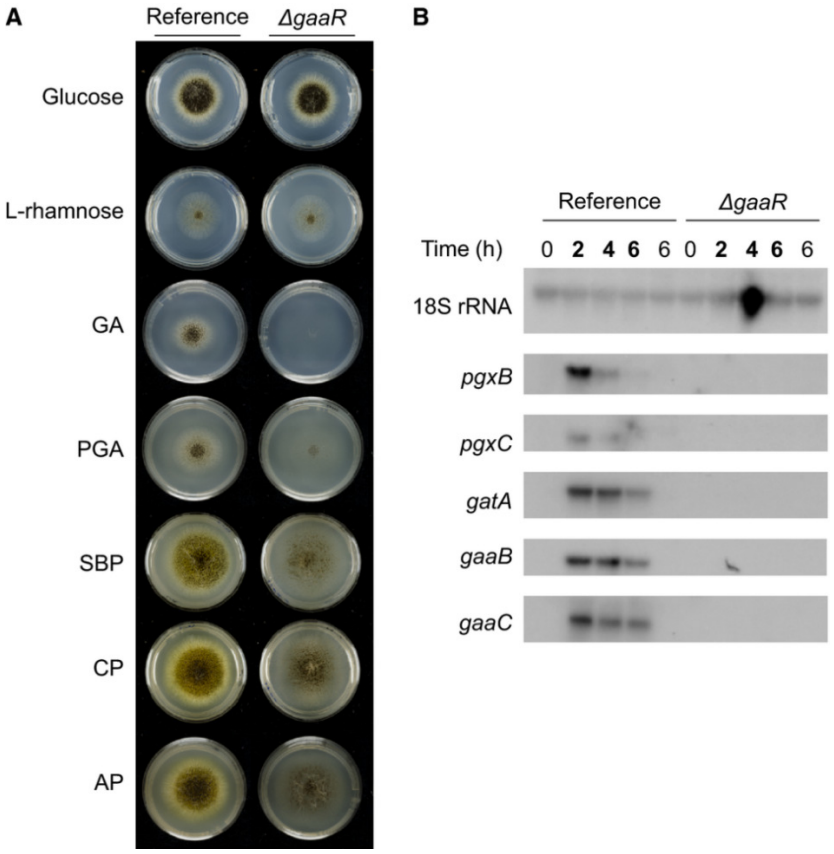

Fig. 1. Phenotypic and gene expression analyses of *A. niger ΔgaaR* **(A)** Growth profile of the reference strain (FP-1132.1) and *ΔgaaR* (FP-1126.1) on MM with 25 mM monomeric and 1% polymeric carbon sources. Strains were grown for 4 days at 30 °C. **(B)** Northern blot analysis of selected GA-induced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

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

The presence of GA has been shown to induce genes involved in PGA degradation (e.g., *pgxB*, *pgxC*), GA transport (*gatA*), and catabolism (*gaaA-D*) [7,14]. As a first indication for the involvement of GaaR in the induction of a subset of these genes on GA, a northern blot analysis was performed. The reference strain and *ΔgaaR* that was made the AB4.1 background were pregrown in D-fructose medium and transferred to either GA or D-fructose medium. For the reference strain, transfer of mycelium to GA resulted in a rapid induction of *pgxB*, *pgxC*, *gatA*, *gaaB*, and *gaaC*, whereas this induction was not observed in *ΔgaaR* (Fig. 1B).

To analyze the expression of a larger number of genes involved in pectin degradation, GA transport and catabolism, a genome-wide gene expression analysis was performed using RNA-seq. The reference strain and *ΔgaaR* in the N593 background were again pregrown in D-fructose medium and transferred to GA medium. RNA-seq analysis indicated that the GA induced expression of all genes that were previously identified as part of the GA-regulon [7] is dependent on GaaR (Table 1 and Fig. 2). The only exception is a putative GA transporter (An03g01620) that is expressed more than threefold less in *ΔgaaR* for which the *P*-value did not pass our significance level (0.05). In general, these observations show that the genes in the suggested GA-regulon [7] showed a significant reduction in *ΔgaaR* compared to the reference strain on GA (Table 1) and that GaaR is required for the induction of those genes.

To identify additional pectinase genes controlled directly or indirectly by GaaR, the expression of all 58 pectinolytic genes [3] was examined (Table S3). An overview of the gene abbreviations and their (putative) function is given in Martens-Uzunova and Schaap [3]. This analysis resulted in the identification of several additional pectinase genes for which the expression on GA is dependent of GaaR (Table 1 and Fig. 2, Fig. S3). This difference could be caused by higher sensitivity of the RNA-seq analysis compared to the previously used Affymetrix microarrays. In general, these newly identified genes were lower expressed compared to the genes in the GA-regulon described previously [7]. The gene encoding the putative pectin methylesterase C (*pmeC*) was missing on the Affymetrix chips, and therefore missed previously, but the RNA-seq study clearly indicated that induction of *pmeC* on GA is GaaR dependent. Inspection of the promoter regions of the newly identified members of the GA-regulon indicated the presence of putative GaaR-binding sites in the promoter regions of most genes (Table 1), enabling us to expand the GA-regulon to a larger set of genes.

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

Both the strongly reduced growth phenotype on GA and PGA and the expression analysis in *ΔgaaR* suggest that that GaaR is required for GA utilization in *A. niger*. Growth and sporulation of *ΔgaaR* on complex pectins such as SBP was also reduced, but not as severe as on GA and PGA (Fig. 1A). This could be explained by two (not mutually exclusive) hypotheses. The first explanation could be that *A. niger* has alternative mechanisms (independent of GaaR) to induce genes involved in GA utilization. The second possibility is

Table 1. RNA-seq analysis on GA of the genes that depend on GaaR for induction. Expression values (FPKM) are averages of duplicates. Fold changes ≥2 and *P*-values ≤0.05 are highlighted. GARE position is given with respect to the transcription start site.

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

Table 1. RNA-seq analysis on GA of the genes that depend on GaaR for induction. Expression values (FPKM) are averages of duplicates. Fold changes ≥2 and *P*-values ≤0.05 are highlighted. GARE position is given with respect to the transcription start site. (Continued)

NRRL3 protein ID	CBS 513.88 gene ID	Gene name	Ref GA 2h	*ΔgaaR* GA 2h	Fold change Ref/*ΔgaaR* GA 2h	*P*-value	GARE (CCNCCAA) position
NRRL3_10559	An18g04810	*rgxC* (putative)	20.00	0.90	22.22	1.26E-02	+ strand -880 and -852 and - strand -250
NRRL3_00684	An14g01130	*rglA*	5.77	1.03	5.63	9.23E-03	- strand -188
NRRL3_01606	An01g00330	*abfA*[c]	87.96	59.62	1.48	5.81E-01	
NRRL3_10865	An08g01710	*abfC* (putative)[a]	201.62	67.16	3.00	4.04E-02	
NRRL3_07094	An16g02730	*abnD* (putative)	4.57	1.53	2.99	2.66E-02	+ strand -246
NRRL3_02479	An01g10350	*lacB* (putative)	137.63	41.24	3.34	1.36E-02	
NRRL3_11738	An06g00290	*lacC* (putative)	28.91	9.08	3.19	3.55E-02	+ strand -267
NRRL3_10643	An18g05940	*galA*	105.64	29.24	3.61	3.53E-02	+ strand -307

[a] Genes identified as the GA-regulon by Martens-Uzunova and Schaap [7]. [b] *pmeC* not present on the Affymetrix microarray. [c] Genes not significantly differentially expressed on GA, but differentially expressed on SBP (see Table 2).

that additional sugars such as L-arabinose, D-galactose, D-xylose, or L-rhamnose that are present in SBP are metabolized and used for growth. To gain insight in the expression of pectinase genes in *ΔgaaR* on complex pectin, the reference strain and *ΔgaaR* were transferred from D-fructose to SBP and grown for 2, 8, and 24 h before harvesting mycelia and extraction of RNA.

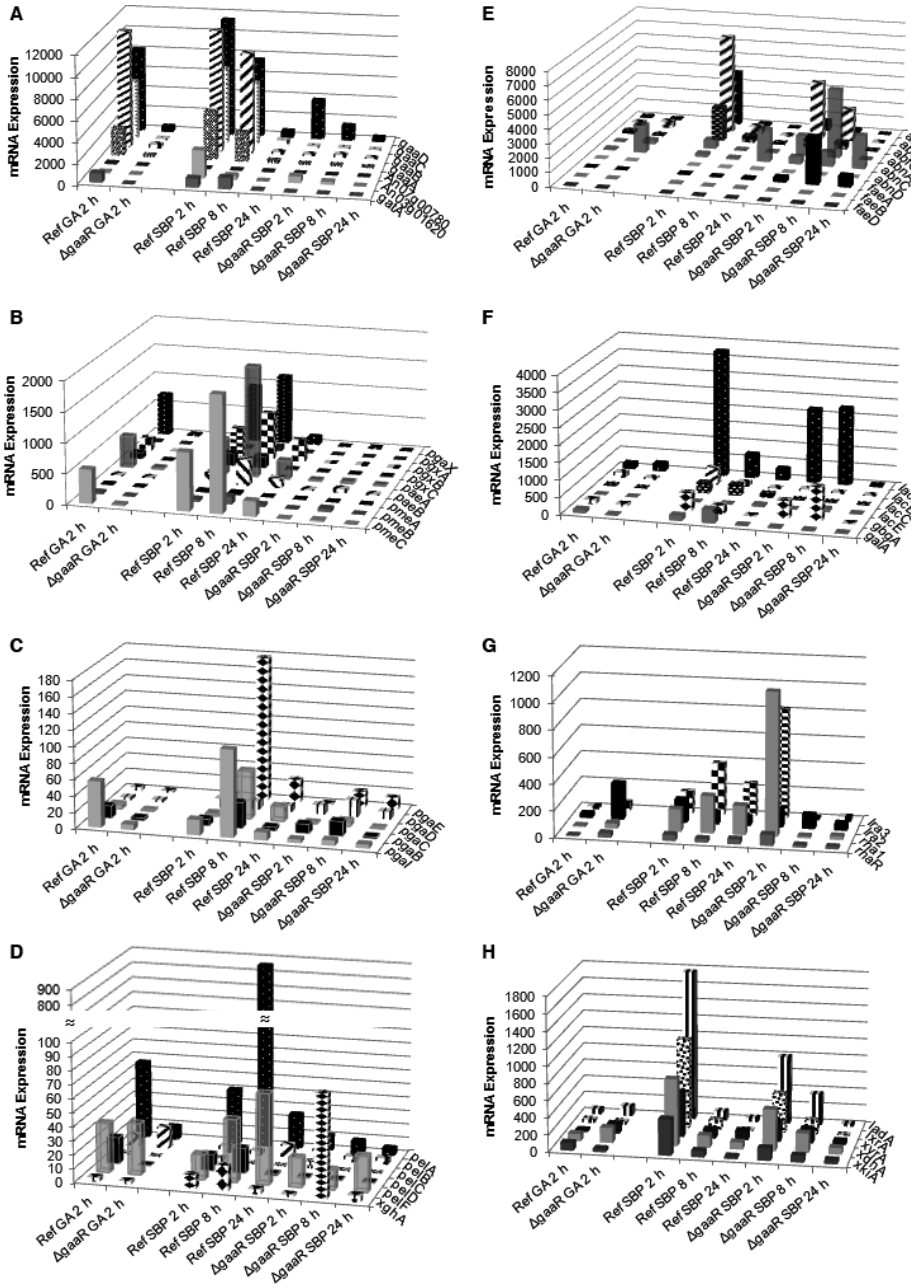

Fig. 2. Transcript levels of pectin utilization genes in *A. niger* reference and *ΔgaaR* on GA or SBP. **(A)** GA transporters and GA catabolic pathway enzymes; **(B)** exo-polygalacturonases and pectin acetyl- and methylesterases; **(C)** endo-polygalacturonases; **(D)** pectin lyases and endo-xylogalacturonan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

Expression profiles of pectinase genes in the reference strain and *ΔgaaR* were pairwise compared for identical time points (Table 2 and Fig. 2, Fig. S3). Most of the genes in the GA-regulon, including those required for GA transport and catabolism, are dependent on GaaR for induction on SBP (Fig. 2A-D). This observation strongly suggests that *ΔgaaR* is not utilizing GA from SBP. The expression of *gaaD*/*larA* can be explained by the dual activity of the enzyme encoded by this gene as both an L-glyceraldehyde reductase and an L-arabinose reductase [11] and the utilization of L-arabinose from SBP in *ΔgaaR* (see below). The expression profile of exo-polygalacturonases, pectin acetyl- and methylesterases, endo-polygalacturonases, and pectin lyases (Table 2 and Fig. 2B-D), all acting on the PGA backbone, support the conclusion that the GaaR target genes are not induced during growth on SBP in *ΔgaaR*.

The results described above indicate that the residual growth of *ΔgaaR* on SBP is due to the utilization of other monosaccharides released from SBP. Analysis of the monosaccharide composition of the SBP used in this study was performed as described previously [36] and showed that it contains 55 mol% GA, as well as 17 mol% L-arabinose, 16 mol% D-galactose, and 10 mol% L-rhamnose. Analysis of the expression of the genes involved in the degradation of RG-I such as exo-rhamnogalacturonases (*rgx*), rhamnogalacturonases (*rhg*), rhamnogalacturonan acetyl esterases (*rgae*), rhamnogalacturonyl hydrolases (*urhg*), arabinofuranosidases (*abf*), endo-arabinanases (*abn*), ferulic acid esterases (*fae*), and *β*-galactosidases (*lac*), and the genes responsible for catabolism of L-rhamnose, L-arabinose, and D-xylose showed that these genes were still expressed in *ΔgaaR* (Fig. 2E-H, Fig. S3), indicating that the degradation and metabolism of RG-I support the growth of *ΔgaaR* on SBP.

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

Table 2. RNA-seq analysis on SBP of the genes that depend on GaaR for induction. Expression values (FPKM) are averages of duplicates. Fold changes ≥2 and *P*-values ≤0.05 are highlighted.

NRRL3 protein ID	CBS 513.88 gene ID	Gene name	Ref SBP 2h	*ΔgaaR* SBP 2h	Fold change Ref/ *ΔgaaR* SBP 2h	*P*-value	Ref SBP 8h	*ΔgaaR* SBP 8h	Fold change Ref/ *ΔgaaR* SBP 8h	*P* -value	Ref SBP 24h	*ΔgaaR* SBP 24h	Fold change Ref/ *ΔgaaR* SBP 24h	*P* -value
NRRL3_00958	An14g04280	*gatA*[a]	849.85	12.60	67.45	6.18E-05	1077.70	21.92	49.17	2.84E-03	57.92	1.82	31.91	5.06E-02
NRRL3_08663	An03g01620	GA transporter (putative)[a]	2647.36	642.68	4.12	1.01E-04	387.81	273.28	1.42	2.41E-01	2.69	3.20	0.84	7.16E-01
NRRL3_04281	An07g00780	GA transporter (putative)[a]	33.99	7.43	4.58	1.89E-01	14.34	6.24	2.30	5.60E-02	34.49	2.63	13.11	8.63E-03
NRRL3_05650	An02g07710	*gaaA*[a]	4649.59	70.19	66.24	2.16E-03	2785.77	78.98	35.27	1.73E-03	215.02	24.06	8.94	9.85E-02
NRRL3_06890	An16g05390	*gaaB*[a]	11722.91	113.85	102.97	2.38E-03	9634.93	229.60	41.96	2.72E-04	208.56	65.62	3.18	8.85E-02
NRRL3_05649	An02g07720	*gaaC*[a]	7306.08	92.88	78.66	2.95E-05	6041.04	78.17	77.29	2.63E-03	527.35	34.38	15.34	1.02E-01
NRRL3_10050	An11g01120	*gaaD*[a]	11412.45	3807.08	3.00	5.06E-03	7573.39	1434.68	5.28	8.61E-04	621.40	409.01	1.52	4.46E-01
NRRL3_03144	An12g07500	*pgaX*[a]	948.28	5.71	166.22	1.16E-02	1154.06	21.72	53.13	4.37E-03	133.41	4.97	26.87	1.59E-03
NRRL3_09810	An11g04040	*pgxA*[a]	19.61	0.11	186.71	1.72E-02	72.87	0.60	121.44	2.09E-03	3.23	0.00	NA	NA
NRRL3_08281	An03g06740	*pgxB*[a]	483.28	3.15	153.66	1.81E-02	784.73	13.84	56.72	6.23E-02	364.35	3.40	107.32	4.37E-02
NRRL3_05260	An02g12450	*pgxC* [a]	206.37	3.08	67.11	8.92E-03	191.37	4.23	45.29	2.90E-04	7.90	8.21	0.96	7.98E-01
NRRL3_06053	An02g02540	*paeA* (putative)[a]	585.00	14.43	40.55	1.48E-02	1836.37	30.62	59.97	1.75E-02	300.70	11.83	25.43	1.46E-02
NRRL3_04916	An07g08940	*paeB* (putative)[c]	166.46	7.51	22.16	1.47E-02	427.08	62.09	6.88	9.60E-02	151.00	10.65	14.18	9.63E-04
NRRL3_08325	An03g06310	*pmeA*	15.92	1.30	12.24	3.81E-02	27.43	0.99	27.70	3.18E-03	0.90	0.12	7.50	2.43E-01
NRRL3_07470	An04g09690	*pmeB* (putative)	41.34	4.84	8.55	3.34E-02	130.79	52.94	2.47	1.77E-02	33.31	2.06	16.17	2.20E-01
NRRL3_05252	An02g12505	*pmeC* (putative)[b]	957.92	7.59	126.29	1.59E-02	1917.18	26.77	71.63	1.09E-02	249.81	0.40	624.53	1.67E-02
NRRL3_02571	An01g11520	*pgaI*	18.75	4.38	4.29	1.44E-02	106.37	6.59	16.14	8.39E-03	9.88	4.83	2.04	4.97E-01
NRRL3_05859	An02g04900	*pgaB*[c]	7.84	11.01	0.71	2.52E-01	31.37	17.51	1.79	3.24E-02	4.29	2.07	2.07	3.62E-01
NRRL3_08805	An05g02440	*pgaC*	2.97	0.65	4.60	3.05E-01	58.43	6.22	9.40	2.60E-02	17.67	0.41	43.09	2.87E-01
NRRL3_02835	An01g14670	*pgaE*[c]	3.16	1.35	2.34	8.30E-02	177.86	19.69	9.03	3.46E-03	28.57	15.13	1.89	1.99E-01

Table 2. RNA-seq analysis on SBP of the genes that depend on GaaR for induction. Expression values (FPKM) are averages of duplicates. Fold changes ≥2 and *P*-values ≤0.05 are highlighted.

NRRL3 protein ID	CBS 513.88 gene ID	Gene name	Ref SBP 2h	*ΔgaaR* SBP 2h	Fold change Ref/ *ΔgaaR* SBP 2h	*P*-value	Ref SBP 8h	*ΔgaaR* SBP 8h	Fold change Ref/ *ΔgaaR* SBP 8h	*P* -value	Ref SBP 24h	*ΔgaaR* SBP 24h	Fold change Ref/ *ΔgaaR* SBP 24h	*P* -value
NRRL3_00965	An14g04370	*pelA*[a]	40.64	11.09	3.67	1.14E-02	861.91	9.02	95.55	2.19E-02	25.43	5.41	4.70	1.82E-01
NRRL3_09811	An11g04030	*pelC*	2.39	4.58	0.52	4.27E-01	0.00	0.00	NA	NA	0.00	0.00	NA	NA
NRRL3_01237	An19g00270	*pelD*	11.72	0.61	19.37	1.10E-02	17.97	0.58	31.24	1.26E-03	1.28	0.42	3.05	8.54E-03
NRRL3_04153	An15g07160	*pelF*[c]	18.03	21.02	0.86	4.67E-01	44.85	14.09	3.18	2.56E-02	64.95	25.60	2.54	5.26E-04
NRRL3_10559	An18g04810	*rgxC* (putative)	86.41	0.94	92.41	6.26E-02	207.96	3.76	55.38	4.39E-02	116.98	1.11	105.39	1.02E-01
NRRL3_00684	An14g01130	*rglA*	2.46	0.00	NA	NA	6.60	2.32	2.84	2.96E-02	3.98	2.30	1.73	6.60E-01
NRRL3_01606	An01g00330	*abfA*[c]	3864.76	705.67	5.48	6.85E-03	133.72	176.93	0.76	3.59E-01	1.56	4.72	0.33	1.76E-01
NRRL3_10865	An08g01710	*abfC* (putative)[a]	2406.40	527.90	4.56	2.46E-02	441.88	483.86	0.91	2.48E-01	97.31	36.96	2.63	2.77E-02
NRRL3_07094	An16g02730	*abnD* (putative)	3.92	2.82	1.39	1.06E-01	14.38	6.20	2.32	4.33E-02	55.45	12.53	4.43	2.40E-01
NRRL3_02479	An01g10350	*lacB* (putative)	487.09	48.89	9.96	1.33E-03	115.20	24.31	4.74	8.71E-03	13.85	8.31	1.67	2.24E-01
NRRL3_11738	An06g00290	*lacC* (putative)	297.53	12.04	24.71	6.23E-03	314.50	8.56	36.74	3.61E-02	113.92	2.81	40.61	3.74E-02
NRRL3_10643	An18g05940	*galA*	154.02	30.49	5.05	1.82E-02	398.55	26.21	15.21	2.85E-04	19.05	20.72	0.92	5.82E-01

[a] Genes identified as the GA-regulon by Martens-Uzunova and Schaap [7]. [b] *pmeC* not present on the Affymetrix microarray. [c] Genes not significantly differentially expressed on GA, but differentially expressed on SBP.

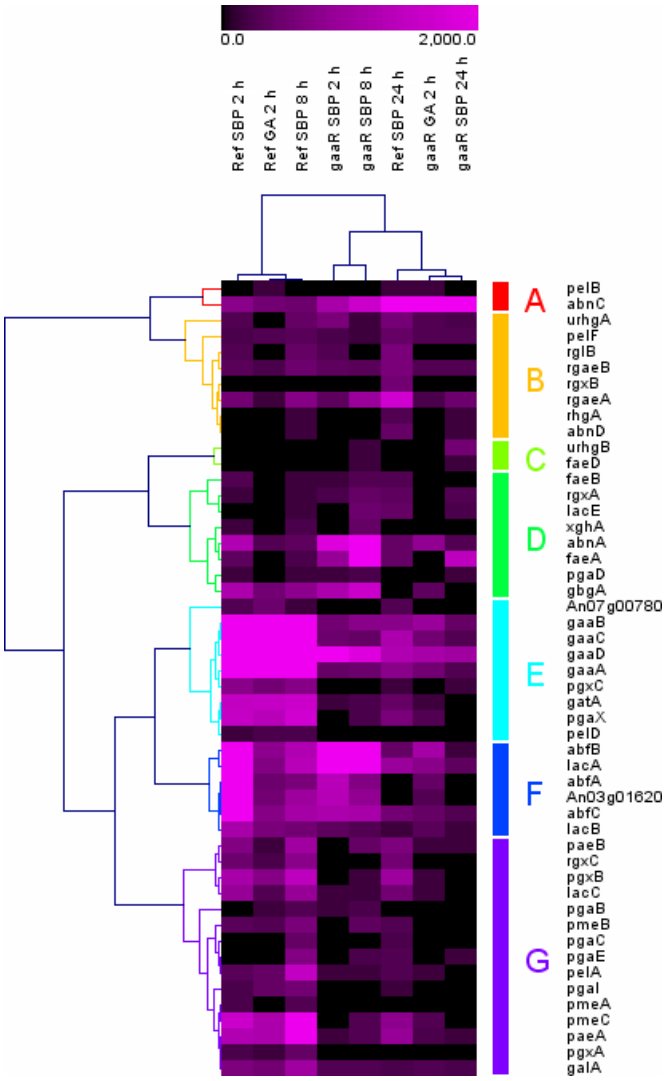

Fig. 3. Hierarchical clustering of pectin utilization genes according to their expression in the reference strain (FP-1132.1) and *ΔgaaR* (FP-1126.1) on GA and SBP. The color code displayed represents the transcript levels of the genes. Clusters E and G include genes that are members of the GA-regulon.

SBP at 24 h or in *ΔgaaR* on GA, suggesting that these genes are likely induced on starvation or derepressed conditions.

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

GaaR in *A. niger*, we identified the missing link to further understand the interplay between several TFs involved in plant cell wall degradation. Insight in the regulation of pectin degradation and GA utilization in *A. niger* can help in exploiting *A. niger* for more efficient pectinase production.

NOTE AFTER PUBLICATION

In RNA-seq experiments, two biological replicates representing each condition were used, and differential gene expression was identified by Student's t-test (*P*-value≤0.05). This approach allows addressing only major differences in gene expression.

ACKNOWLEDGEMENTS

EA and JN were supported by a grant from BE-Basic (Flagship 10) and the China Scholarship Council, respectively. JEK and MVAP were supported by a grant of the Dutch Technology Foundation STW, Applied Science division of NWO, and the Technology Program of the Ministry of Economic Affairs 016.130.609 to RPdV. We thank Patricia vanKuyk and Mark Arentshorst for constructing the kusA deletion strains. The pectin composition analysis was conducted at Complex Carbohydrate Research Centre and supported by the Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, U.S. Department of Energy grant (DEFG02-93ER20097) to Parastoo Azadi.

AUTHOR CONTRIBUTIONS

EA, JN and JEK performed experiments, MP, MVAP and EA performed bioinformatics analysis, all authors analyzed results, EA, JEK, JALK, JV, RPV and AFJR wrote the manuscript with input of all authors, RPV and AFJR designed experiments and supervised the study.

SUPPORTING INFORMATION

Additional Supporting Information may be found online (http://onlinelibrary.wiley.com/doi/10.1002/1873-3468.12211/abstract) in the supporting information tab for this article.

REFERENCES

1 Mohnen D (2008) Pectin structure and biosynthesis. *Curr Opin Plant Biol* 11, 266–277.

2 Leijdekkers AGM, Huang JH, Bakx EJ, Gruppen H and Schols HA (2015) Identification of novel isomeric pectic oligosaccharides using hydrophilic interaction chromatography coupled to traveling-wave ion mobility mass spectrometry. *Carbohydr Res* 404, 1–8.

3 Martens-Uzunova ES and Schaap PJ (2009) Assessment of the pectin degrading enzyme network of *Aspergillus niger* by functional genomics. *Fungal Genet Biol* 46, S170–S179.

4 Coutinho PM, Andersen MR, Kolenova K, Vankuyk PA, Benoit I, Gruben BS, Trejo-Aguilar B, Visser H, van Solingen P, Pakula T et al. (2009) Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. *Fungal Genet Biol* 46, S161–S169.

5 Espagne E, Lespinet O, Malagnac F, Da Silva C, Jaillon O, Porcel BM, Couloux A, Aury JM, Segurens B, Poulain J et al. (2008) The genome sequence of the model ascomycete fungus *Podospora anserina*. *Genome Biol* 9, 1–22.

6 Sloothaak J, Schilders M, Schaap PJ and de Graaff LH (2014) Overexpression of the *Aspergillus niger* GatA transporter leads to preferential use of D-galacturonic acid over D-xylose. *AMB Express* 4, 1–9.

7 Martens-Uzunova ES and Schaap PJ (2008) An evolutionary conserved D-galacturonic acid metabolic pathway operates across filamentous fungi capable of pectin degradation. *Fungal Genet Biol* 45, 1449–1457.

8 Mojzita D, Wiebe M, Hilditch S, Boer H, Penttila M and Richard P (2010) Metabolic engineering of fungal strains for conversion of D-galacturonate to *meso*-galactarate. *Appl Environ Microbiol* 76, 169–175.

9 Kuivanen J, Mojzita D, Wang Y, Hilditch S, Penttila M, Richard P and Wiebe MG (2012) Engineering filamentous fungi for conversion of D-galacturonic acid to L-galactonic acid. *Appl Environ Microbiol* 78, 8676–8683.

10 Wiebe MG, Mojzita D, Hilditch S, Ruohonen L and Penttila M (2010) Bioconversion of D-galacturonate to keto-deoxy-L-galactonate (3-deoxy-L-threo-hex-2-ulosonate) using filamentous fungi. *BMC Biotechnol* 10, 1–8.

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

12 Kowalczyk JE, Benoit I and de Vries RP (2014) Regulation of plant biomass utilization in *Aspergillus*. *Adv Appl Microbiol* 88, 31–56.

13 de Vries RP, Jansen J, Aguilar G, Parenicova L, Joosten V, Wulfert F, Benen JA and Visser J (2002) Expression profiling of pectinolytic genes from *Aspergillus niger*. *FEBS Lett* 530, 41–47.

14 Niu J, Homan TG, Arentshorst M, de Vries RP, Visser J and Ram AF (2015) The interaction of induction and repression mechanisms in the regulation of galacturonic acid-induced genes in *Aspergillus niger*. *Fungal Genet Biol* 82, 32–42.

15 Chang P and Ehrlich KC (2013) Genome-wide analysis of the $Zn(II)_2Cys_6$ zinc cluster-encoding gene family in *Aspergillus flavus*. *Appl Microbiol Biotechnol* 97, 4289–4300.

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

17 Gruben BS, Zhou M, Wiebenga A, Ballering J, Overkamp KM, Punt PJ and de Vries RP (2014) *Aspergillus niger* RhaR, a regulator involved in L-rhamnose release and catabolism. *Appl Microbiol Biotechnol* 98, 5531–5540.

18 Battaglia E, Visser L, Nijssen A, van Veluw GJ, Wösten HAB and de Vries R (2011) Analysis of regulation of pentose utilisation in *Aspergillus niger* reveals evolutionary adaptations in Eurotiales. *Stud Mycol* 69, 31–38.

19 van Peij NNME, Visser J and de Graaff LH (1998) Isolation and analysis of *xlnR*, encoding a transcriptional activator co-ordinating xylanolytic expression in *Aspergillus niger*. *Mol Microbiol* 27, 131–142.

20 Gruben BS, Zhou M and de Vries RP (2012) GalX regulates the D-galactose oxido-reductive pathway in *Aspergillus niger*. *FEBS Lett* 586, 3980–3985.

21 Raulo R, Kokolski M and Archer DB (2016) The roles of the zinc finger transcription factors XlnR, ClrA and ClrB in the breakdown of lignocellulose by *Aspergillus niger*. *AMB Express* 6, 1–12.

22 Zhang L, Lubbers RJM, Simon A, Stassen JHM, Ribera PRV, Viaud M and van Kan JAL (2015) A novel Zn_2Cys_6 transcription factor BcGaaR regulates D-galacturonic acid utilization in *Botrytis cinerea*. *Mol Microbiol* 100, 247–262.

23 Goosen T, Bloemheuvel G, Gysler C, de Bie DA, van den Broek HWJ and Swart K (1987) Transformation of *Aspergillus niger* using the homologous orotidine-5'-phosphate-decarboxylase gene. *Curr Genet* 11, 499– 503.

24 Meyer V, Arentshorst M, El-Ghezal A, Drewsd A, Kooistra R, van den Hondel CAMJJ and Ram AFJ (2007) Highly efficient gene targeting in the *Aspergillus niger kusA* mutant. *J Biotechnol* 128, 770–775.

25 Carvalho NDSP, Arentshorst M, Kwon MJ, Meyer V and Ram AFJ (2010) Expanding the *ku70* toolbox for filamentous fungi: establishment of complementation vectors and recipient strains for advanced gene analyses. *Appl Microbiol Biotechnol* 87, 1463–1473.

26 Arentshorst M, Ram AFJ and Meyer V (2012) Using non-homologous end-joining-deficient strains for functional gene analyses in filamentous fungi. *Methods Mol Biol* 835, 133–150.

27 Arentshorst M, Niu J and Ram AFJ (2015) Efficient generation of *Aspergillus niger* knock out strains by combining NHEJ mutants and a split marker approach. In Genetic Transformation Systems in Fungi. Vol. 1 (van den Berg MA and Maruthachalam K, eds), pp. 263–272. Springer International Publishing, Cham, Switzerland.

28 Sambrook J and Russell DW (2001) Molecular Cloning: A Laboratory Manual. 3rd edn. Cold Spring Harbor Press, New York, NY, USA.

29 Langmead B, Trapnell C, Pop M and Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10, 25–34.

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

31 Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ and Pachter L (2010) Transcript assembly and abundance estimation from RNA-Seq reveals thousands of new transcripts and switching among isoforms. *Nat Biotechnol* 28, 511–515.

32 Li B and Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12, 1–16.

33 Edgar R, Domrachev M and Lash AE (2002) Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30, 207–210.

34 Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215, 403–410.

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

36 Santander J, Martin T, Loh A, Pohlenz C, Gatlin DM and Curtiss R (2013) Mechanisms of intrinsic resistance to antimicrobial peptides of *Edwardsiella ictaluri* and its influence on fish gut inflammation and virulence. *Microbiology* 159, 1471–1486.

CHAPTER 4

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

Niu J, Alazi E, Reid ID, Arentshorst M, Punt PJ, Visser J, Tsang A, Ram AFJ

Published in *Genetics* (2017) 205:169-183

ABSTRACT

The expression of genes encoding extracellular polymer-degrading enzymes and the metabolic pathways required for carbon utilization in fungi are tightly controlled. The control is mediated by transcription factors that are activated by the presence of specific inducers, which are often monomers or monomeric derivatives of the polymers. A D-galacturonic acid-specific transcription factor named GaaR was recently identified and shown to be an activator for the expression of genes involved in galacturonic acid utilization in *Botrytis cinerea* and *Aspergillus niger*. Using a forward genetic screen, we isolated *A. niger* mutants that constitutively express GaaR-controlled genes. Reasoning that mutations in the *gaaR* gene would lead to a constitutively activated transcription factor, the *gaaR* gene in 11 of the constitutive mutants was sequenced, but no mutations in *gaaR* were found. Full genome sequencing of five constitutive mutants revealed allelic mutations in one particular gene encoding a previously uncharacterized protein (NRRL3_08194). The protein encoded by NRRL3_08194 shows homology to the repressor of the quinate utilization pathway identified previously in *Neurospora crassa* (*qa-1S*) and *Aspergillus nidulans* (QutR). Deletion of NRRL3_08194 in combination with RNA-seq analysis showed that the NRRL3_08194 deletion mutant constitutively expresses genes involved in galacturonic acid utilization. Interestingly, NRRL3_08194 is located next to *gaaR* (NRRL3_08195) in the genome. The homology to the quinate repressor, the chromosomal clustering, and the constitutive phenotype of the isolated mutants suggest that NRRL3_08194 is likely to encode a repressor, which we name GaaX. The GaaR–GaaX module and its chromosomal organization is conserved among ascomycetes filamentous fungi, resembling the quinate utilization activator-repressor module in amino acid sequence and chromosomal organization.

INTRODUCTION

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

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

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

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

MATERIALS AND METHODS

Strains, media, and growth conditions

All strains in this study are listed in Table 1. Strains were grown in liquid or solidified (1.5% agar) minimal medium (MM) containing 7 mM KCl, 8 mM KH_2PO_4, 70 mM $NaNO_3$, and 2 mM $MgSO_4$ (pH adjusted to 5.5) as described (Bennett and Lasure 1991). MM was supplemented with 50 mM glucose, 50 mM D-galacturonic acid, 50 mM fructose, or 50 mM sorbitol as carbon source. Complete medium (CM) was also used and consists of MM supplemented with 0.1% casamino acids and 0.5% w/v yeast extract and 50 mM glucose. MM agar plates containing acetamide as sole nitrogen source were made as described previously (Arentshorst *et al*. 2012).

Isolation of mutants with constitutive expression of genes involved in PGA utilization

A. niger strain JN29.2 (Table 1) was used for the selection of mutants with constitutive expression of genes involved in PGA utilization. Spontaneous mutants were obtained by plating out freshly harvested and myracloth-filtered conidia ($1x10^4$ conidia per plate) on MM glucose/acetamide plates and incubated at 30^o for 5 days. In addition, mutants were obtained after mild UV-mutagenesis (80% survival) as described (Damveld *et al*. 2008). Individual mutants growing on the primary MM-glucose/acetamide selection plates were purified twice on the MM-glucose/acetamide agar plates. In total, 14 spontaneous mutants and 59 UV-mutants were isolated that grew well on MM-glucose/acetamide agar plates and they were potential mutants with constitutive expression of genes involved in PGA utilization. To identify mutants constitutively producing PGA-degrading enzymes, all 73 mutants were grown by inoculating $5x10^7$ spores in 50 ml MM-glucose medium for 36 hr at 30^o with shaking at 150 rpm. Supernatant of each culture was harvested by filtration. The extracellular culture fluid and the mycelia were stored at -80^o for enzymatic assays and RNA extraction, respectively. A total of 10 ml supernatant of each sample was spotted on PGA plates made by dissolving 0.2% PGA (Sigma) in NaAc buffer (pH 4.2) with 1% agarose (Sphaero). The PGA plate assay was modified from the protocol used to detect cellulase activity (Teather and Wood 1982). Plates were incubated at 37^o for 17 hr after spotting. PGA was stained by flooding the plates with a filter-sterile 0.05% solution of

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

Name	Genotype/description	Reference/source
N402	*cspA1*, derivative of N400	Bos *et al.* 1988
AB4.1	*pyrG*$^{-}$,derivative of N402	van Hartingsveldt *et al*. 1987
MA234.1	*Δku70::DR_amdS_DR* in MA169.4	Alazi *et al.* 2016
MA70.15	*Δku70::amdS* in AB4.1	Meyer *et al.* 2007
MA299.2	*Δku70* in MA70.15	Niu *et al.* 2015
MA323.1	*Δku70::amdS, ΔnicB*$^{-}$*, pyrG*$^{-}$	Niu *et al.* 2016
JC1.5	*pgaX-amdS* in MA299.2, *pyrG*$^{+}$	Niu *et al.* 2015
JN29.2	*ΔcreA*::*hygB* in JC1.5	Niu *et al.* 2015
JN38	spontaneous mutation S1 in JN29.2	This study
JN39	spontaneous mutation S2 in JN29.2	This study
JN42	spontaneous mutation S5 in JN29.2	This study
JN44	spontaneous mutation S7 in JN29.2	This study
JN52	UV1 in JN29.2	This study
JN53	UV2 in JN29.2	This study
JN54	UV3 in JN29.2	This study
JN55	UV4 in JN29.2	This study
JN56	UV5 in JN29.2	This study
JN57	UV6 in JN29.2	This study
JN58	UV7 in JN29.2	This study
JN59	UV8 in JN29.2	This study
JN60	UV9 in JN29.2	This study
JN61	UV10 in JN29.2	This study
JN62	UV11 in JN29.2	This study
JN63	UV12 in JN29.2	This study
JN64	UV13 in JN29.2	This study
JN122.1, JN122.2, JN122.3	*ΔgaaX::phleo* in JN29.2	This study
JN123.1, JN123.2, JN123.3	*ΔgaaX::hygB* in JC1.5	This study
JN125.1	*ΔgaaX::nicB* in MA323.1	This study
JN126.2, JN126.5, JN126.6	PgaaX::GaaX::GFP::TgaaX in JN125.1	This study
JN127.1, JN127.2, JN127.3	PgaaX::GFP::GaaX::TgaaX in JN125.1	This study

Congo Red (Sigma) dissolved in Milli-Q water for 15 min. The Congo Red solution was then poured off and the plates were washed with Milli-Q water and further treated by flooding with 1 M NaCl for 15 min. The formation of a clear zone of hydrolysis indicated PGA degradation.

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

DNA sequencing and data analysis

Sequencing of the *gaaR* gene from 11 constitutive mutants was performed by PCR amplification of the *gaaR* gene, including 137 bp upstream and 152 bp downstream sequences, using genomic DNA of the mutants as template and primers gaaRP7f and gaaRP8r (Table S1). Genomic DNA was isolated as described (Arentshorst *et al*. 2012). The PCR fragment (2765 bp in size) was sequenced in both directions using *gaaR* sequencing primers (Table S1). Sequencing was performed by Macrogen Europe (Amsterdam, The Netherlands).

Genomic DNA of three spontaneous mutants and two UV-mutants was isolated as described (Arentshorst *et al*. 2012), and was further purified with DNA Isolation Kit (MO BIO Laboratories) for whole-genome DNA sequencing. The mutant genomes were sequenced at the McGill University Génome Québec Innovation Centre (QC, Canada) using the Illumina HiSeq platform to ~50-fold coverage. The DNA reads were aligned to the NRRL3 genome with Bowtie2 (Langmead and Salzberg 2012) and sequence differences were detected with Freebayes (Garrison and Marth 2012).

Deletion of *gaaX* gene

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

was isolated as described (Arentshorst *et al*. 2012) and Southern blot hybridizations, using PCR-amplified fragments generated with primers listed in Table S1 as probes, were performed to confirm proper deletion and to exclude additional integrations.

Bioreactor cultivation

Controlled bioreactor cultivations for *A. niger* strains MA234.1 and JN123.1 were performed in 6.6-L BioFlo3000 bioreactors (New Brunswick Scientific) as previously described (Jørgensen *et al*. 2010). Briefly, autoclaved bioreactor vessels were filled with 5 liter of sterile MM with 0.75% fructose. During cultivation at 30^{o}, the controller was set to maintain pH 3 by addition of titrants (2 M NaOH or 1 M HCl). Sterile air was supplied at a rate of 1 liter/min. Prior to inoculation, 1.5 ml of 10% (w/v) filter-sterilized yeast extract was added to enhance conidial germination. Cultures were inoculated with freshly harvested spores at a concentration of $7.0x10^{8}$ conidia per liter. To reduce the loss of hydrophobic conidia during germination, the stirrer speed was set to 250 rpm and the culture was aerated via the headspace during the first 6 hr after inoculation. Subsequently, the stirrer speed was increased to 750 rpm, 0.5 ml of polypropyleneglycol P2000 was added as an antifoam agent, and air was supplied via the sparger. Cultures broth was harvested at regular intervals from batch cultures and mycelial biomass was retained by vacuum filtration using glass microfiber filters (Whatman). Both biomass and filtrate were quickly frozen in liquid nitrogen and subsequently stored at -80^{o}. Dry biomass concentrations were gravimetrically determined from lyophilized mycelia originating from a known mass of culture broth.

Transcriptome analysis

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

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

To construct fusions of GFP to the N-terminus or C-terminus of GaaX, PgaaX_GFP::GaaX_TgaaX and PgaaX_GaaX::GFP_TgaaX constructs were generated using a fusion-PCR approach in which N402 genomic DNA, as well as plasmid PagsA_eGFP_TtrpC (Damveld *et al*. 2008), were used as template DNA. For constructing the PgaaX_GFP::GaaX_TgaaX construct, the promoter region of *gaaX* was PCR amplified using primers PgaaX_P7f-NotI and PgaaX_P11r, GFP was PCR amplified from plasmid PagsA_eGFP_TtrpC using primers GFP_P1f and GFP_P3r, *gaaX* and the terminator region of *gaaX* was PCR amplified using primers gaaX_P12f and TgaaX_P10r-NotI, and the three fragments were combined together in a two-step fusion PCR. Two amino acids (Gly-Ala) were introduced as spacer between GFP and GaaX. Subsequently, the fusion fragment was cloned into vector pJet1.2 to give plasmid pJN34. For the PgaaX_GaaX:: GFP_TgaaX construct, *gaaX* with the promoter region of *gaaX* was PCR amplified using primers PgaaX_P7f-NotI and gaaX_P8r, GFP was PCR amplified using primers GFP_P1f and GFP_P2r, the terminator region of *gaaX* was PCR amplified using primers gaaX_P9f and TgaaX_P10r-NotI, and the three fragments were combined together in a two-step fusion PCR. Again, a Gly-Ala spacer was introduced between GaaX and GFP. Subsequently, the fusion fragment was cloned into vector pJet1.2 to give plasmid pJN35.

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

Microscopy

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

Data availability

Strains are listed in Table 1 and are available upon request. Table S2 contains SNPs and indels detected in genomes of mutants. Table S3A contains transcript per million (TPM) values of NRRL3 gene models in wild type and the *gaaX* mutant, and Table S3B contains

their DEseq2 analysis. The DNA reads described in this study are deposited in the Short Read Archive under accession number SRP078415. The RNA reads described in this study are deposited in the Short Read Archive under accession number SRP078485.

RESULTS

Mutants constitutively expressing genes related to galacturonic acid utilization

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

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

To further demonstrate that the presumed trans-acting mutations indeed affected expression of multiple genes related to GA utilization and belonging to the GA-induced genes (Martens-Uzunova and Schaap 2008; Niu *et al*. 2015; Alazi *et al*. 2016), the expression of three GA-induced genes (*pgaX*, *gatA*, and *gaaB*) was examined by Northern blot analysis in a subset of mutants after growth on glucose. The *pgaX*, *gatA*, and *gaaB* genes encode an exo-polygalacturonase, a GA-specific transporter, and the L-galactonic acid dehydratase involved in the GA release from PGA, the uptake of GA, and subsequent metabolism of GA, respectively. As shown in Figure 1B, expression of these genes was not detected in the wild type (N402) and the parent JN29.2 (*ΔcreA*, P*pgaX-amdS*) strains, whereas these genes were expressed in the constitutive mutants that displayed increased polygalacturonase activity. The presumed *cis*-acting mutants from the plate assay (UV2,

UV7, UV9, UV10, UV11, and S2) did not constitutively express *pgaX*, *gatA*, and *gaaB*, and showed a small halo on PGA plates, indicating the halo assay can be used to discriminate between *cis*- and *trans*-acting mutants. To determine whether a *cis*-acting mutation in the *pgaX* promoter in front of the *amdS* gene was responsible for the ability of this class of mutants to grow on acetamide, the *pgaX* promoter in front of the *amdS* gene of all eight *cis*-acting mutants was PCR amplified using *pyrG*- and *amdS*-specific primers (Table S1). This analysis revealed no mutations in the *pgaX* promoter region of any of the eight presumed *cis*-acting mutants. Hence, the nature of the mutation(s) in these strains that allows growth on acetamide remains unknown. A possibility could be the activation of expression of endogenous *amdS* genes, as at least four *amdS*-like genes are present in the genome of *A. niger*.

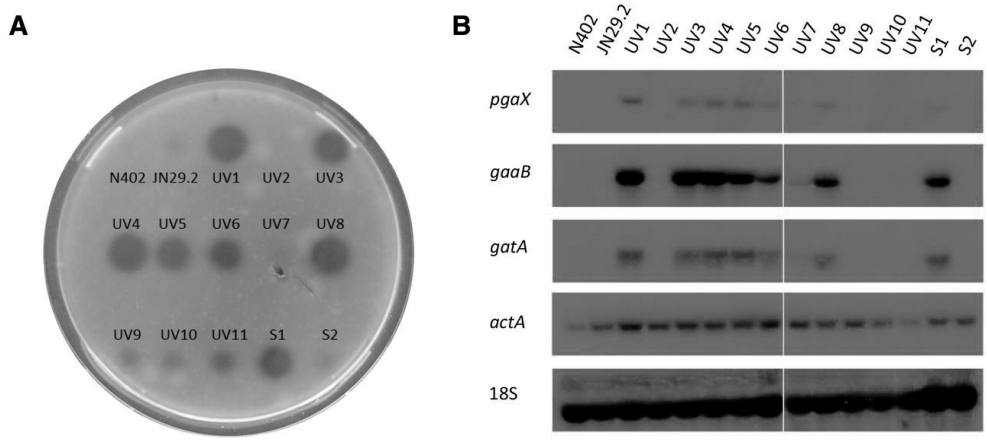

Figure 1. Enzymatic and RNA blot analysis of mutants with constitutive expression of genes involved in polygalacturonic acid utilization. **(A)** Supernatant (10 ml) from glucose-grown cultures of reference strains N402 and JN29.2, 11 UV-mutants, and two spontaneous mutants were spotted on polygalacturonic acid agarose medium to detect polygalacturonase activity. **(B)** Northern blot analysis of selected GA-responsive genes in the reference strains N402 and JN29.2, 11 UV-mutants, and two spontaneous mutants.

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

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

To determine whether the constitutive expression of GA-induced genes in these mutants involves a functioning GaaR transcription factor, we deleted the *gaaR* gene in seven of the constitutive mutants (UV1, UV3, UV4, UV5, UV6, UV8, and S1) and analyzed constitutive

expression using the P*pgaX-amdS* reporter. All seven mutants were unable to grow on glucose/acetamide plates (data not shown), indicating that the constitutive expression of the GA-induced genes requires a functional *gaaR* gene.

To identify mutation(s) in the gene(s) responsible for the constitutive phenotype, the genomes of five mutants (UV1, UV8, S1, S5, and S7) and the parental strain JN29.2 were sequenced. Table 2 summarizes the number of SNPs and indels detected in the five mutant strains and Table S2 lists positions and type of all SNPs and indels detected. Spontaneous mutant S7 contains only 11 SNPs or indels, of which 10 are located in intergenic regions and only one SNP mutated a gene (NRRL3_08194). Remarkably, in all four other mutants a mutation was found in the same gene. Two of the mutants carry nonsense mutations (UV1 and S5) and one carries a frame-shift mutation (S3), all leading to premature stop codons and predicted to result in truncated proteins (Table 2). Mutants S7 and UV8 have missense mutations in the C-terminal part of the protein. These results strongly suggest that the constitutive expression of genes encoding pectin degrading enzymes in the five mutants is caused by a loss of function of the protein encoded by NRRL3_08194.

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

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

NRRL3_08194) are clustered in the genome, similar to the quinic acid utilization transcriptional activator (QutA/Qa-1F) and repressor (QutR/ Qa-1S) in *Aspergillus nidulans* and *Neurospora crassa*, respectively (Geever *et al*. 1989; Levesley *et al*. 1996).

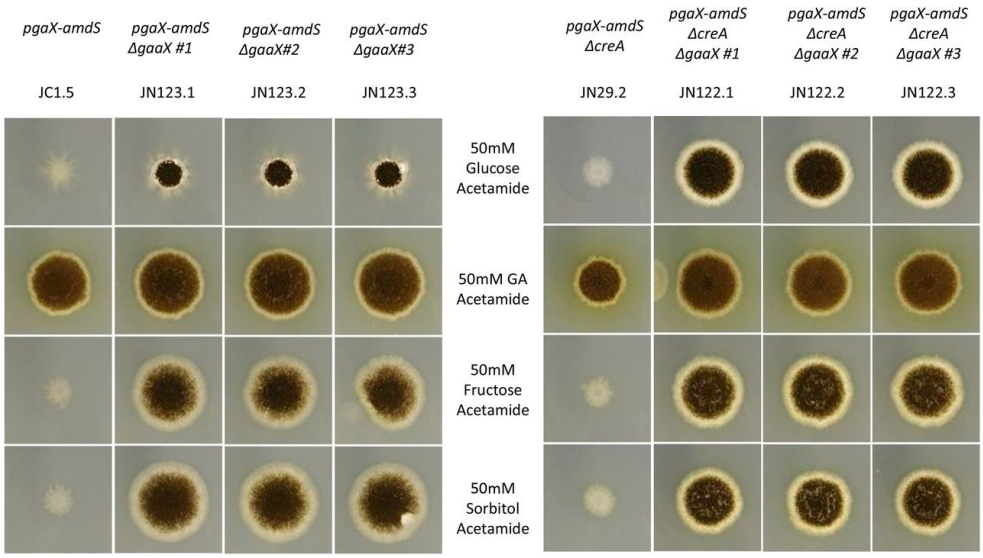

Figure 2. Regulation of the *pgaX* expression is controlled by GaaX and by CreA-mediated glucose repression. Growth of P*pgaX-amdS* reporter strains in *gaaX* deletion and their parental strains was examined on various carbon sources. Parental strains and corresponding deletion strains were grown on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

The GaaR–GaaX target gene regulon

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

To identify differentially expressed genes in the *ΔgaaX* strain as compared to its parental strain, RNA-seq was performed on RNA isolated from exponentially growing cells at the time point at which about 75–80% of the maximum biomass yield was reached. RNA-seq reads were mapped to the NRRL3/N400 genome as this is the parent of the laboratory strain N402 and derivatives used in this study. TPM values were calculated using Salmon (Patro *et al*. 2016) (Table S3). Analysis of differential gene expression, based on a stringent false discovery rate (FDR) <0.001 and a fold change (FC) >4.0, identified 37 upregulated genes (Table 3). Gene Ontology (GO) enrichment analysis using FetGOat (Nitsche *et al*. 2012) and manual inspection of the genes upregulated in the *ΔgaaX* mutant indicated that genes involved in pectin catabolism were highly enriched. Of the 37 genes, 16 are predicted to encode extracellular enzymes acting on the GA backbone of pectin or acting on pectin side chains (Table 3). Nine genes in the group of 37 upregulated genes in the *ΔgaaX* strain are predicted to encode intracellular proteins. Four of these nine genes (*gaaA*–*gaaD*) are required for the conversion of GA into pyruvate and glycerol (Martens-Uzunova and Schaap 2008). The exact role of the other five genes and their possible role in GA catabolism is currently unknown. The group of 37 upregulated genes also includes seven genes predicted to encode sugar transporter proteins. Of these seven transporter-encoding genes, only GatA has been studied in detail and shown to be able to transport GA (Sloothaak *et al*. 2014). Apart from the genes encoding extracellular enzymes (16), transporters (seven), and enzymes possibly involved in GA catabolism (nine), the remaining five genes in this group encode proteins with unknown functions or with similarities to known proteins that, for now, cannot be directly linked to GA metabolism. The deletion of *gaaX* has the most profound effect on the transcript levels of the genes encoding the first three steps of the GA utilization pathway (*gaaA*, *gaaB*, and *gaaC*) and on the expression of *gatA*. Deletion of *gaaX* resulted in a 1.24-fold (P = 0.000035) increase in *gaaR* gene activity. Since the upregulation of *gaaR* in the *ΔgaaX* mutant is modest, it is likely that the repressing activity of GaaX is mediated at the protein level (e.g., by interacting with GaaR) rather than by transcriptional control of *gaaR*. Seventeen of the 37 genes upregulated in the *ΔgaaX* mutant were previously identified as part of the GA regulon (Martens-Uzunova and Schaap 2008; Alazi *et al*. 2016) (Figure 3 and Table 3). Sixteen of the 17 genes found in common with previous studies are predicted or demonstrated to encode extracellular pectin-degrading enzymes. These results indicate that loss of function of *gaaX* affects the expression of the GA regulon. The other 20 genes were identified as significantly upregulated in the *gaaX* mutant, but these were not identified previously as being part of the GA regulon (Figure 3 and Table 3). A re-examination of the expression of these 20 genes in the RNA-seq data published earlier (Alazi *et al*. 2016) indicated that 10 of the genes (indicated in Table 3 by the asterisk, Table S4)were also GA-induced or GaaR-dependent for induction in this previous study. On the other hand, 15 genes identified to be GA-induced in a GaaR-dependent manner in the previous study (Alazi *et al*. 2016) were not significantly upregulated in the *gaaX* deletion strain (Figure 3). These results therefore suggest that full induction of GA-inducible genes requires more than the loss of GaaX activity, and that an additional induction mechanism plays a role.

An additional GA-induced gene identified in the study of Martens-Uzunova and Schaap (2008) but missing in the GaaR study is *gaaX* itself. Expression of *gaaX* was not examined

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

Strain	Total number of SNPs and indels	SNPs or indels in coding region	Mutation in NRRL3_ 08194	Position of mutation relative to ATG of NRRL3_08194	Mutation in codon (bold)	Amino acid change	Predicted protein length (Full length protein is 697 amino acids)
29.2_UV1	40	19	G-T	1372	**G**AG-**T**AG	E to Stop	457 aa
JN29.2_UV8	21	4	G-A	1577	G**G**A-G**A**A	G to E (526)	697 aa
JN29.2_S1	68	34	Extra G	1958	GTT-G**G**T	V to G out of frame	663 aa
JN29.2_S5	48	11	C-T	1105	**C**AA-**T**AA	Q to Stop	368 aa
JN29.2_S7	11	1	T-C	2027	C**T**G-C**C**G	L-P (676)	697 aa

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

Gene ID (NRRL3)	Gene ID (CBS513.88)	Gene name	Description	Average expression value WT[a]	Average expression value *ΔgaaX*[a]	FC *ΔgaaX* VS WT[b]	False discovery rate[b]	Predicted localization	GARE element[c]
NRRL3_05649	An02g07720[d,e]	*gaaC*	L-threo-3-deoxy-hexylosonate aldolase	12.54	2283.77	169.39	0.00E+00	intracellular	(-) -292, -606
NRRL3_09863[f]	An11g03500		alpha-hydroxy acid dehydrogenase	0.82	165.84	160.07	0.00E+00	intracellular	(+)-543, (-)-182
NRRL3_00958	An14g04280[d,e]	*gatA*	MFS-type sugar/inositol transporter	3.47	524.95	140.38	0.00E+00	membrane	(+) -360
NRRL3_06890	An16g05390[d,e]	*gaaB*	L-galactonic acid dehydratase	47.77	6256.70	121.49	0.00E+00	intracellular	(+) -326
NRRL3_03291	An12g05600		heterokaryon incompatibility protein	0.00	6.78	82.98	6.48E-78	intracellular	(+)-737, -325
NRRL3_05650	An02g07710[d,e]	*gaaA*	D-galacturonic acid reductase	19.92	1515.44	71.84	0.00E+00	intracellular	(+) -414, -100
NRRL3_03144	An12g07500[d,e]	*pgaX*	exo-polygalacturonase	1.36	51.63	32.66	2.60E-272	extracellular	(+)-388
NRRL3_05252	An02g12505[e]	*pmeC*	pectin methylesterase	1.06	31.32	25.03	7.10E-189	extracellular	(+)-275, -246, -35
NRRL3_06244[f]	An02g00140		glycoside hydrolase family 43 protein	0.81	22.19	23.46	8.13E-193	extracellular	(+)-96, (-)-712

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

Gene ID (NRRL3)	Gene ID (CBS513.88)	Gene name	Description	Average expression value WT[a]	Average expression value *ΔgaaX*[a]	FC *ΔgaaX* VS WT[b]	False discovery rate[b]	Predicted localization	GARE element[c]
NRRL3_08281	An03g06740[d,e]	*pgxB*	exo-polygalacturonase Pgx28B	0.00	2.10	22.95	3.59E-43	extracellular	(-) -298, -823
NRRL3_10559	An18g04810[e]	*rgxC*	glycoside hydrolase family 28 protein	0.08	3.11	17.32	1.33E-45	extracellular	(+)-852, (-) -250
NRRL3_05260	An02g12450[d,e]	*pgxC*	exo-polygalacturonase Pgx28C	0.95	16.77	15.22	1.83E-144	extracellular	(+) -268, (-) -642
NRRL3_08663	An03g01620[d,e]		MFS-type sugar/inositol transporter	0.27	5.62	14.28	8.28E-46	membrane	(+) -673
NRRL3_03342[f]	An12g04990		short-chain dehydrogenase/reductase	0.81	14.46	13.51	2.10E-62	intracellular	none
NRRL3_10865	An08g01710[d,e]	*abfC*	alpha-N-arabinofuranosidase	0.80	11.21	11.92	1.19E-84	extracellular	none
NRRL3_09862	An11g03510		hypothetical protein	0.00	0.93	11.05	4.32E-18	unknown	(-) -517, -829, -843
NRRL3_10050	An11g01120[d,e]	*gaaD/ larA*	NADPH-dependent erythrose reductase Err1	256.41	2732.37	10.16	0.00E+00	intracellular	(-) -538, -583, -801 -813
NRRL3_00957	An14g04260		B3/B4 domain-containing protein	1.53	18.56	10.05	2.46E-64	unknown	none
NRRL3_00502	An09g06200		hypothetical protein	0.93	12.38	9.36	1.08E-32	unknown	(-) -189
NRRL3_10558[f]	An18g04800		alpha-L-rhamnosidase	0.35	3.85	9.09	2.19E-61	extracellular	(+) -365
NRRL3_11710	An06g00620		MFS-type sugar/inositol transporter	2.86	25.29	8.05	1.39E-118	membrane	(+)-487, (-) -368
NRRL3_06053	An02g02540[d,e]	*paeA*	carbohydrate esterase family 16 protein	2.06	17.62	7.76	5.39E-107	extracellular	(+) -1238
NRRL3_01073	An14g05840		O-methyltransferase, COMT-type	0.54	5.57	7.16	1.06E-21	intracellular	(+) -300
NRRL3_07382	An16g00540		alpha-L-fucosidase	0.04	0.67	7.07	5.92E-14	extracellular	(-) -606
NRRL3_00660[f]	An14g00860		carboxylesterase	0.21	1.87	6.98	1.47E-31	extracellular	none

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

Gene ID (NRRL3)	Gene ID (CBS513.88)	Gene name	Description	Average expression value WT[a]	Average expression value *ΔgaaX*[a]	FC *ΔgaaX* VS WT[b]	False discovery rate[b]	Predicted localization	GARE element[c]
NRRL3_08499	An03g03960		uncharacterized protein	0.48	4.58	6.71	3.84E-21	unknown	none
NRRL3_07470	An04g09690[e]	*pmeB*	pectin methylesterase	0.75	5.72	6.36	3.59E-43	extracellular	(+) -389
NRRL3_02479	An01g10350[e]	*lacB*	exo-beta-1,4-galactanase	3.48	22.29	6.03	1.37E-179	extracellular	none
NRRL3_03292	An12g05590		carboxylesterase	0.18	1.86	5.26	1.75E-09	extracellular	(-) -1000
NRRL3_01237	An19g00270[e]	*pelD*	pectin lyase	0.17	1.47	5.25	6.14E-14	extracellular	(-) -409, -465
NRRL3_03467	An12g03550		MFS-type transporter	1.13	6.73	5.05	4.03E-28	membrane	none
NRRL3_01127[f]	An14g06500	DHAK	dihydroxyacetone kinase	18.90	100.64	4.98	1.61E-102	intracellular	(+) -69
NRRL3_11054[f]	An08g04040		MFS-type sugar/inositol transporter	6.62	31.89	4.57	2.57E-126	membrane	(+)-778, -726, (-) -646
NRRL3_08833[f]	n.a.		hypothetical protein	0.31	2.02	4.52	1.58E-12	unknown	none
NRRL3_01398[f]	An13g02090		MFS-type transporter	2.81	13.10	4.16	5.62E-30	membrane	(-) -270
NRRL3_08325	An03g06310[e]	*pmeA*	pectin methylesterase Pme8A	0.04	0.56	4.10	1.49E-07	extracellular	(+)-983, (-) -308
NRRL3_02770	An01g13880		MFS-type transporter	0.87	4.07	4.06	1.57E-20	membrane	(+) -578

Gray boxes indicate genes proposed to be part of the GaaR/GaaX regulon.

[a] Expression values (TPM) are averages of biological triplicates for the wild type (wt) and duplicates for the *ΔgaaX* strain.

[b] Fold changes and FDR-values are calculated with DESeq2, and genes with fold changes ≥4 and FDR-values ≤0.001 are listed.

[c] GA-responsive element (GARE) position (CCNCCAA) is given with respect to the translational start site.

[d] Genes identified as the GA-regulon by Martens-Uzunova and Schaap (2008).

[e] Genes identified as the GA regulon by Alazi *et al*. (2016).

[f] Genes not identified as components of the GA regulons in the studies by Martens-Uzunova and Schaap (2008) and Alazi *et al*. (2016), but are likely to be under the regulation of GaaR (see text for details).

in the Alazi *et al*. (2016) study as its function was not yet directly linked to GA utilization. However, re-evaluation of the dataset revealed that the induced expression of *gaaX* on GA is dependent on GaaR (FC of wild type vs. *ΔgaaR*: 18.7; P = 0.003; Table S4). Combining the expression data of the *ΔgaaX* mutant (this study), the *ΔgaaR* mutant (Alazi *et al*. 2016) and the genes induced on GA (Martens-Uzunova and Schaap 2008), we propose a panregulon of 53 GaaR–GaaX controlled genes and a core GaaR–GaaX regulon of at least 27 genes (Figure 3, Table 3, and Table S4). These 27 genes include 11 genes present in the intersection of all three data sets, six genes present in the intersection of the *ΔgaaX* data and the *ΔgaaR* data (Alazi *et al*. 2016), nine genes identified by examining the *gaaX* dataset with supporting evidence from previous studies, and *gaaX* (Figure 3). Of these 27 genes, all except NRRL3_00660 (carboxyesterase), NRRL3_10865 (*α*-*N*-arabinofuranosidase), NRRL3_03342 (short-chain dehydrogenase/reductase), NRRL3_08833 (hypothetical protein), and NRRL3_02479 (*β*-galactosidase), have at least one predicted GARE motif in the upstream regions of the coding region (Table 3). It is interesting to note that among the genes listed in Figure 3 and Table 3 that are upregulated in the *ΔgaaX*, some of them (NRRL3_00957 and NRRL3_00958; NRRL3_09862 and NRRL3_09863; NRRL3_03291 and NRRL3_03292) are clustered. Except for NRRL3_00958, which encodes a GA-specific transporter (Sloothaak *et al*. 2014), the possible role of these genes in pectin degradation is currently unknown.

GaaX is induced on galacturonic acid and localized in the cytosol

GaaX was previously identified as a GA-induced gene with unknown function (Martens-Uzunova and Schaap 2008). To monitor the induction of GaaX and to localize the GaaX protein in the cell, GaaX was fused to GFP at either the N- or C-terminal part of GaaX and expressed from the endogenous GaaX promoter. Fusion constructs were targeted to the *pyrG* locus of *A. niger* in a strain lacking endogenous *gaaX* (JN125.1) to be able to test complementation of the GFP-GaaX and GaaX-GFP fusion proteins (Figure S3). As shown in Figure 4A, JN125.1 (*ΔgaaX*::*nicB*) constitutively expressed pectinases indicated by the halo on PGA plates, while both the C-terminally tagged as well as the N-terminally tagged versions of GaaX (JN126.2 and JN127.3, respectively) complemented the constitutive expression phenotype, indicating that both fusion proteins are functional. Confocal fluorescent microscopy was performed on GFP-tagged strains to localize GaaX (Figure 4B). Spores were germinated either on GA or on fructose (a nonrepressing carbon source) and a fluorescent signal was only detectable in the GFP-labeled strains after growth on GA. This observation confirms the results from the expression data that indicate that GaaX is lowly expressed under noninducing conditions and is induced on GA. The expression of GaaX is low on fructose and no GFP signal above the background level was detected on fructose. Based on the fluorescent pictures, GaaX is likely to be localized in the cytosol.

DISCUSSION

The forward genetic screen with a positive selection strategy for the isolation of *A. niger* mutants with constitutive expression of genes involved in PGA degradation resulted in the identification of a repressor protein (NRRL3_08194), which we named GaaX. Both the

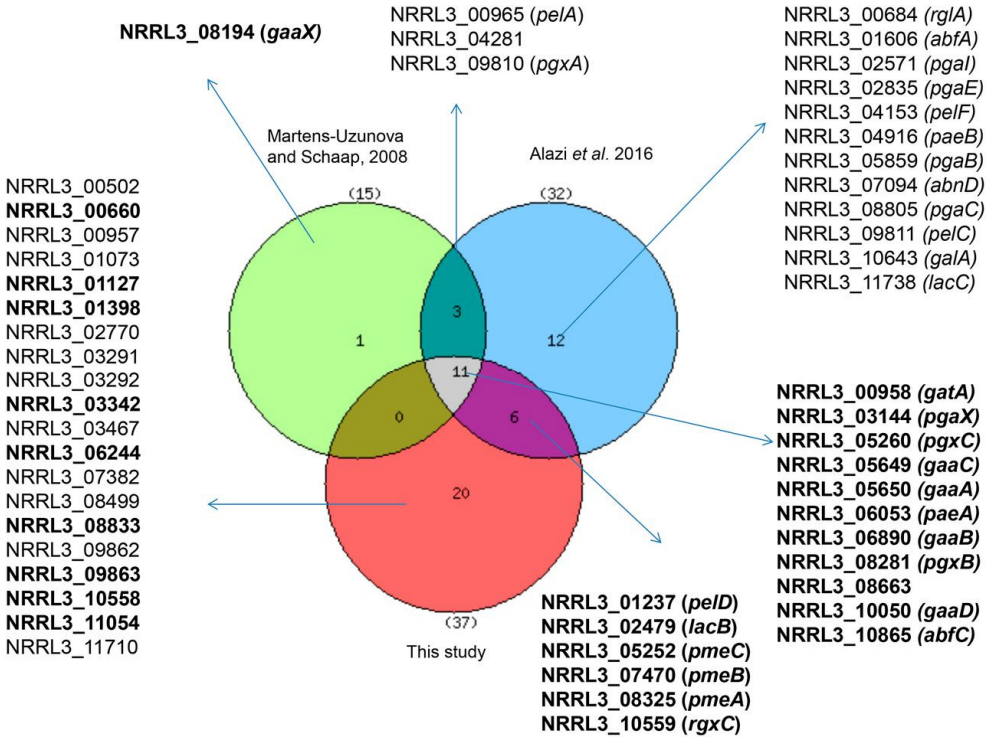

Figure 3. Venn diagram showing the overlaps between upregulated genes in the wt_fructose vs. wt_GA study (Martens-Uzunova and Schaap 2008), the upregulated genes between *ΔgaaR*_GA vs. wt_GA (Alazi *et al.* 2016), and the upregulated genes in wt_fructose vs. *ΔgaaX*_fructose (this study) to identify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

genome sequencing of five independently obtained mutants, as well as the analysis of a targeted deletion mutant (*ΔgaaX*), showed that the loss of function of *gaaX* leads to constitutive expression of genes previously identified as GA-induced genes (Martens-Uzunova and Schaap 2008) and genes encoding pectinolytic enzymes that are activated via the transcription factor GaaR (Alazi *et al*. 2016). Deletion of *gaaX* did not result in a growth alteration on any carbon source tested (Figure 2 and data not shown). Transcriptome analysis (Table S3) strongly suggests that deletion of *gaaX* only affects the expression of genes related to the degradation and metabolism of (poly)galacturonic acid. Genes encoding enzymes involved in the hydrolysis of nonpectin polysaccharides are not differentially regulated in *ΔgaaX*. In addition, GO enrichment analysis of *ΔgaaX* transcriptome shows a strong correlation only between the activity of GaaX and the expression of GA-induced genes. In agreement with these observations, the phenotype of the *gaaR* deletion mutant was specific for (poly)galacturonic acid, with no growth defect observed on other substrates tested (glucuronic acid, rhamnose, xylose, and arabinose) (Alazi *et al*. 2016). Taken together, these findings indicate that GaaR and GaaX are specifically involved in the regulation of pectin catabolism.

Interestingly, the *gaaX* gene is located next to the recently identified GA-specific transcriptional activator *gaaR* (NRRL3_08195). The GaaR transcriptional activator is

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

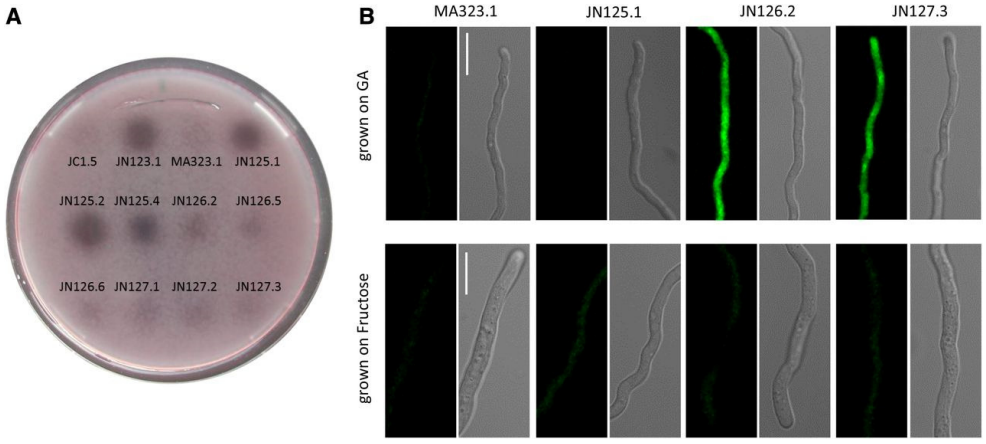

Figure 4. (A) Complementation analysis of GaaX-GFP fusions. Polygalacturonase activities of *gaaX* deletion strains, *gaaX-GFP* and *GFP-gaaX* complementation strains, and their parental strains were detected by spotting 50 ml supernatant from fructose-grown cultures on polygalacturonic acid agarose.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

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

The GaaR transcriptional activator has previously been reported to be conserved in other Ascomycetes belonging to the Pezizomycotina subdivision, including members of the Eurotiomycetes (*Penicillium*, *Talaromyces* spp.), Leotiomycetes (*Botrytis*, *Oidiodendron*), Sordariomycetes (*Neurospora*, *Myceliophthora*, *Magnaporthe*, *Trichoderma*, and *Fusarium* spp.), and Dothideomycetes (*Zymoseptoria* (Mycosphaerella), *Aureobasidium*, and

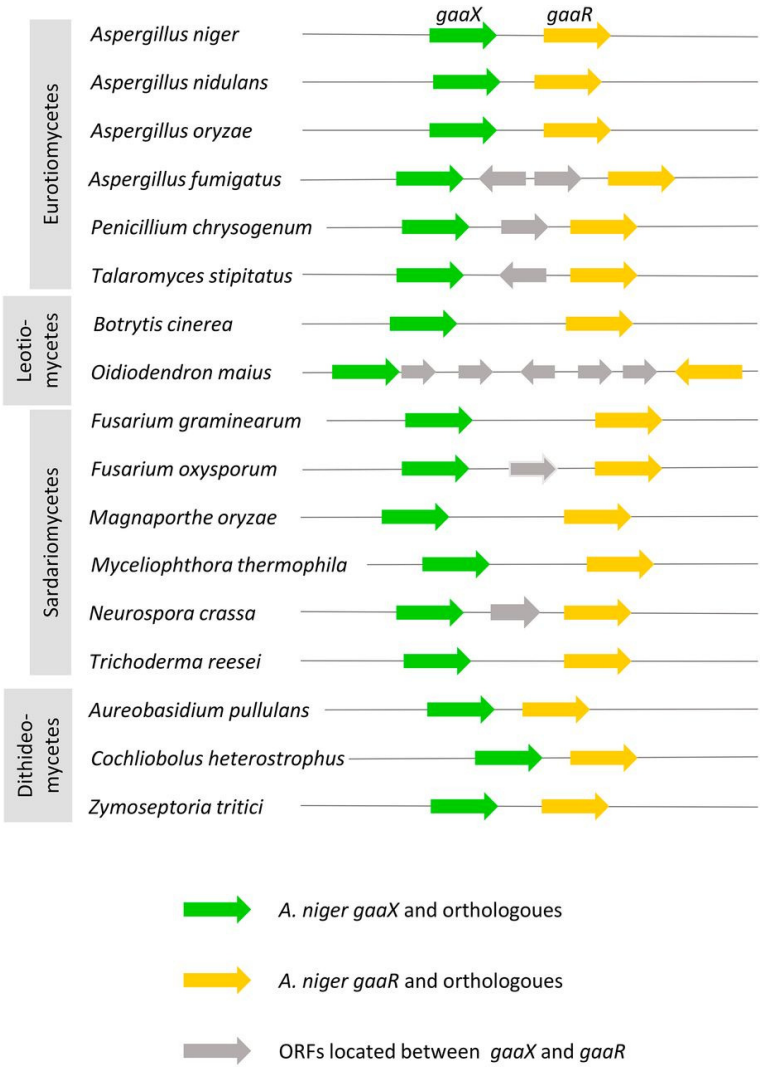

Figure 5. Schematic overview of the conservation of the *gaaX-gaaR* gene pair in 17 *Pezizomycotina* species. GaaX orthologs (green), GaaR orthologs (yellow), and ORFs between *gaaX* and *gaaR* (gray) are indicated. Arrow heads indicate the direction of transcription.

Cochliobolus spp.) (Zhang *et al*. 2016). Synteny analysis of 17 species belonging to four classes of Pezizomycetes (Eurotiomycetes, Leotiomycetes, Sordariomycetes, and Dothideomycetes) revealed a strong conservation of the genomic clustering of *gaaR* and *gaaX* orthologs (Figure 5 and Table S5). For most fungal species analyzed, *gaaR* and *gaaX* are next to each other on the chromosome or close to each other and separated by one to five genes (Figure 5). The head to tail orientation of *gaaR–gaaX* driving expression of *gaaR* and *gaaX* from different promoters is conserved in all species except in *Oidiodendron maius*. Like GaaR, GaaX was found only in the Pezizomycotina and not in ascomycete yeasts, zygomycetes, or basidiomycetes.

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

Previous studies have identified genes specifically induced by GA (Martens-Uzunova and Schaap 2008) and pectinolytic genes that were dependent on the GaaR transcriptional activator for induction by GA (Alazi *et al*. 2016). Eleven of the 15 GA-induced genes identified by Martens-Uzunova and Schaap were upregulated in the *gaaX* mutant (Figure 3 and Table 3). The three genes that are considered GA-inducible but not detected as differentially expressed in the *gaaX* mutant are predicted to encode a transporter (NRRL3_04281), an exo-polygalacturonase (NRRL3_09810, *pgxA*), and a pectin lyase (NRRL3_00965, *pelA*). These three genes were not classified as differentially expressed according to the stringent statistical settings in our current study. The fourth gene induced on GA in the study of Martens-Uzunova and Schaap (2008), but missing in our study, is *gaaX* itself.

In our recent study on the GaaR transcriptional activator, we identified 32 pectinolytic genes whose expression on GA was dependent on GaaR (Alazi *et al*. 2016). These genes overlap largely with the previously identified GA-responsive genes (Martens-Uzunova and Schaap 2008) (Figure 3 and Table 3), but also include 18 new potential GaaR target genes. Six of these genes (including NRRL3_02479 (*lacB*) NRRL3_05252 (*pmeC*), NRRL3_08325 (*pmeA*), NRRL3_07470 (*pmeB*), NRRL3_10559 (*rgxC*), and NRRL3_01237 (*pelD*)) were also found to be significantly upregulated in *ΔgaaX* (Figure 3 and Table 3) and are therefore considered to be part of the core GA regulon. The remaining 12 genes identified as being GaaR dependent for induction on GA (Alazi *et al*. 2016) were not identified as differentially expressed based on the stringent settings in this study. Whether these genes are indeed directly controlled by GaaR and GaaX, and therefore part of the core GA regulon, awaits further study.

The GaaX protein is predicted to be 697 aa long and displays significant similarity to the last three domains in the C-terminal half of the AROM protein. AROM is a large (1586 aa in *A. niger*) pentafunctional protein composed of five domains and the individual domains are involved in five different enzymatic steps representing the prechorismate shikimate pathway, which is required for aromatic amino acid biosynthesis (Duncan *et al*. 1987; Hawkins and Smith 1991).The last three domains of the AROM protein encode the shikimate kinase, 3-dehydroquinate dehydratase, and shikimate dehydrogenase and are

homologous to the respective bacterial enzymes (*aroL*, *aroD*, and *aroE*) (Lamb *et al*. 1996). The AROM protein is present in fungi, including yeasts, and Euglena. The evolutionary origin of AROM is likely to be bacterial and it has been suggested that the AROM protein is the result of gene fusion events (Richards *et al*. 2006). Sequence alignment and BLASTP searches showed that the GaaX protein has significant sequence homology with the last three domains of the AROM protein. The observation of a transcriptional activator (GaaR) located next to a possible repressor protein (GaaX) that displays significant homology to AROM is analogous to the clustered transcriptional activator/repressor module regulating quinic acid utilization (Geever *et al*. 1989; Lamb *et al*. 1990). Like GaaX, the quinate repressor protein shows significant sequence similarities with the last three C-terminal domains of AROM (Lamb et al. 1996).

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

As noted earlier, *gaaX* (NRRL3_08194) was identified as an upregulated gene when an *A. niger* culture pregrown for 18 hr with 2% fructose was transferred to a medium containing 1% GA as the sole carbon source (Martens-Uzunova and Schaap 2008). The expression of a functional GFP-tagged version of GaaX confirmed the induced expression and showed cytosolic localization of GaaX in the presence of GA (Figure 4). In the promoter region of *gaaX*, a GA-responsive element (GARE) was found, suggesting that activation of the transcription factor results in increased levels of repressor protein. Although this might seem contradictory at the first glance, it could be an elegant mechanism to ensure that the expression of GA-induced genes is tightly controlled and quickly responds to the presence or absence of GA. The induction of the expression of the repressor is partially

analogous the activation/repression system of the qa cluster in *N. crassa*. In *N. crassa* it has been shown that both the activator (*qa-1F*) and the repressor (*qa-1S*) are transcriptionally induced in the presence of quinic acid (Patel *et al*. 1981; Giles *et al*. 1991). In the GA regulation system of *A. niger*, only the repressor protein is induced and not the activator. It should be noted that in almost all of 17 species analyzed, the *gaaX* and *gaaR* genes do not share the same promoter region (head to tail orientation; Figure 5), while the *qa-1S* and *qa-1F* genes of *N. crassa* share the same promoter region, which might function as a bidirectional promoter.

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

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

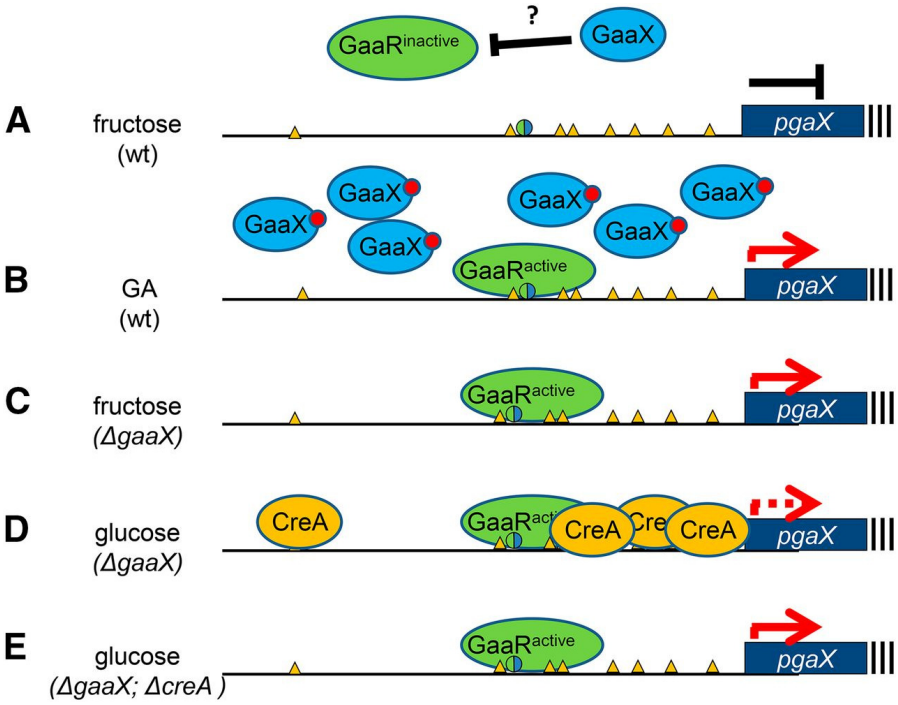

Figure 6. Model for the regulation of GA-induced gene expression in *A. niger*. **(A)** GA-induced gene expression, with *pgaX* as an example, is controlled via interaction of the transcriptional activator (GaaR) and transcriptional repressor (GaaX) in combination with CreA-mediated carbon catabolite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

to AROM. Based on these observations, we suggest that the GAL repressor module has evolved independently from that of GaaX/QutR.

In addition to GaaX (NRRL3_08194) and QutR (NRRL3_11039), we identified two additional paralogues in the *A. niger* genome (NRRL3_08276 and NRRL3_07605). All four paralogues showed significant homology to the *A. niger* AROM protein, as well as limited homology towards each other. Both NRRL3_08276 and NRRL3_07605 are also located next to predicted $Zn(II)_2Cys_6$ domain transcription factors, NRRL3_08275 and NRRL3_07604, respectively. Whereas the function of the GaaR/GaaX and QutA/QutR modules are related to GA and quinic acid metabolism, respectively, the function of the two other pairs that are present in *A. niger* remains to be elucidated. The sequence similarity of NRRL3_08276

and NRRL3_07605 to QutR and GaaX and their genome clustering with predicted transcription factors suggest that the proposed activator/repressor modules observed for GaaR–GaaX and QutA-QutR is an evolutionarily conserved mechanism to control gene expression in filamentous ascomycete fungi. The number of similar activator/repressor modules varies among Pezizomycotina species (Figure S4 and Figure S5). Most Pezizomycotina species contain the galacturonic acid and quinic acid related transcriptional activator/repressor modules. It is interesting to note that some fungi, e.g., *Talaromyces stipitatus* and *B. cinerea*, seem to have lost the quinic acid specific repressor, which suggests they might have lost the capacity to utilize quinic acid. The GaaR/GaaX and QutA/QutR activator/repressor modules and their variants are specific for Pezizomycotina and missing in ascomycete yeasts, zycomycetes, and basidiomycetes.

NOTE AFTER PUBLICATION

In RNA-seq experiments, two or three biological replicates representing each condition were used, and differential gene expression was identified using DESeq2 (FDR≤0.001 and FC≥4). This approach allows addressing only major differences in gene expression.

ACKNOWLEDGEMENTS

We thank Cees van den Hondel for helpful discussions and Tim Knetsch and Jos Reijngoud for assistance with bioreactor cultivations. J.N. is supported through a grant from the China Scholarship Council. E.A. is supported by a grant from BE-Basic within flagship 10 Advanced Microbial Biofuel and chemicals production (AMBIC). This work was in part supported by Genome Canada and Génome Québec.

SUPPORTING INFORMATION

Supplemental material is available online at www.genetics.org/lookup/suppl/doi:10.1534/genetics.116.194050/-/DC1.

LITERATURE CITED

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

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

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

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

Bennett, J. W., and L. Lasure, 1991 Growth media, pp. 441–458 in More Gene Manipulations in Fungi, edited by J. W. Bennett, and L. L. Lasure. Academic Press, San Diego.

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

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

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

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

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

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

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

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

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

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

Grant, S., C. F. Roberts, H. Lamb, M. Stout, and A.R. Hawkins 1988 Genetic regulation of the quinic acid utilization (QUT) gene cluster in *Aspergillus nidulans*. J. Gen. Microbiol. 134: 347–358.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Patro, R., G. Duggal, M. I. Love, R. A. Irizarry, and C. Kingsford, 2016 Salmon provides accurate, fast, and bias-aware transcript expression estimates using dual-phase inference. bioRxiv 021592.

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

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

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

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

Sambrook, J., and D. W. Russell, 2001 Molecular Cloning: A Laboratory Manual, 3rd Ed. Cold Spring Harbor Press, New York.

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

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

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

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

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

vanKuyk, P. A., J. A. Benen, H. A. Wösten, J. Visser, 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.

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

Yuan, X. L., J. A. Roubos, C. A. van den Hondel, and A. F. Ram, 2008b Identification of InuR, a new $Zn(II)_2Cys_6$ transcriptional activator involved in the regulation of inulinolytic genes in *Aspergillus niger*. Mol. Genet. Genomics 279: 11–26.

Zhang, L., R. J. Lubbers, A. Simon, J. H. Stassen, P. R. Vargas Ribera et al., 2016 A novel Zn_2Cys_6 transcription factor BcGaaR regulates D-galacturonic acid utilization in *Botrytis cinerea*. Mol. Microbiol. 100: 247–262.

CHAPTER 5

The pathway intermediate 2-keto-3-deoxy-L-galactonate mediates the induction of genes involved in D-galacturonic acid utilization in *Aspergillus niger*

Alazi E, Khosravi C, Homan TG, du Pré S, Arentshorst M, Di Falco M, Pham TTM, Peng M, Aguilar-Pontes MV, Visser J, Tsang A, de Vries RP, Ram AFJ

Published in *FEBS Letters* (2017) 591:1408-1418

ABSTRACT

In *Aspergillus niger*, the enzymes encoded by *gaaA*, *gaaB*, and *gaaC* catabolize D-galacturonic acid (GA) consecutively into L-galactonate, 2-keto-3-deoxy-L-galactonate, pyruvate, and L-glyceraldehyde, while GaaD converts L-glyceraldehyde to glycerol. Deletion of *gaaB* or *gaaC* results in severely impaired growth on GA and accumulation of L-galactonate and 2-keto-3-deoxy-L-galactonate, respectively. Expression levels of GA-responsive genes are specifically elevated in the *ΔgaaC* mutant on GA as compared to the reference strain and other GA catabolic pathway deletion mutants. This indicates that 2-keto-3-deoxy-L-galactonate is the inducer of genes required for GA utilization.

INTRODUCTION

Pectins are heterogeneous plant cell wall polysaccharides rich in D-galacturonic acid (GA). They represent a natural carbon source for many saprotrophic fungi including Aspergillus *niger* [1,2]. The *A. niger* genome contains 58 genes encoding pectin-degrading enzymes [2,3]. GA, the most abundant uronic acid in pectin, is transported by *A. niger* into the cell via the transporter GatA [4] and then catabolized into pyruvate and glycerol by consecutive action of four enzymes: GaaA, D-galacturonate reductase; GaaB, L-galactonate dehydratase; GaaC, 2-keto-3-deoxy-L-galactonate aldolase; and GaaD, L-glyceraldehyde reductase [5–8] (Fig. 1A). This four-step GA catabolic pathway is evolutionarily conserved in Pezizomycotina fungi [5], and has been studied in detail in *Botrytis cinerea* [9] and *Trichoderma reesei* [10–13]. In *B. cinerea*, the first enzymatic step is catalyzed by two functionally redundant enzymes, BcGar1 and the *A. niger* GaaA ortholog BcGar2 [9]. In *T. reesei*, GA is converted into L-galactonate by TrGar1 [10]. In addition, GaaA and GaaD (LarA) of *A. niger* have been shown to be involved in D-glucuronate and L-arabinose catabolism, respectively [14,15].

Degradation of plant cell wall polysaccharides and subsequent transport and catabolism of released sugars are tightly controlled [16]. Genes required for pectin degradation, GA transport, and GA catabolism are subject to carbon catabolite repression via CreA [17,18]. They are specifically induced in the presence of GA [5,17] and are regulated by the GaaR/GaaX activator-repressor module [19,20]. The conserved $Zn(II)_2Cys_6$ transcription factor GaaR is required for growth on GA and for the activation of the GA-responsive genes in both *B. cinerea* and *A. niger* [19,21].

The mechanism of activation of transcription factors can be diverse, and possibly requires so-called inducer molecules. These inducer molecules are often metabolites related to the substrate [22]. Only a few examples of activation of a transcription factor via an inducer have been elucidated in fungi. Probably the best studied example is the $Zn(II)_2Cys_6$ transcription factor Gal4p that regulates galactose utilization in *Saccharomyces cerevisiae*. Gal4p is repressed under noninducing conditions because the transcriptional activation domain of Gal4p is bound to the corepressor Gal80p. In the presence of galactose and ATP (inducing conditions), the sensor protein Gal3p binds to the Gal4p/Gal80p complex leading to dissociation of Gal4p and subsequent Gal4p-dependent transcription [23–27]. In the regulation of leucine biosynthesis, the $Zn(II)_2Cys_6$ transcription factor Leu3p interacts directly with a metabolic intermediate. The middle domain of the Leu3p protein masks the C-terminal activation domain by an intramolecular interaction in the absence of α-isopropylmalate (α-IPM), a metabolic intermediate of the leucine biosynthesis pathway. In the presence of α-IPM, which accumulates during leucine starvation, this self-masking is prevented, resulting in active Leu3p and activation of leucine biosynthesis genes [28–30]. The Gal4p and Leu3p transcription factors localize to the nucleus regardless of the presence or absence of inducer molecules [31,32]. On the other hand, the transcriptional activator AmyR, involved in starch degradation in *Aspergillus nidulans* and *Aspergillus oryzae*, is translocated from the cytoplasm to the nucleus only in the presence of its inducer isomaltose [33–35].

In *A. niger*, GA or a derivative of GA was suggested to act as an inducer required for the activation of GA-responsive genes [17]. In *B. cinerea*, BcGaaR was shown to translocate

from the cytoplasm to the nucleus in response to such an inducer [21]. Previous studies of *A. niger* and *B. cinerea* mutants disrupted in GA catabolic pathway did not unambiguously identify a specific inducer [6–9]. In this study, we constructed GA catabolic pathway deletion mutants (*ΔgaaA*, *ΔgaaB*, *ΔgaaC*, and *ΔgaaD*) to gain insight into regulation of GA-responsive genes in *A. niger*. Comparative analysis of these mutants indicates that 2-keto-3-deoxy-L-galactonate acts as the physiological inducer of the GA-responsive genes.

MATERIALS AND METHODS

Strains, media and growth conditions

All strains used in this study are listed in Table S1. MA249.1 was obtained by transformation of N593.20 (*cspA1*, *pyrG-*, *kusA::amdS*) [19] with a 3.8-kb *Xba*I fragment containing the *A. niger pyrG* gene, resulting in the full restoration of the *pyrG* locus.

Media were prepared as described previously [36]. Radial growth phenotype analyses were performed with minimal medium (MM) (pH 5.8) containing 1.5% (w/v) agar (Scharlau, Barcelona, Spain) and various carbon sources: 50 mM glucose (VWR International, Amsterdam, the Netherlands), D-fructose (Sigma-Aldrich, Zwijndrecht, the Netherlands), GA (Chemodex, St Gallen, Switzerland), L-rhamnose (Fluka, Zwijndrecht, the Netherlands), L-arabinose (Sigma-Aldrich) or glycerol (Glycerol 87%; BioChemica AppliChem, Darmstadt, Germany), or 1% (w/v) polygalacturonic acid (PGA) (Sigma), apple pectin (AP) (Sigma-Aldrich), or galactan (Acros Organics, Geel, Belgium). Filter sterilized D-fructose or GA solution was added after autoclaving MM with agar. Other carbon sources were autoclaved together with the medium. The plates were inoculated with 5 µL 0.9% NaCl containing 10^4 freshly harvested spores and cultivated at 30 °C for 7 days. For microtiter plate growth phenotype analysis, wells in a 96-well, flat bottom plate (Sarstedt AG & Co., Nümbrecht, Germany) were filled with 180 µL MM (pH 5.8) containing 55 mM GA as the sole carbon source, and 20 µL freshly harvested spores ($7.5x10^5$ spores mL^{-1}). The plate was incubated with lids in EnSpire Multimode Plate Reader (PerkinElmer, Waltham, MA, USA) at 30 °C. Lid temperature was set to 32 °C to prevent condensation on the lid. Optical density at 600 nm was measured every hour. The average OD from the GA-containing control wells was subtracted from the OD of the test wells and negative values were corrected as zero.

For gene expression and metabolic analyses, 10^8 freshly harvested spores were inoculated and grown in 100 mL complete medium (CM) (pH 5.8) with 2% (w/v) D-fructose for 16 h, and mycelia were harvested by filtration through sterile myracloth. For northern blot and metabolic analyses, pregrown mycelia were washed twice with MM with no carbon source (pH 4.5) and 1.5 g (wet weight) mycelia were transferred and incubated in 50 mL MM (pH 4.5) with 50 mM D-fructose or 50 mM GA for 2 h. For metabolic analysis, 1.5 g (wet weight) mycelia were transferred and incubated in 50 mL MM (pH 4.5) with 50 mM GA for 55 h. Additionally, 30 g (wet weight) mycelia of SDP20.6 (*ΔgaaC*) were transferred and incubated in 1 L MM (pH 4.5) with 50 mM GA for 55 h. For RNA-seq analysis, pregrown mycelia were washed with MM with no carbon source (pH 6) and 2.5 g (wet weight) were transferred to 50 mL MM (pH 6) with 25 mM GA and grown for 2 h. All incubations were carried out in a rotary shaker at 30 °C and 250 r.p.m.

Construction of gene deletion strains

Protoplast-mediated transformation of *A. niger*, purification of the transformants and genomic DNA extraction were performed as described [36].

To construct the deletion cassettes, 50 and 30 flanks of the *gaaA*, *gaaB*, *gaaC*, and *gaaD* genes were PCR-amplified using the primer pairs listed in Table S2 with N402 genomic DNA as template. For all cloning experiments *Escherichia coli* strain DH5*α* was used. To create SDP22.1 (*ΔgaaA*), SDP21.5 (*ΔgaaB*), and SDP20.6 (*ΔgaaC*), gene deletion cassettes were made using MultiSite Gateway Three-fragment Vector Construction Kit (Invitrogen, Carlsbad, CA, USA) according to the supplier's instructions. *Aspergillus oryzae pyrG* gene flanked by AttB1 and AttB2 sites was amplified by PCR using the primer pair listed in Table S2 and plasmid pMA172 [37] as template. *gaaA*, *gaaB*, and *gaaC* deletion cassettes containing 5' and 3' flanks of the target genes with *A. oryzae pyrG* gene in between were obtained by restriction digestion. To create EA1.1 (*ΔgaaD*), 5' flank of *gaaD* was ligated into pJET1.2/ blunt cloning vector (Thermo Fisher Scientific, Carlsbad, CA, USA) and amplified in *E. coli*. Following plasmid isolation, the 5' flank was excised using restriction enzymes *Kpn*I and *Xho*I, ligated into *Kpn*I-*Xho*I opened pBluescript II SK(+) (Agilent Technologies, La Jolla, CA, USA) and amplified in *E. coli*. *Aspergillus oryzae pyrG* gene was obtained from plasmid pMA172 [37] by restriction digestion with *Hin*dIII and *Xho*I. Isolated pBluescript II SK(+) plasmid containing the 5' flank was opened with restriction enzymes *Xho*I and *Not*I, and the *A. oryzae pyrG* gene as *Xho*I-*Not*I fragment and *Hin*dIII-*Not*I fragment of the *gaaD* 3' flank were ligated into the plasmid. Ligation product was amplified in *E. coli* and the linear deletion cassette was obtained by PCR amplification from the plasmid using primers gaaDP1-KpnI and gaaDP4-NotI. Deletion cassettes were introduced into the *pyrG*$^{-}$ strain N593.20. Gene deletions were confirmed via southern blot analysis.

Gene expression analysis

Northern blot and RNA-seq analyses were performed as described [19] with minor modifications: For northern blot analysis, total RNA was extracted using TRIzol reagent (Life Technologies, Carlsbad, CA, USA). Probes were PCR-amplified using the N402 genomic DNA and the primer pairs listed in Table S2.

Chemical analysis

One milliliter culture samples were taken 7, 24, 31, 48, and 55 h after the transfer of mycelia to MM with GA. About 250 µL of each culture sample was centrifuged at 16 000 g for 30 min and the supernatant was transferred to a new microfuge tube. After adding 1xvolume of cold methanol (-20 °C), the sample was incubated on ice for 15 min and centrifuged at 16 000 g for 30 min. The supernatant was collected in a new microfuge tube and 1xvolume of 0.1% formic acid was added. Metabolites in the extracellular culture fluids were analyzed by high pressure liquid chromatography-high-resolution mass spectrometry. Aliquots were loaded, using a Series 200 micropump (PerkinElmer), onto a reversed-phase Eclipse C18 2.1x150 mm column (Agilent, Santa Clara, CA, USA) connected

in-line to a 7 Tesla LTQ-FT-ICR mass spectrometer (Thermo Electron Corporation, San Jose, CA, USA) and negative mode electrospray ionization spectra were acquired at a resolution of 100 000 at 200 *m/z*. Absolute GA concentration was calculated using a standard dilution calibration curve of commercially obtained GA (Chemodex). Standards for L-galactonate and 2-keto-3-deoxy-L-galactonate were not available, therefore, these metabolites were assigned based on accurate mass alone (matched within a 5 p.p.m. *m/z* window) and relative amounts in terms of extracted ion chromatograms peak areas were compared. One liter culture of SDP20.6 (*ΔgaaC*) was filtered through sterile myracloth 55 h after the transfer of mycelia to MM with GA, and the filtrate was stored at -80 °C. After freeze-drying, dry materials from SDP20.6 (*ΔgaaC*) extracellular culture fluid were dissolved in D_2O (Sigma Aldrich) for structural investigation by Nuclear Magnetic Resonance Spectroscopy (NMR). Spectra were recorded with a Varian VNMRS- 500 MHz at 25 °C. The presence of 2-keto-3-deoxy-L-galactonate was confirmed by ^{1}H-NMR and ^{13}C-NMR.

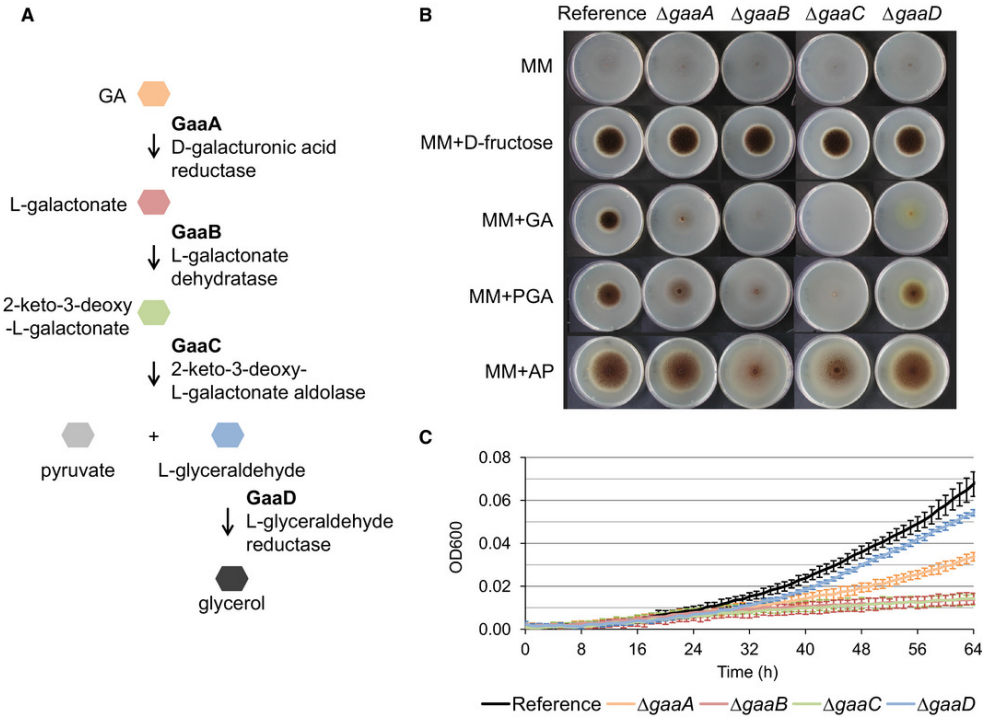

Fig. 1. (A) The evolutionarily conserved GA catabolic pathway in filamentous fungi as proposed by Martens-Uzunova and Schaap [5]. GA is converted in pyruvate and glycerol by consecutive action of GaaA, GaaB, GaaC, and GaaD enzymes. Growth profile of the reference strain (MA249.1) and GA catabolic pathway deletion mutants *ΔgaaA*, *ΔgaaB*, *ΔgaaC*, and *ΔgaaD* **(B)** on solid MM without any carbon source, or with 50 mM monomeric or 1% polymeric carbon sources after 7 days at 30 °C, and **(C)** in microtiter plate in liquid medium with 50 mM GA at 30 °C. Error bars represent standard deviation of six biological replicates.

Bioinformatics

RNA-seq data were analyzed as described previously [19]. Differential expression was identified by Student's *t*-test with a *P*-value cut-off of 0.05. RNA-seq data for FP-1132.1 (reference strain) and SDP20.6 (*ΔgaaC*) were submitted to Gene Expression Omnibus [38] with accession numbers GSE80227 [19] and GSE95776 (this study), respectively.

RESULTS

Growth analysis of D-galacturonic acid catabolic pathway deletion mutants

Aspergillus niger GA catabolic pathway deletion mutants, *ΔgaaA*, *ΔgaaB*, *ΔgaaC*, and *ΔgaaD*, were constructed and were verified by southern blot analysis (Fig. S1). We compared the growth phenotype of the strains on monomeric and polymeric carbon sources (Fig. 1, Fig. S2). Disruption of *gaaA* and *gaaD* resulted in reduced growth and sporulation on plates containing GA or PGA as carbon source. However, both mutants showed better growth on plates containing MM with GA compared to plates containing MM with no carbon source, indicating that they can still metabolize GA. The *ΔgaaB* and *ΔgaaC* mutants showed a more drastically reduced growth on plates containing GA, PGA, or AP (Fig. 1B). The growth defects of the GA catabolic pathway deletion mutants on GA plates were confirmed in microtiter plate-based growth assays (Fig. 1C, Fig. S2A). None of the GA catabolic pathway deletion mutants exhibited defects in growth on other carbon sources tested, except that the deletion of *gaaD*, also known as the L-arabinose reductase gene *larA*, resulted in a poor growth on L-arabinose (Fig. S2B), confirming previous observations [15]. The inability of *ΔgaaB* or *ΔgaaC* to use GA as a carbon source suggests that there are no functionally redundant enzymes capable of replacing GaaB and GaaC.

ΔgaaB and *ΔgaaC* accumulate the D-galacturonic acid catabolic pathway intermediates L-galactonate and 2-keto-3-deoxy-L-galactonate, respectively

Since the roles of GaaB and GaaC in GA catabolism cannot be replaced by redundant enzymes, we expect the accumulation in the medium of the corresponding enzyme substrate in *ΔgaaB* and *ΔgaaC*, as shown previously [7,8]. The extracellular GA concentration and the extracellular metabolites were examined by FTICR mass spectrometry over time during growth in GA. This analysis revealed that the reference strain utilized all GA in the medium within 48 h of growth, whereas in the GA catabolic pathway deletion mutants GA was still present in the medium after 55 h of growth (Fig. 2A). In *ΔgaaA* and *ΔgaaD*, the concentration of GA gradually decreased to approximately 45% of the initial GA concentration in the medium, which reflects the slow catabolism of GA in these mutants. *ΔgaaB* consumed about 35% of the initial GA in 55 h and secreted L-galactonate. The time course consumption of GA by *ΔgaaB* was proportional to its release of L-galactonate (Fig. 2A). The *ΔgaaC* mutant took up about 78% of the initial GA in 55 h, and extracellular 2-keto-3-deoxy-L-galactonate accumulated in the medium of the *ΔgaaC* mutant over time (Fig. 2A). The presence of 2-keto-3-deoxy-L-galactonate in the extracellular culture fluid of the *ΔgaaC* mutant was confirmed by structural resolution by ^{1}H-NMR and ^{13}C-NMR (Fig. S3).

Expression of D-galacturonic acid-responsive genes is increased in *ΔgaaC*

Genes involved in the degradation of the pectic substructures PGA (e.g., NRRL3_03144 exo-polygalacturonase and *pgx28B*) and rhamnogalacturonan I (RG-I) (e.g., NRRL3_10865 alpha-N-arabinofuranosidase), GA transport (*gatA*), and GA catabolism (*gaaA-D*) have

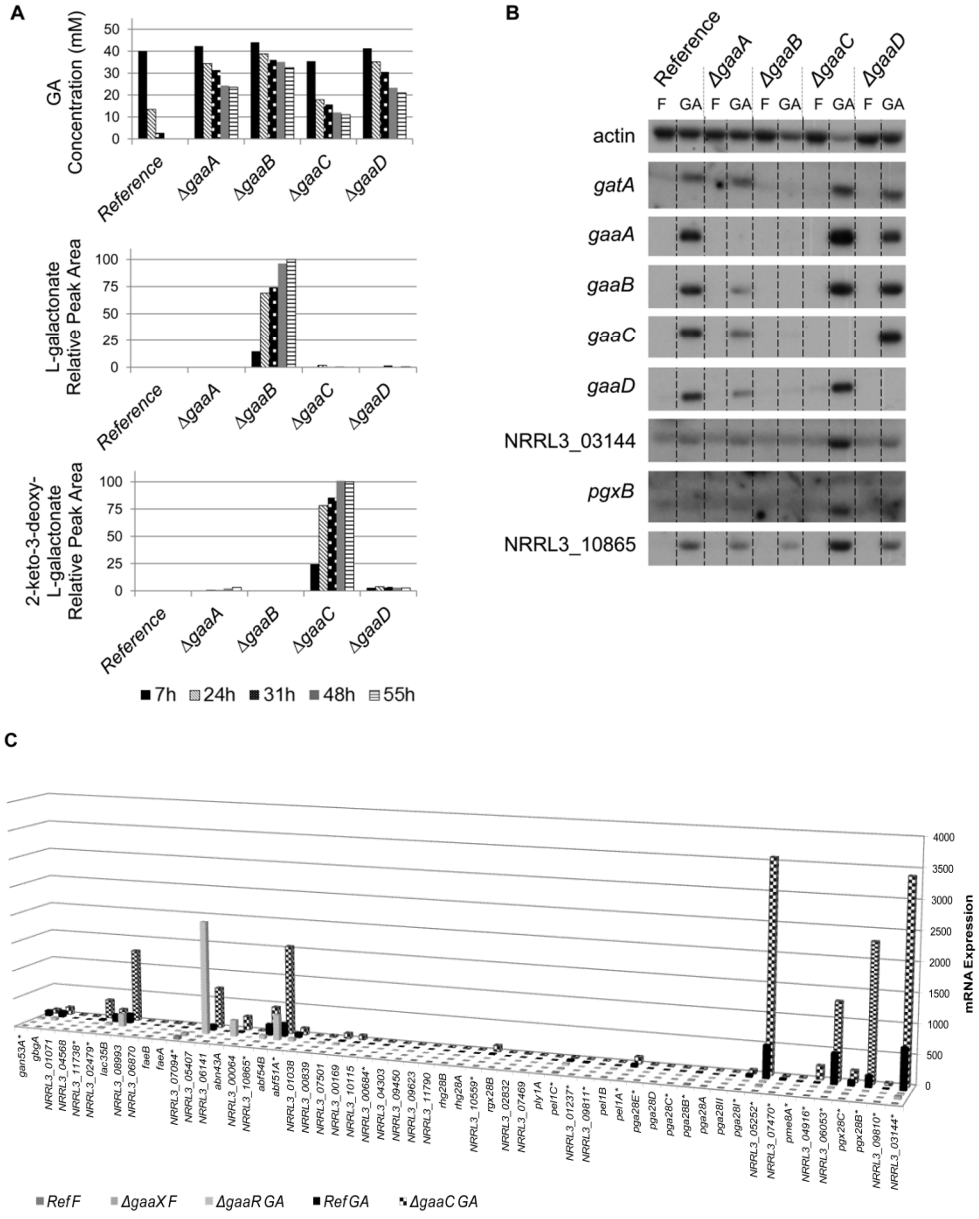

Fig. 2. Metabolic and gene expression analyses of *Aspergillus niger* GA catabolic pathway deletion mutants *ΔgaaA*, *ΔgaaB*, *ΔgaaC*, and *ΔgaaD* **(A)** Extracellular GA, L-galactonate, and 2-keto-3-deoxy-L-galactonate concentration in cultures of the reference strain (FP-1132.1) and GA catabolic pathway deletion mutants. GA concentration is given in mM and L-galactonate and 2-keto-3-deoxy-L-galactonate amounts are presented as ion chromatogram peak areas relative to *ΔgaaB* 55 h and *ΔgaaC* 55 h samples, respectively. **(B)** Northern blot analysis of selected GA-responsive genes in the reference strain (MA249.1) and GA catabolic pathway deletion mutants. Actin (NRRL3_03617) was used as a control. **(C)** RNA-seq analysis of pectinase genes in the reference strain (FP-1132.1) and *ΔgaaC* in GA (FPKM). Expression in *ΔgaaR* in GA (FPKM) [19] and in the reference strain (MA234.1) and *ΔgaaX* in D-fructose (TPM) [20] was shown for comparison. Pectinase genes that belong to the GaaR/GaaX panregulon [20] are indicated with an asterisk. Strains were pregrown in CM with 2% D-fructose. For metabolic analysis, mycelia were transferred to and grown in MM containing 50 mM GA. For northern blot analysis, mycelia were transferred to and grown in MM containing 50 mM D-fructose (F) or GA for 2 h. For RNA-seq analysis, mycelia were transferred to and grown in MM containing 25 mM GA for 2 h.

been shown to be induced in the presence of GA [5,18] and are part of the proposed GaaR/GaaX-controlled gene regulon [20]. To test the effect of the GA catabolic pathway gene deletions on the induction of GA-responsive genes, northern blot analysis was performed. The reference and *ΔgaaA*, *ΔgaaB*, *ΔgaaC* and *ΔgaaD* strains were pregrown in D-fructose medium and transferred to either GA or D-fructose medium. Rapid induction of *gatA*, *gaaA*, *gaaB*, *gaaC*, *gaaD*, and NRRL3_10865 was observed in the reference strain upon transfer from D-fructose to GA as expected (Fig. 2B). Induction of these genes upon transfer to GA was also found in *ΔgaaA*, but at lower levels compared to the reference strain. The induction of GA-responsive genes was nearly absent in *ΔgaaB*. As shown in Fig. 2B, deletion of *gaaC* resulted in a hyperinduction of GA-responsive genes, especially pectinases (NRRL3_03144, *pgx28B*, and NRRL3_10865). Expression of *gatA*, *gaaA*, *gaaB*, *gaaC*, and the pectinases in *ΔgaaD* was similar to the expression in the reference strain (Fig. 2B).

Transcriptome analysis of *ΔgaaC*

In order to analyze the expression of a larger number of genes controlled by GaaR/GaaX activator–repressor module in *ΔgaaC*, a genome-wide gene expression analysis was performed using RNA-seq. The reference strain and the *ΔgaaC* mutant were pregrown in D-fructose medium and transferred to GA medium. Seventeen of the 53 GaaR/GaaX panregulon genes were significantly upregulated (FC ≥ 2 and *P*-value ≤ 0.05) in the *ΔgaaC* mutant cultured in GA as compared to the reference strain (Table 1, Table S3). These 17 genes include *gaaA* and 6 pectinases (NRRL3_03144, *pgx28B*, NRRL3_05252, NRRL3_04916, NRRL3_10559, and NRRL3_11738), as well as genes encoding four transporters and six genes for which the function has not yet been established. The expression of 24 of the remaining GaaR/GaaX panregulon genes was higher in *ΔgaaC* compared to the reference strain, but differences were relatively small and did not pass the stringent *P*-value of ≤ 0.05.

In addition to GaaR/GaaX-controlled genes, we also compared the expression of all 58 pectinases identified in the genome of *A. niger* [2] between the reference strain and the *ΔgaaC* mutant (Table S4, Fig. 2C). Apart from the six pectinases that depend on GaaR for

induction [19], nine additional pectinases acting on the RG-I backbone and arabinan and arabinogalactan side chains were significantly upregulated (FC ≥ 2 and *P*-value ≤ 0.05) in the *ΔgaaC* mutant compared to the reference strain (Table 2). It has been reported that many of these genes are regulated by transcription factors RhaR (NRRL3_02832, NRRL3_07501, NRRL3_07501, and *faeB*), XlnR (NRRL3_05407 and *lac35B*), or AraR (*lac35B*), which are required for the utilization of L-rhamnose, xylan/D-xylose, and arabinan/ L-arabinose, respectively [39–42]. To address the possibility that deletion of *gaaC* affected the expression of these genes via their specific transcription factors, the expression of *rhaR*, *xlnR*, and *araR* was analyzed in more detail. Expression of *rhaR* (FC = 5.84 and P-value = 4.76E-03) and *xlnR* (FC = 2.68 and P-value = 5.60E-03) was significantly higher in *ΔgaaC*, which might explain the upregulation observed in these genes. The *araR* gene was not significantly differentially regulated in the *ΔgaaC* mutant.

DISCUSSION

In this study, we used GA catabolic pathway deletion mutants to investigate the induction mechanism of the GA-responsive genes in *A. niger*. We observed that the *gaaA* and the *gaaD* deletion mutants show reduced growth on GA or PGA compared to the reference strain, whereas growth of *ΔgaaB* and *ΔgaaC* is more severely reduced on GA, PGA, or AP (Fig. 1B,C). These results are in line with the previous reports showing the inability of *ΔgaaB* and *ΔgaaC* to grow on GA [7,8]. *ΔgaaA* was reported to be unable to grow on GA in a previous study [6], where the tenuous growth of *ΔgaaA* could have been interpreted as no growth. GA catabolic pathway deletion mutants derived from N593.20 in this study and from ATCC1015 in previous studies [6–8] showed the same growth defects on GA (unpublished results), excluding the possibility of a phenotypic difference caused by strain background.

Deletion of *gaaB* and *gaaC* severely impaired growth on MM containing GA (Fig. 1B,C), indicating that there are no alternative enzymes replacing GaaB and GaaC. The residual growth of *ΔgaaA* and *ΔgaaD* on GA indicates that GA is catabolized in these reductase deletion mutants via partially redundant enzymes. In *B. cinerea*, there are two nonhomologous D-galacturonate reductases, BcGar1, and BcGar2. While single gene deletion mutants (*ΔBcgar1* or *ΔBcgar2*) could still grow on GA, the double gene deletion mutant *ΔBcgar1 ΔBcgar2* showed a complete loss of growth [9]. *A. niger* also contains a BcGar1 ortholog, NRRL3_06930, which shows no protein homology to GaaA. As in *B. cinerea*, NRRL3_06930 might enable the residual growth of *ΔgaaA* on GA. However, the expression of NRRL3_06930 is considerably lower than the expression of *gaaA* in GA, and unlike the expression of *gaaA*, does not depend on GaaR or GaaX [19,20]. It is also possible that the two dehydrogenases belonging to the GaaR/GaaX panregulon, NRRL3_03342, and NRRL3_09863, partially replace GaaA or GaaD.

The recently proposed model related to the regulation of GA-responsive gene expression [20] postulates that under noninducing conditions the repressor GaaX inhibits the transcriptional activity of GaaR. The repressing activity of GaaX is suggested to be lost in the presence of an inducer and subsequent activation of GaaR, resulting in the induction of GA-responsive genes in *A. niger* [20]. The results of metabolic and northern blot

Table 1. RNA-seq analysis of 53 genes of the GaaR-GaaX panregulon [20] in *ΔgaaC* in GA. 27 genes belonging to GaaR-GaaX core regulon [20] are written in bold. Expression values (FPKM) are averages of duplicates. Significantly upregulated genes (FC≥2 and *P*-value≤0.05) are highlighted.

Gene ID NRRL3	Gene ID CBS513.88	Description[a]	Gene Name	Ref	*ΔgaaC*	FC *ΔgaaC*/Ref	*P*-value
NRRL3_00958	**An14g04280**	**D-galacturonic acid transporter GatA**	***gatA***	888.35	1062.68	1.20	6.95E-02
NRRL3_03144	**An12g07500**	**exo-polygalacturonase**		698.90	3384.63	4.84	1.34E-02
NRRL3_05260	**An02g12450**	**exo-polygalacturonase Pgx28C**	***pgx28C***	99.93	192.85	1.93	9.11E-02
NRRL3_05649	**An02g07720**	**2-keto-3-deoxy-L-galactonate aldolase GaaC**	***gaaC***	5658.32	14.60	0.00	2.88E-04
NRRL3_05650	**An02g07710**	**D-galacturonic acid reductase GaaA**	***gaaA***	2599.98	6710.72	2.58	1.04E-02
NRRL3_06053	**An02g02540**	**carbohydrate esterase family 16 protein**		522.81	1301.08	2.49	8.01E-02
NRRL3_06890	**An16g05390**	**L-galactonate dehydratase GaaB**	***gaaB***	11309.00	13990.90	1.24	1.91E-01
NRRL3_08281	**An03g06740**	**exo-polygalacturonase Pgx28B**	***pgx28B***	200.31	2306.06	11.51	2.82E-02
NRRL3_08663	**An03g01620**	**MFS-type sugar/inositol transporter**		106.09	227.29	2.14	1.71E-01
NRRL3_10050	**An11g01120**	**L-glyceraldehyde reductase GaaD**	***gaaD***	8104.43	7499.78	0.93	5.79E-01
NRRL3_10865	**An08g01710**	**alpha-N-arabinofuranosidase**		201.62	440.98	2.19	1.92E-01
NRRL3_01237	**An19g00270**	**pectin lyase**		18.95	3.68	0.19	9.55E-03
NRRL3_02479	**An01g10350**	**exo-beta-1,4-galactanase**		137.63	170.01	1.24	5.21E-01
NRRL3_05252	**An02g12505**	**pectin methylesterase**		558.37	3569.08	6.39	2.07E-02
NRRL3_07470	**An04g09690**	**pectin methylesterase**		30.16	12.81	0.42	4.22E-02
NRRL3_08325	**An03g06310**	**pectin methylesterase Pme8A**	***pme8A***	6.54	6.74	1.03	8.79E-01
NRRL3_10559	**An18g04810**	**glycoside hydrolase family 28 protein**		20.00	97.18	4.86	1.19E-02
NRRL3_00965	An14g04370	pectin lyase Pel1A	*pel1A*	56.54	113.40	2.01	3.58E-01
NRRL3_04281	An07g00780	MFS-type transporter		90.41	106.00	1.17	5.05E-01
NRRL3_09810	An11g04040	exo-polygalacturonase		10.65	35.99	3.38	7.58E-02
NRRL3_08194	**An04g00790**	**Repressor of D-galacturonic acid utilization**	***gaaX***	381.34	529.21	1.39	1.97E-01
NRRL3_00684	An14g01130	rhamnogalacturonan lyase		5.77	13.23	2.29	2.61E-01
NRRL3_01606	An01g00330	alpha-N-arabinofuranosidase Abf51A	*abf51A*	87.96	111.63	1.27	4.97E-01
NRRL3_02571	An01g11520	endo-polygalacturonase Pga28I	*pga28I*	56.38	59.67	1.06	5.83E-01
NRRL3_02835	An01g14670	endo-polygalacturonase Pga28E	*pga28E*	4.26	13.51	3.17	9.99E-02
NRRL3_04153	An15g07160	pectin lyase		35.48	19.78	0.56	3.56E-02
NRRL3_04916	An07g08940	carbohydrate esterase family 16 protein		13.41	221.16	16.49	4.37E-02
NRRL3_05859	An02g04900	endo-polygalacturonase Pga28B	*pga28B*	15.10	4.12	0.27	9.36E-02
NRRL3_07094	An16g02730	endo-1,5-alpha-arabinanase		4.57	3.48	0.76	2.43E-01

Table 1. RNA-seq analysis of 53 genes of the GaaR-GaaX panregulon [20] in *ΔgaaC* in GA. 27 genes belonging to GaaR-GaaX core regulon [20] are written in bold. Expression values (FPKM) are averages of duplicates. Significantly upregulated genes (FC≥2 and *P*-value≤0.05) are highlighted. (Continued)

Gene ID NRRL3	Gene ID CBS513.88	Description[a]	Gene Name	Ref	*ΔgaaC*	FC *ΔgaaC*/Ref	*P*-value
NRRL3_08805	An05g02440	endo-polygalacturonase Pga28C	*pga28C*	5.26	7.27	1.38	1.85E-01
NRRL3_09811	An11g04030	pectin lyase		0.51	0.11	0.21	6.88E-02
NRRL3_10643	An18g05940	arabinogalactanase Gan53A	*gan53A*	105.64	67.21	0.64	2.70E-01
NRRL3_11738	An06g00290	beta-galactosidase		28.91	319.96	11.07	4.60E-02
NRRL3_00502	An09g06200	hypothetical protein		14.07	41.41	2.94	1.16E-01
NRRL3_00660	**An14g00860**	**carboxylesterase**		74.22	825.36	11.12	4.58E-02
NRRL3_00957	An14g04260	B3/B4 domain-containing protein		7.87	13.03	1.66	2.87E-01
NRRL3_01073	An14g05840	O-methyltransferase, COMT-type		3.22	11.45	3.55	1.39E-02
NRRL3_01127	**An14g06500**	**dihydroxyacetone kinase**		584.25	203.94	0.35	1.55E-02
NRRL3_01398	**An13g02090**	**MFS-type transporter**		26.10	96.31	3.69	1.69E-02
NRRL3_02770	An01g13880	MFS-type transporter		3.71	6.43	1.73	9.57E-02
NRRL3_03291	An12g05600	heterokaryon incompatibility protein		0.80	6.04	7.60	6.39E-02
NRRL3_03292	An12g05590	carboxylesterase		0.25	1.72	6.88	3.30E-01
NRRL3_03342	**An12g04990**	**short-chain dehydrogenase/reductase**		151.58	706.28	4.66	1.05E-02
NRRL3_03467	An12g03550	MFS-type transporter		4.91	92.55	18.85	2.61E-02
NRRL3_06244	**An02g00140**	**glycoside hydrolase family 43 protein**		80.90	137.44	1.70	1.81E-01
NRRL3_07382	An16g00540	alpha-L-fucosidase		2.29	8.06	3.53	4.41E-02
NRRL3_08499	An03g03960	uncharacterized protein		13.64	45.86	3.36	6.05E-03
NRRL3_08833	**n.a.**	**hypothetical protein**		4.29	1.87	0.44	2.27E-02
NRRL3_09862	An11g03510	hypothetical protein		0.43	0.20	0.45	5.62E-01
NRRL3_09863	**An11g03500**	**alpha-hydroxy acid dehydrogenase, FMN-dependent**		59.53	64.98	1.09	2.85E-01
NRRL3_10558	**An18g04800**	**alpha-L-rhamnosidase**		17.04	109.06	6.40	3.54E-02
NRRL3_11054	**An08g04040**	**MFS-type sugar/inositol transporter**		693.37	4713.62	6.80	8.89E-03
NRRL3_11710	An06g00620	MFS-type sugar/inositol transporter		341.35	1977.10	5.79	2.76E-02

[a] Descriptions were obtained from manual annotation (manuscript in preparation).

Table 2. RNA-seq analysis of 9 pectinase genes that were significantly upregulated in *ΔgaaC* in GA and do not belong to the GaaR-GaaX panregulon [20].

Gene ID NRRL3	Gene ID CBS513.88	Description[a]	Gene Name	Ref	*ΔgaaC*	FC *ΔgaaC*/Ref	*P*-value
NRRL3_02832	An01g14650	glycoside hydrolase family 28 protein		1.49	12.95	8.69	1.21E-02
NRRL3_09450	An11g08700	endo-rhamnogalacturonase		1.75	4.34	2.48	3.39E-02
NRRL3_07501	An04g09360	carbohydrate esterase family 12 protein		17.42	87.29	5.01	4.60E-02
NRRL3_00839	An14g02920	glycoside hydrolase family 105 protein		3.61	22.81	6.32	5.98E-03
NRRL3_05407	An02g10550	endo-1,5-alpha-arabinanase		103.20	702.79	6.81	1.45E-02
NRRL3_02931	An12g10390	feruloyl esterase FaeB	*faeB*	4.17	16.38	3.93	3.08E-02
NRRL3_02630	An01g12150	beta-galactosidase Lac35B	*lac35B*	172.89	1259.38	7.28	3.28E-02
NRRL3_04568	An07g04420	exo-beta-1,4-galactanase		0.23	9.58	41.63	7.17E-03
NRRL3_01071	An14g05820	beta-galactosidase		0.75	8.06	10.74	2.90E-02

[a] Descriptions were obtained from manual annotation (manuscript in preparation).

analyses indicate that accumulation of 2-keto-3-deoxy-L-galactonate in *ΔgaaC* is responsible for the induction of the GA-responsive genes. In other words, the pathway intermediate 2-keto-3-deoxy-L-galactonate, and not GA or L-galactonate, is the physiological inducer of the GA-responsive genes in *A. niger* (Fig. 2A,B). In the *ΔgaaA* mutant, we postulate that GA is converted into L-galactonate via partially redundant enzymes (see above) and the 2-keto-3-deoxy-L-galactonate produced is enough for the induction of GA-responsive genes. However, this induction is lower compared to the reference strain (Fig. 2B). This result is supported by a previous finding that *gaaB* and *gaaC* were expressed at lower levels in *ΔgaaA* compared to the reference strain [6]. In contrast, *ΔgaaB* possibly does not produce 2-keto-3-deoxy-L-galactonate from L-galactonate, since the growth phenotype of the *ΔgaaB* mutant suggests that there are no functionally redundant enzymes replacing GaaB. As a result, expression of GA-responsive genes is not induced in *ΔgaaB* (Fig. 2B). Reduced expression of *gatA*, *gaaA*, and *gaaC* in the *ΔgaaB* mutant was also observed previously [7].

RNA-seq analysis of *ΔgaaC* revealed significant upregulation of several genes from the GaaR/GaaX panregulon involved in pectin breakdown and GA utilization, as well as genes with currently unknown link to GA utilization, such as transporters that might facilitate the faster GA transport in *ΔgaaC* compared to other GA catabolic pathway deletion mutants observed both in this study (Fig. 2A) and previous studies [6–8]. Deletion of *gaaC* also induced the expression of several pectinases acting on RG-I that do not belong to GaaR/GaaX panregulon (Table 2). A possible explanation is that starvation in *ΔgaaC* results in the induction of these genes. Several pectinases acting on side chains of RG-I, including NRRL3_05407, *lac35B* and NRRL3_07501, were previously reported to be induced upon starvation [43]. Another explanation is that the increased transcript levels of *rhaR* and *xlnR* results in an increase in the expression of these genes that were suggested to be under control of RhaR and XlnR (see above).

Although both *ΔgaaB* and *ΔgaaC* cannot utilize GA, residual growth of *ΔgaaC* was observed on AP, whereas the growth of *ΔgaaB* on AP was more impaired (Fig. 1B). This could be explained by the high capacity of *ΔgaaC* to secrete pectinases acting on RG-I and release monosaccharides (L-arabinose, L-rhamnose, D-galactose) other than GA to support growth, and the less efficient pectinase production in *ΔgaaB*.

Previously, we identified 53 genes as the GaaR/GaaX panregulon downregulated in *ΔgaaR* under inducing condition and/or upregulated in *ΔgaaX* under noninducing condition. However, only a core set of 27 genes was significantly differentially regulated under both conditions [19,20], and only 17 of 53 panregulon genes, 10 of which belong to the core regulon, were hyperinduced in response to deletion of *gaaC* (Table 1), demonstrating the complex regulation of GA-responsive gene expression. A dynamic equilibrium is suggested to exist between the free and DNA-bound states of a transcription factor, and the binding of a transcription factor to the promoters of its target genes depends on its concentration, as well as its cooperative/competitive interactions with other proteins and the chromatin accessibility [44,45]. Deletion of *gaaR* would result in the lack of GaaR in the cell, whereas deletion of *gaaX* or intracellular accumulation of 2-keto-3-deoxy-L-galactonate in *ΔgaaC* would, possibly to different degrees, increase the concentration of active GaaR by elimination or reducing the repressing activity of GaaX. GaaR concentration might also be

regulated transcriptionally: *gaaX* is highly upregulated in GA [5], whereas *gaaR* expression is significantly increased in the *ΔgaaC* mutant (FC = 5.10 and P-value = 7.88E-03). Moreover, different levels of CreA mediated repression on different GA-responsive genes [18] and accessibility of the promoter regions of these genes under different conditions might play a role in the observed differences in gene regulation. Condition specific cross-regulation between transcription factors and coregulation of target genes might add additional complexity to GA-responsive gene expression, as discussed above.

To conclude, in this study we identified the GA catabolic pathway intermediate 2-keto-3-deoxy-L-galactonate as the probable inducer of the GA-responsive genes in *A. niger*. Considering that both the GA catabolic pathway enzymes and the GaaR/GaaX activator–repressor module is evolutionarily conserved in the Pezizomycotina subdivision of Ascomycetes [5,20], it is highly probable that the mechanism by which 2-keto-3-deoxy-L-galactonate acts as an inducer and interacts with the activator-repressor module is also conserved.

NOTE AFTER PUBLICATION

In RNA-seq experiments, two biological replicates representing each condition were used, and differential gene expression was identified by Student's t-test (*P*-value≤0.05). This approach allows addressing only major differences in gene expression.

ACKNOWLEDGEMENTS

EA was supported by a grant from BE-Basic (Flagship 10). CK and MVAP were supported by a grant of the Dutch Technology Foundation STW, Applied Science Division of NWO, and the Technology Program of the Ministry of Economic Affairs 016.130.609 to RPdV. This works was in part supported by the NSERC Industrial Biocatalysis Network. We thank Peter Richard (VTT Technical Research Centre of Finland) for providing *A. niger* strains ATCC1015 *ΔgaaA*, *ΔgaaB*, and *ΔgaaC*.

AUTHOR CONTRIBUTIONS

EA, CK, TGH, SdP, MA, MDF, TTMP performed experiments. EA, MDF, MP, MVAP performed bioinformatics analysis. EA, JV, AT, RPdV, and AFJR wrote the manuscript with input of all authors.

SUPPORTING INFORMATION

Additional Supporting Information may be found online (http://onlinelibrary.wiley.com/doi/10.1002/1873-3468.12654/abstract) in the supporting information tab for this article.

REFERENCES

1 Mohnen D (2008) Pectin structure and biosynthesis. *Curr Opin Plant Biol* 11, 266–277.

2 Martens-Uzunova ES and Schaap PJ (2009) Assessment of the pectin degrading enzyme network of *Aspergillus niger* by functional genomics. *Fungal Genet Biol* 46, S170–S179.

3 Coutinho PM, Andersen MR, Kolenova K, Vankuyk PA, Benoit I, Gruben BS, Trejo-Aguilar B, Visser H, van Solingen P, Pakula T et al. (2009) Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. *Fungal Genet Biol* 46, S161–S169.

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

5 Martens-Uzunova ES and Schaap PJ (2008) An evolutionary conserved D-galacturonic acid metabolic pathway operates across filamentous fungi capable of pectin degradation. *Fungal Genet Biol* 45, 1449–1457.

6 Mojzita D, Wiebe M, Hilditch S, Boer H, Penttila M and Richard P (2010) Metabolic engineering of fungal strains for conversion of D-galacturonate to *meso*-galactarate. *Appl Environ Microbiol* 76, 169–175.

7 Kuivanen J, Mojzita D, Wang Y, Hilditch S, Penttila M, Richard P and Wiebe MG (2012) Engineering filamentous fungi for conversion of D-galacturonic acid to L-galactonic acid. *Appl Environ Microbiol* 78, 8676– 8683.

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

9 Zhang L, Thiewes H and van Kan JA (2011) The D-galacturonic acid catabolic pathway in *Botrytis cinerea*. *Fungal Genet Biol* 48, 990–997.

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

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

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

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

14 Kuivanen J, Sugai-Guerios MH, Arvas M and Richard P (2016) A novel pathway for fungal D-glucuronate catabolism contains an L-idonate forming 2-keto-L-gulonate reductase. *Sci Rep* 6, 26329.

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

16 Huberman LB, Liu J, Qin L and Glass NL (2016) Regulation of the lignocellulolytic response in filamentous fungi. *Fungal Biol Rev* 30, 101–111.

17 de Vries RP, Jansen J, Aguilar G, Parenicova L, Joosten V, Wulfert F, Benen JA and Visser J (2002) Expression profiling of pectinolytic genes from *Aspergillus niger*. *FEBS Lett* 530, 41–47.

18 Niu J, Homan TG, Arentshorst M, de Vries RP, Visser J and Ram AF (2015) The interaction of induction and repression mechanisms in the regulation of galacturonic acid-induced genes in Aspergillus niger. Fungal Genet Biol 82, 32–42.

19 Alazi E, Niu J, Kowalczyk JE, Peng M, Aguilar Pontes MV, van Kan JAL, Visser J, de Vries RP and Ram AFJ (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.

20 Niu J, Alazi E, Reid ID, Arentshorst M, Punt PJ, Visser J, Tsang A and Ram AFJ (2017) An evolutionarily conserved transcriptional activator-repressor module controls expression of genes for D-galacturonic acid utilization in *Aspergillus niger*. *Genetics* 205, 169–183.

21 Zhang L, Lubbers RJM, Simon A, Stassen JHM, Ribera PRV, Viaud M and van Kan JAL (2015) A novel Zn_2Cys_6 transcription factor BcGaaR regulates D-galacturonic acid utilization in *Botrytis cinerea*. *Mol Microbiol* 100, 247–262.

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

23 Keegan L, Gill G and Ptashne M (1986) Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. *Science* 231, 699–704.

24 Johnston M (1987) Genetic evidence that zinc is an essential co-factor in the DNA binding domain of GAL4 protein. *Nature* 328, 353–355.

25 Ma J and Ptashne M (1987) A new class of yeast transcriptional activators. *Cell* 51, 113–119.

26 Sil AK, Alam S, Xin P, Ma L, Morgan M, Lebo CM, Woods MP and Hopper JE (1999) The Gal3p-Gal80p-Gal4p transcription switch of yeast: Gal3p destabilizes the Gal80p-Gal4p complex in response to galactose and ATP. *Mol Cell Biol* 19, 7828–7840.

27 Lavy T, Kumar PR, He H and Joshua-Tor L (2012) The Gal3p transducer of the GAL regulon interacts with the Gal80p repressor in its ligand induced closed conformation. *Genes Dev* 26, 294–303.

28 Poulou M, Bell D, Bozonelos K, Alexiou M, Gavalas A, Lovell-Badge R and Remboutsika E (2010) Development of a chromosomally integrated metabolite-inducible Leu3p-α-IPM “off-on” gene switch. *PLoS One* 5, e12488.

29 Sze JY, Woontner M, Jaehning JA and Kohlhaw GB (1992) In vitro transcriptional activation by a metabolic intermediate: activation by Leu3 depends on a-isopropylmalate. *Science* 258, 1143–1145.

30 Wang D, Zheng F, Holmberg S and Kohlhaw GB (1999) Yeast transcriptional regulator Leu3p. Self-masking, specificity of masking, and evidence for regulation by the intracellular level of Leu3p. *J Biol Chem* 274, 19017–19024.

31 Wightman R, Bell R and Reece RJ (2008) Localization and interaction of the proteins constituting the GAL genetic switch in *Saccharomyces cerevisiae*. *Eukaryot Cell* 7, 2061–2068.

32 Kirkpatrick CR and Schimmel P (1995) Detection of leucine-independent DNA site occupancy of the yeast Leu3p transcriptional activator in vivo. *Mol Cell Biol* 15, 4021–4030.

33 Makita T, Katsuyama Y, Tani S, Suzuki H, Kato N, Todd RB, Hynes MJ, Tsukagoshi N, Kato M and Kobayashi T (2009) Inducer-dependent nuclear localization of a $Zn(II)_2Cys_6$ transcriptional activator, AmyR, in *Aspergillus nidulans*. *Biosci Biotechnol Biochem* 73, 391–399.

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

35 Suzuki K, Tanaka M, Konno Y, Ichikawa T, Ichinose S, Hasegawa-Shiro S, Shintani T and Gom K (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.

36 Arentshorst M, Ram AFJ and Meyer V (2012) Using non-homologous end-joining-deficient strains for functional gene analyses in filamentous fungi. In Plant Fungal Pathogens: Methods and Protocols, Methods in Molecular Biology (Bolton MD and Thomma BPHJ, eds), pp. 133–150. Vol. 835. Springer Science + Business media, New York, NY.

37 Carvalho NDSP, Arentshorst M, Kwon MJ, Meyer V and Ram AFJ (2010) Expanding the *ku70* toolbox for filamentous fungi: establishment of complementation vectors and recipient strains for advanced gene analyses. *Appl Microbiol Biotechnol* 87, 1463–1473.

38 Edgar R, Domrachev M and Lash AE (2002) Gene expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30, 207–210.

39 Gruben BG (2012) Novel transcriptional activators of *Aspergillus* involved in plant biomass utilization. PhD Thesis, Utrecht University.

40 Gruben BS, Zhou M, Wiebenga A, Ballering J, Overkamp KM, Punt PJ and de Vries RP (2014) *Aspergillus niger* RhaR, a regulator involved in L-rhamnose release and catabolism. *Appl Microbiol Biotechnol* 98, 5531–5540.

41 van Peij NNME, Visser J and de Graaff LH (1998) Isolation and analysis of *xlnR*, encoding a transcriptional activator co-ordinating xylanolytic expression in *Aspergillus niger*. *Mol Microbiol* 27, 131–142.

42 Battaglia E, Visser L, Nijssen A, van Veluw GJ, Wosten HAB and de Vries R (2011) Analysis of regulation of pentose utilisation in *Aspergillus niger* reveals evolutionary adaptations in *Eurotiales*. *Stud Mycol* 69, 31–38.

43 van Munster JM, Daly P, Delmas S, Pullan ST, Blythe MJ, Malla S, Kokolski M, Noltorp EC, Wennberg K, Fetherston R et al. (2014) The role of carbon starvation in the induction of enzymes that degrade plant-derived carbohydrates in *Aspergillus niger*. *Fungal Genet Biol* 72, 34–47.

44 Biggin MD (2011) Animal transcription networks as highly connected, quantitative continua. *Dev Cell* 21, 611–626.

45 Granek JA and Clarke ND (2005) Explicit equilibrium modeling of transcription-factor binding and gene regulation. *Genome Biol* 6, Article R87.

CHAPTER 6

Inducer-independent production of pectinases in *Aspergillus niger* by overexpression of the D-galacturonic acid responsive transcription factor *gaaR*

Alazi E, Knetsch T, Di Falco M, Reid ID, Arentshorst M, Visser J, Tsang A, Ram AFJ

Published in *Applied Microbiology and Biotechnology* (2018) 102:2723-2736

ABSTRACT

The transcription factor GaaR is needed for the expression of genes required for pectin degradation and transport and catabolism of the main degradation product, D-galacturonic acid (GA) in *Aspergillus niger*. In this study, we used the strong constitutive *gpdA* promoter of *Aspergillus nidulans* to overexpress *gaaR* in *A. niger*. Overexpression of *gaaR* resulted in an increased transcription of the genes encoding pectinases, (putative) GA transporters, and catabolic pathway enzymes even under non-inducing conditions, i.e., in the absence of GA. Exoproteome analysis of a strain overexpressing *gaaR* showed that this strain secretes highly elevated levels of pectinases when grown in fructose. The genes encoding exo-polygalacturonases were found to be subjected to CreA-mediated carbon catabolite repression, even in the presence of fructose. Deletion of *creA* in the strain overexpressing *gaaR* resulted in a further increase in pectinase production in fructose. We showed that GaaR localizes mainly in the nucleus regardless of the presence of an inducer, and that overexpression of *gaaR* leads to an increased concentration of GaaR in the nucleus.

INTRODUCTION

Aspergillus niger is an important filamentous fungus for the industrial production of pectinases (Pedrolli et al. 2009). Pectinases are widely used in the food industry (Kashyap et al. 2001; Toushik et al. 2017; Khan et al. 2013) and are important enzymes in the utilization of pectin-rich feedstock in biofuel production (Edwards and Doran-Peterson 2012). Pectin is a complex plant cell wall polysaccharide and four substructures have been de defined which include polygalacturonic acid (PGA), rhamnogalacturonan I, rhamnogalacturonan II, and xylogalacturonan. PGA is the most abundant pectic substructure and consists of D-galacturonic acid (GA) residues. GA is also present in the backbones of rhamnogalacturonan I, rhamnogalacturonan II, and xylogalacturonan (Caffall and Mohnen 2009).

A. niger contains a large number of enzymes potentially acting on pectin substructures (Martens-Uzunova and Schaap 2009; Coutinho et al. 2009; de Vries et al. 2017). In the presence of GA, the main sugar acid in pectin, the expression of the genes encoding pectinases (Martens-Uzunova and Schaap 2009; Alazi et al. 2016), the GA transporter GatA (Sloothaak et al. 2014), and the GA catabolic pathway enzymes GaaA, GaaB, GaaC, and GaaD (Martens-Uzunova and Schaap 2008; Alazi et al. 2016) are induced via the Zn_2Cys_6 type transcription factor GaaR (Alazi et al. 2016). Apart from the transcriptional activator (GaaR), the expression of GA-responsive genes is controlled by a repressor protein, GaaX. Loss of function of GaaX leads to constitutive and inducer-independent expression of pectinases (Niu et al. 2017). The repressor protein GaaX is postulated to inhibit the transcriptional activity of GaaR under non-inducing conditions, i.e., in the absence of GA. The presence of an inducer is suggested to inhibit the repressing activity of GaaX, thereby leading to the transcriptional induction of GA-responsive genes via GaaR (Niu et al. 2017). The GA catabolic pathway intermediate 2-keto-3-deoxy-L-galactonate has recently been identified as the physiological inducer of the GA-responsive genes (Alazi et al. 2017).

Overexpression of transcription factors has been shown to be an effective method to increase the expression of their target genes in *Saccharomyces cerevisiae*, even under conditions in which the transcription factors under consideration are normally not active (Chua et al. 2006). Similarly, overexpression of transcription factors involved in plant biomass degradation in filamentous fungi, such as *xlnR* (Noguchi et al. 2009) and *manR* (Ogawa et al. 2012) in *Aspergillus oryzae* and *xyr1* in *Trichoderma reesei* (Jiang et al. 2016), was previously reported to result in elevated expression of their target genes in the presence of inducers. Inducer-independent production of cellulases was also observed in *T. reesei* strains overexpressing *xyr1* (Lv et al. 2015; Wang et al. 2013).

In this study, we demonstrate that overexpression of *gaaR* results in constitutive transcription and secretion of pectinases under non-inducing conditions, probably by disturbing the stoichiometric balance of GaaR and GaaX in favor of GaaR. We further show that the effect of *gaaR* overexpression on pectinase production is sensitive to CreA-mediated carbon catabolite repression even when fructose, a less repressing carbon source compared to glucose, was used. A further increase in pectinase production on fructose upon *gaaR* overexpression was accomplished when the CreA-mediated carbon catabolite repression was inactivated via *creA* deletion.

MATERIALS AND METHODS

Strains, media, and growth conditions

All *A. niger* strains used in this study are listed in Online Resource 1. Media were prepared as described previously (Arentshorst et al. 2012). Radial growth assays of the strains were performed on minimal medium (MM) (pH 5.8) containing 1.5% (w/v) agar (Scharlau, Barcelona, Spain) and various carbon sources: 50 mM glucose (VWR International, Amsterdam, The Netherlands), fructose (Sigma-Aldrich, Zwijndrecht, The Netherlands), or GA (Chemodex, St Gallen, Switzerland), or 1% (w/v) PGA (Sigma, Zwijndrecht, The Netherlands) or apple pectin (AP) (Sigma-Aldrich, Zwijndrecht, The Netherlands). Plates were inoculated with 5 µL 0.9% NaCl containing 10^4 freshly harvested spores and cultivated at 30 °C for 7 days. MM (pH 5.8) containing 1.5% (w/v) agar, 10 mM acetamide (Sigma-Aldrich, Steinheim, Germany) as the sole nitrogen source, and acetate (Merck, Darmstadt, Germany), glucose, fructose, sorbitol (Roth, Karlsruhe, Germany) or GA as the carbon source was prepared as described previously (Arentshorst et al. 2012). Plates were inoculated with 5 µL 0.9% NaCl containing $5x10^4$ freshly harvested spores. Filter sterilized carbon source solutions were added after autoclaving MM containing agar. PGA and AP were autoclaved together with the medium. All growth experiments were performed in duplicate.

For enzymatic analysis, 10^6 freshly harvested spores were inoculated per mL in 100 mL shake flasks that include 25 or 50 mL MM (pH 5.8) containing 50 mM glucose, fructose, sorbitol, or GA and were grown for 36 h in a rotary shaker at 30 °C and 250 rpm. Experiments were performed in duplicate.

For microscopic analysis of the co-localization of the nuclear specific SYTO59 dye (Invitrogen, Eugene, Oregon, USA) with the eGFP-tagged H2B protein, conidia of the MA26.1 strain were propagated on complete medium containing 1.5% (w/v) agar. $2x10^5$ freshly harvested spores were placed on cover slips in a Petri dish with 20 mL MM containing 50 mM fructose and grown at 30 °C. After 16 h, the cover slips were rinsed twice with water and transferred to a new Petri dish with 20 ml MM containing 50 mM GA and growth was continued at 30 °C for 1.5 h. For microscopic analysis of the co-localization of the nuclear specific SYTO59 dye with the eGFP-tagged GaaX or GaaR proteins, conidia of the JN126.2, EA19.2, and EA20.10 strains were propagated on MM containing 1.5% (w/v) agar and 50 mM GA. $3x10^5$ freshly harvested spores were inoculated on cover slips in Petri dishes that include 3 mL MM containing 10 mM GA and 0.003% yeast extract and grown at 30 °C for approximately 22 h. For microscopic analysis of the fluorescence intensity, conidia of the EA19.2 and EA20.10 strains were propagated on complete medium containing 1.5% (w/v) agar. $2x10^5$ freshly harvested spores were placed on cover slips in Petri dishes that include 20 mL MM containing 50 mM fructose, or 50 mM GA and 1 mM fructose, and grown at 30 °C for 17.5 h. For each condition, two biological replicates were performed.

Construction of strains overexpressing *gaaR*

Protoplast-mediated transformation of *A. niger*, purification of the transformants and extraction of the genomic DNA were performed as described by Arentshorst et al. (2012).

The plasmid pEA4 containing the P*gpdA-gaaR*-T*trpC* construct was created as follows: the *Aspergillus nidulans gpdA* promoter was obtained from plasmid pAN52.1-NOTI (Punt et al. 1987) by restriction digestion with *Not*I and *Nco*I. The *gaaR* gene was amplified by PCR using the primer pairs listed in Online Resource 2 with *A. niger* N402 genomic DNA as template, ligated into pJET1.2/blunt cloning vector (Thermo Fisher Scientific, Carlsbad, CA) and amplified in *Escherichia coli* DH5*α*. Following plasmid isolation, *gaaR* was excised using restriction enzymes *Psc*I and *Bgl*II. The *Not*I-*Nco*I fragment of P*gpdA* and the *Psc*I-*Bgl*II fragment of *gaaR* were ligated into *Not*I-*Bam*HI opened pAN52.1-NOTI. pEA4 was sequenced to ensure no PCR errors have occurred and proper ligation and orientation of the fragments. To create strains EA21.3, EA21.5, EA21.6, and EA21.8, pEA4 was cotransformed into strain JN36.1 together with the plasmid pMA357 containing the *A. nidulans amdS* gene behind the *A. nidulans gpdA* promoter (Alazi et al. 2016). Transformants were selected on plates containing acetamide as the sole nitrogen source. To create strain TK1.1, strain JN36.1 was co-transformed with pEA4 and the plasmid p3SR2 (Hynes et al. 1983). p3SR2 contains the *A. nidulans amdS* gene behind the endogenous *amdS* promoter (Hynes et al. 1983). Transformants were selected on plates containing acrylamide as the sole nitrogen source. Ectopic integration of the P*gpdA-gaaR*-T*trpC* construct was confirmed via Southern blot analysis. Genomic DNA was restricted overnight with *Nco*I restriction enzyme. A 501 bp fragment containing the *gaaR* gene was PCR-amplified using the primer pairs listed in Online Resource 2 with N402 genomic DNA as template, and was used as a probe.

Construction of strains (over)expressing *eGFP-gaaR*

The *gaaR* and *eGFP* genes were amplified by PCR using the primer pairs listed in Online Resource 2 with N402 genomic DNA and the plasmid pFG029 (unpublished vector, containing P*gpdA-eGFP*-T*trpC*) as template, respectively. *eGFP* and *gaaR* were combined by fusion PCR using primers eGFP_P1_NcoI and gaaR_comp_P2_BglII, and the *eGFP-gaaR* fusion product was ligated into pJET1.2/blunt cloning vector and amplified in *E. coli* DH5*α*. Following plasmid isolation, the *eGFP-gaaR* fusion product was excised in two parts using restriction enzymes *Nco*I and *Bgl*II, resulting in an *Nco*I-*Nco*I fragment and an *Nco*I-*Bgl*II fragment.

The plasmid pEA3 containing the P*gpdA-eGFP-gaaR*-T*trpC* construct was created as follows: The *Not*I-*Nco*I fragment of P*gpdA* and the *Nco*I-*Bgl*II fragment of *eGFP-gaaR* were ligated into *Not*I-*Bam*HI opened pAN52.1-NOTI. The resulting plasmid was digested with *Nco*I and ligated with the *Nco*I-*Nco*I fragment of *eGFP-gaaR*. pEA3 was sequenced to ensure no PCR errors and proper ligation and orientation of the fragments. Strain EA20.10 was created by cotransformation of strain JN36.1 with pEA3 together with the plasmid pMA357.

To construct plasmid pEA2 (P*gaaR-eGFP-gaaR*-T*trpC*), the *gaaR* promoter was PCR-amplified using the primer pairs listed in Online Resource 2 with N402 genomic DNA as template, ligated into pJET1.2/blunt cloning vector and amplified in *E. coli* DH5*α*. Following plasmid isolation, P*gaaR* was excised using restriction enzymes *Not*I and *Nco*I. pEA2 was created in a similar way to pEA3, except that the *Not*I-*Nco*I fragment of P*gaaR*

was used instead of P*gpdA*. pEA2 was sequenced to ensure no PCR errors and proper ligation and orientation of the fragments. Strain EA19.2 was created by cotransformation of strain JN36.1 with pEA2 together with the plasmid pMA357 and transformants were selected on plates containing acetamide as the sole nitrogen source. Ectopic integrations of the P*gpdA-eGFP-gaaR*-T*trpC* and P*gaaR-eGFP-gaaR*-T*trpC* constructs were confirmed by diagnostic PCRs (data not shown).

Construction of *creA* deletion strains

Loss of the *pyrE* gene in EA21.6 was mediated by counter selection on MM-5'-FOA plates (Arentshorst et al. 2012), resulting in the strain EA23.6. The split marker approach was employed in the deletion of the *creA* gene (Arentshorst et al. 2015). 5' and 3' flanks of creA were PCR-amplified using the primer pairs listed in Online Resource 2 with N402 genomic DNA as template. The *A. nidulans pyrF* gene (named *pyrE* in *A. niger*) was PCR-amplified as two fragments using the primer pairs listed in Online Resource 2 with *A. nidulans* strain A234 genomic DNA as template. Split marker fragments with the *pyrF* selection marker were created by fusion PCR and used to transform the strain EA23.6, resulting in the strain TK2.1. Proper deletion of *creA* was confirmed by diagnostic PCR (data not shown).

MA342.2 was also constructed using the split marker approach (Arentshorst et al. 2015). 5' and 3' flanks of *creA* were PCR-amplified using the primer pairs listed in Online Resource 2 with N402 genomic DNA as template. The hygromycin resistance cassette was PCR-amplified using primers hygP3f and hygP4r and a derivative of pAN7.1 (Punt et al. 1987) as template. *creA-hygR* split marker fragments were created by fusion PCR and transformed to strainMA234.1, resulting in the *ΔcreA* strain MA342.2. Proper deletion of *creA* was confirmed by diagnostic PCR (data not shown).

Bioreactor cultivations and transcriptome and exoproteome analyses

Controlled bioreactor cultivations of MA234.1 (the reference strain) (in triplicate) and JN123.1 (*ΔgaaX*) (in duplicate) in MM containing 0.75% fructose and the subsequent transcriptome analyses were performed previously (Niu et al. 2017). Controlled bioreactor cultivations of the EA21.6strain (*OEgaaR*) (in duplicate) under exactly the same growth conditions and the subsequent RNA-seq analyses were performed as previously described by Niu et al. (2017). Both biomass accumulation (offline) and base addition (online) were determined to monitor exponential growth.

Broth samples were taken during exponential growth after every 4 mL of base addition. RNA isolated from exponentially growing cells at the sample point at which about 75–80% of the maximum biomass yield was reached was used for the RNA-seq experiment. RNA-seq data were submitted to the Sequence Read Archive under accession number SRP078485 for MA234.1 and JN123.1 (Niu et al. 2017) and accession number SRP114830 for EA21.6 (this study).

Supernatant samples from an exponentially growing culture of each strain at two successive sample points (based on base addition) following the RNA-seq sample point

were withdrawn and filtered. The filtered supernatants were lyophilized, resuspended in 1 mL 50 mM citric acid buffer pH 5.0, and used for the exoproteome analysis. For each sample, proteins were precipitated with TCA (trichloroacetic acid), the pellet was washed twice with acetone and resuspended in 75 µL 200 mM ammonium bicarbonate and 0.1% AALS II. Protein concentrations were determined by the RCDC assay Kit (BioRad, Mississauga, Ontario). Three micrograms of total protein were loaded on 12% SDS-PAGE gel. The gel was colored with silver stain and developed for 3 min. Five micrograms of total protein were trypsin digested in solution overnight at 37 °C. Samples were desalted with C18 ziptips (Millipore, Billerica, MA), the eluate was dried and the peptides were resuspended in 50 µL 5% acetonitrile and 0.1% formic acid. Five microlitres of peptide digest were analyzed by LC-MS/MS on a Velos LTQ-Orbitrap. Extracted Ion Chromatogram (EIC) peak area values of proteins were calculated by averaging the top three most abundant peptide ion EIC values assigned to each protein as per the Proteome Discoverer 1.4 (Thermo Fisher, San Jose, CA) precursor peak area quantification workflow. Protein EIC area values were normalized using the determined value of a fixed spiked amount of trypsin-digested bovine serum albumin.

Enzymatic analysis

Supernatants from bioreactor or shake flask cultures were obtained by filtration through glass microfiber filters (Whatman, Buckinghamshire, UK) or sterile miracloth, and the filtrate was stored at − 80 °C. PGA plate assays were performed as described by Niu et al. (2017). Twenty-five microlitres of supernatant from each culture was spotted on plates containing 0.2% or 0.5% PGA, and plates were incubated at 37 °C for 16 or 20 h. PGA degradation was indicated by the formation of a clear zone of hydrolysis.

Microscopy

The cover slips with adherent germlings were placed upside down on glass slides and observed under a Zeiss Observer confocal laser scanning microscope (Zeiss, Jena, Germany). For nuclear staining, 0.5 mL of 25 µM SYTO59 dye solution was dropped on glass slides before placing the cover slips, and imaging was performed approximately after 2 h. The GFP and SYTO59 fluorescence were excited using 488 and 625 nm laser lines, respectively. Images were analyzed using the ImageJ software (Abramoff et al. 2004). To analyze the fluorescence intensity, 1–2 images were taken for each biological replicate. On each image, the exact same brightness and contrast adjustments were applied, 3–10 nuclear and 3–10 cytoplasmic fluorescence intensities in a defined area were measured, and calibrated for the background fluorescence.

RESULTS

Expression of pectinase genes is upregulated in strains overexpressing *gaaR*

To create strains that overexpress the GA-responsive transcription factor gene *gaaR* (*OEgaaR*), the *A. niger gaaR* gene was fused with the strong constitutive *A. nidulans gpdA*

promoter and transformed into a *ΔgaaR* strain. Southern blot analysis indicated that the *gaaR* overexpression construct was ectopically integrated in one or two copies in the genomes of the *OEgaaR* strains EA21.3, EA21.5, EA21.6, and EA21.8 (Fig. ESM_3.1 b). An additional multicopy *OEgaaR* strain, TK1.1, was obtained by using a more stringent selection method using acrylamide, and Southern blot analysis confirmed ectopic integration of at least four copies of the *gaaR* overexpression construct in its genome (Fig. ESM_3.1 c).We compared the radial growth of the *OEgaaR* strains on different monomeric and polymeric carbon sources (Fig. ESM_3.2). The *ΔgaaR* strain showed a strongly reduced growth on GA, PGA, and AP as previously shown (Alazi et al. 2016). Reintroducing the *gaaR* gene expressed from the *gpdA* promoter resulted in growth on GA, PGA, and AP in EA21.3, EA21.5, EA21.6, and EA21.8, indicating the presence of a functional GaaR. However, the *OEgaaR* strains showed partial and different levels of complementation of growth on GA containing carbon sources (Fig. ESM_3.2 a). The TK1.1 strain, containing the highest copy number of the *gaaR* overexpression construct, showed a severely impaired growth on GA, PGA, and AP (data not shown and Fig. ESM_3.2 b). Growth of all *OEgaaR* strains on glucose or fructose was similar to the growth of the reference strain.

To assess the pectinase production capacity of the *OEgaaR* strains, the strains were grown in shake flasks in minimal medium containing non-inducing (50 mM glucose, 50 Mm fructose, or 50 mM sorbitol) or inducing (50 mM GA) carbon sources, and the culture supernatants were spotted on PGA plates. As indicated by the clear zones of hydrolysis on PGA plates, the polygalacturonase activity in the culture supernatants of the *OEgaaR* strains EA21.3, EA21.5, EA21.6, and EA21.8 grown in glucose or fructose was higher compared to the reference strain (Fig. ESM_3.3 a). The culture supernatant of EA21.6 grown on sorbitol or GA displayed the highest polygalacturonase activity compared to EA21.3, EA21.5, EA21.8, and the reference strain (Fig. ESM_3.3 a and data not shown). EA21.6 was selected to be used in further experiments based on the growth profiles and pectinase production capacities of the *OEgaaR* strains. A further increased level of polygalacturonase activity was observed in the culture supernatant of TK1.1 grown on fructose compared to the reference strain as well as to EA21.6 (Fig. ESM_3.3 b).

Transcriptome analysis of the *OEgaaR* strain

To investigate the expression of the multitude of genes involved in pectin degradation, GA transport, and catabolism in a strain overexpressing *gaaR*, we performed a genome wide gene expression analysis using RNA-seq (Online Resource 4). The reference and *OEgaaR* (EA21.6) strains were grown in bioreactors on fructose, a carbon source that does not induce the expression of GA-responsive genes (Martens-Uzunova and Schaap 2008). Growth of the *OEgaaR* strain under the controlled bioreactor conditions (μ_{max} 0.200 ± 0.001 g dry weight/kg/h, Y_{max} 4.117 ± 0.167 g dry weight/kg (n = 2)) was similar to the growth of the reference strain (μ_{max} 0.214 ± 0.007 g dry weight/kg/h, Y_{max} 4.151 ± 0.134 g dry weight/kg (n = 3)). Analysis of RNA-seq data showed that the expression of *gaaR* was highly increased in the *OEgaaR* strain compared to the reference strain with a fold change of 63.8 (Online Resource 4 and Table 1). Overexpression of *gaaR* resulted in the upregulation (FC ≥ 4, FDR ≤ 0.001) of 19 of 48 genes encoding extracellular enzymes specifically assigned to the degradation of pectin according to de Vries et al. (2017) (see

Table 1). Almost all of these genes (18 out of 19) belong to the GaaR/GaaX panregulon (Niu et al. 2017) and include several exo- and endo-polygalacturonases, pectin methylesterases, and pectin lyases, all acting on the PGA backbone of pectin, as well as the xylogalacturonase NRRL3_07469 acting on xylogalacturonan, and the arabinogalactanase encoded by *gan53A* and the α-L-rhamnosidase NRRL3_10558 acting on rhamnogalacturonan I. Nine of 19 pectinases that were upregulated in the *OEgaaR* strain were previously shown to be upregulated in *ΔgaaX* (JN123.1), the repressor deletion mutant, on fructose compared to the reference strain (FC ≥ 4, FDR ≤ 0.001) (Niu et al. 2017). The expression of these nine pectinase genes, except NRRL3_05252, was generally much lower in *ΔgaaX* than in the *OEgaaR* strain, and no additional pectinases were found to be upregulated in *ΔgaaX* (Table 1). Constitutive production of pectinases in the *OEgaaR* strain and to a lesser extent in *ΔgaaX* grown in bioreactors on fructose was also observed via a PGA plate assay (Fig. 1a). This analysis clearly indicates that overexpression of *gaaR* results in a more dramatic increase in the expression of several pectinases compared to deletion of the repressor *gaaX*.

Comparison of the expression of the genes encoding the (putative) GA transporters and the GA catabolic pathway enzymes (*gaaA*, *gaaB*, *gaaC*, and *gaaD*) between the *OEgaaR*, the reference and *ΔgaaX* strains revealed that *gatA*, the putative GA transporter NRRL3_04281, *gaaA*, *gaaB*, and *gaaC* were significantly upregulated (FC ≥ 4, FDR ≤ 0.001) in the *OEgaaR* strain (Table 1). The expression of *gaaD* was also significantly (FDR ≤ 0.001) increased in the *OEgaaR* strain compared to the reference strain with a fold change of 2.5 (Table 1). Interestingly, the expression of the genes encoding the GA catabolic pathway enzymes were moderately induced in the *OEgaaR* strain and expressed at much higher levels in the *ΔgaaX* strain. In contrast, many of the genes encoding the extracellular enzymes were expressed at higher levels in the *OEgaaR* strain compared to the *ΔgaaX* strain (see “Discussion”).

We also analyzed the effect of overexpression of *gaaR* on the expression of all 375 genes predicted to encode carbohydrate active enzymes (CAZymes) in *A. niger* strain NRRL3 (Online Resource 4). In addition to the above mentioned 19 pectinases (belonging to CAZy families CE8, GH28, GH53, GH78, PL1_4, and PL4_3), 20 CAZymes acting specifically on cellulose (AA9, GH5_5), starch (GH13_5, GH31), and xyloglucan (GH12, GH74, GH95) or acting on multiple substrates (CE16, GH18, GH3, GH35, GH43, GH51, GH54 GH79) (de Vries et al. 2017) were highly upregulated in the *OEgaaR* strain on fructose (Table 2).

In total, 124 genes were significantly upregulated (FC ≥ 4, FDR ≤ 0.01) in the *OEgaaR* strain compared to the reference strain (Online Resource 4). The promoter regions of 110 upregulated genes for which the *A. niger* CBS 513.88 gene ID was available were screened for the presence of transcription factor binding sites using TFBSF (Meyer et al. 2009), and it was found that 69 genes contain the galacturonic acid-responsive element GARE (CCNCCAA) (Martens-Uzunova and Schaap 2008) required for GA-responsive gene induction (Niu et al. 2015) in their 1 kb upstream sequences. A gene ontology enrichment analysis via FetGOat (Nitsche et al. 2011) indicated that the genes upregulated in the *OEgaaR* strain were highly enriched with genes involved in carbohydrate (xyloglucan, pectin, lactose) metabolism. Out of 53 genes belonging to the GaaR/GaaX panregulon (Niu et al. 2017), 34 were upregulated in the *OEgaaR* strain, including *gaaX* with a fold change

Table 1. Transcriptome and exoproteome analyses of predictive pectinases, GA transporters and catabolic pathway genes, *gaaR* and *gaaX* in the reference, *OEgaaR* and *ΔgaaX* strains. Transcript and extracellular protein levels are represented as TPM and normalized protein (EIC) area values, respectively. Transcript levels with a fold change≥4 and FDR≤0.001 compared to the reference strain are indicated with an asterisk. Genes belonging to the GaaR/GaaX panregulon (Niu et al. 2017) are written in bold. Transcriptome data for the reference and *ΔgaaX* strains were taken from Niu et al. 2017.

Gene ID NRRL3	Gene ID CBS 513.88	Description	CAZy family	Transcript level (ave TPM)			Extracellular protein level (ave normalized protein (EIC) area)		
				Reference	*ΔgaaX*	*OEgaaR*	Reference	*ΔgaaX*	*OEgaaR*
NRRL3_00169	An09g02160	rhamnogalacturonan acetylesterase	CE12	0.455	0.680	0.822	#N/A	#N/A	#N/A
NRRL3_07501	An04g09360	carbohydrate esterase family 12 protein	CE12	1.313	1.134	1.680	0.00	0.00	1.31
NRRL3_08325	**An03g06310**	**pectin methylesterase Pme8A**	CE8	0.037	0.556*	24.737*	0.00	1.48	105.91
NRRL3_07470	**An04g09690**	**pectin methylesterase**	CE8	0.755	5.722*	57.977*	0.55	7.49	45.73
NRRL3_05252	**An02g12505**	**pectin methylesterase-like protein**	CE8	1.061	31.320*	17.513*	0.00	109.39	48.98
NRRL3_00839	An14g02920	glycoside hydrolase family 105 protein	GH105	4.997	5.090	4.027	#N/A	#N/A	#N/A
NRRL3_01038	An14g05340	glycoside hydrolase family 105 protein	GH105	0.330	0.367	0.289	#N/A	#N/A	#N/A
NRRL3_06782	An16g06990	endo-polygalacturonase Pga28A	GH28	0.000	0.000	0.000	8.56	11.51	19.74
NRRL3_05859	An02g04900	endo-polygalacturonase Pga28B	GH28	16.541	18.675	27.407	0.00	0.22	1.76
NRRL3_08805	**An05g02440**	**endo-polygalacturonase Pga28C**	GH28	0.053	0.596	376.634*	0.00	0.00	321.96
NRRL3_00263	An09g03260	endo-polygalacturonase Pga28D	GH28	0.929	1.066	1.322	0.605	0	0
NRRL3_02835	**An01g14670**	**endo-polygalacturonase Pga28E**	GH28	0.483	1.152	479.334*	0.00	0.00	415.78
NRRL3_02571	**An01g11520**	**endo-polygalacturonase Pga28I**	GH28	0.215	1.243	23.092*	0.00	0.00	8.40
NRRL3_04000	An15g05370	endo-polygalacturonase Pga28II	GH28	0.293	1.055	2.076*	0.00	2.94	11.24
NRRL3_03144	**An12g07500**	**exo-polygalacturonase**	GH28	1.361	51.633*	65.839*	0.00	23.17	46.86
NRRL3_08281	**An03g06740**	**exo-polygalacturonase Pgx28B**	GH28	0.000	2.102*	14.290*	0.00	1.04	7.85
NRRL3_05260	**An02g12450**	**exo-polygalacturonase Pgx28C**	GH28	0.955	16.768*	45.050*	0.00	10.00	27.24
NRRL3_09810	**An11g04040**	**exo-polygalacturonase**	GH28	0.007	0.120	1.221*	#N/A	#N/A	#N/A
NRRL3_07469	An04g09700	xylogalacturonase	GH28	0.167	0.344	4.828*	0.00	0.00	10.86

Table 1. Transcriptome and exoproteome analyses of predictive pectinases, GA transporters and catabolic pathway genes, *gaaR* and *gaaX* in the reference, *OEgaaR* and *ΔgaaX* strains. Transcript and extracellular protein levels are represented as TPM and normalized protein (EIC) area values, respectively. Transcript levels with a fold change≥4 and FDR≤0.001 compared to the reference strain are indicated with an asterisk. Genes belonging to the GaaR/GaaX panregulon (Niu et al. 2017) are written in bold. Transcriptome data for the reference and *ΔgaaX* strains were taken from Niu et al. 2017. (Continued)

Gene ID NRRL3	Gene ID CBS 513.88	Description	CAZy family	Transcript level (ave TPM)			Extracellular protein level (ave normalized protein (EIC) area)		
				Reference	*ΔgaaX*	*OEgaaR*	Reference	*ΔgaaX*	*OEgaaR*
NRRL3_09126	An12g00950	endo-rhamnogalacturonase Rhg28A	GH28	5.967	6.717	4.859	#N/A	#N/A	#N/A
NRRL3_00953	An14g04200	endo-rhamnogalacturonase Rhg28B	GH28	0.000	0.009	0.000	0.69	2.31	0.00
NRRL3_04303	An07g01000	endo-rhamnogalacturonase	GH28	0.183	0.249	0.256	#N/A	#N/A	#N/A
NRRL3_09450	An11g08700	endo-rhamnogalacturonase	GH28	1.380	1.585	1.286	#N/A	#N/A	#N/A
NRRL3_11790	An06g02070	endo-rhamnogalacturonase	GH28	0.217	0.061	0.034	#N/A	#N/A	#N/A
NRRL3_09623	An11g06320	endo-rhamnogalacturonase	GH28	0.169	0.301	0.035	#N/A	#N/A	#N/A
NRRL3_02832	An01g14650	glycoside hydrolase family 28 protein	GH28	3.472	5.049	3.482	#N/A	#N/A	#N/A
NRRL3_08631	An03g02080	alpha-L-rhamnosidase Rha28B	GH28	0.000	0.021	0.159	#N/A	#N/A	#N/A
NRRL3_10559	**An18g04810**	**glycoside hydrolase family 28 protein**	GH28	0.084	3.112*	16.880*	0.00	2.34	7.39
NRRL3_10643	**An18g05940**	**arabinogalactanase Gan53A**	GH53	1.738	3.850	28.517*	0.00	2.91	22.91
NRRL3_02162	An01g06620	alpha-L-rhamnosidase	GH78	9.524	8.130	8.251	#N/A	#N/A	#N/A
NRRL3_03279	An12g05700	alpha-L-rhamnosidase-like protein	GH78	1.326	1.329	0.921	#N/A	#N/A	#N/A
NRRL3_03924	An15g04530	glycoside hydrolase family 78 protein	GH78	0.000	0.000	0.000	#N/A	#N/A	#N/A
NRRL3_04245	An07g00240	alpha-L-rhamnosidase-like protein	GH78	0.070	0.077	0.043	#N/A	#N/A	#N/A
NRRL3_06304	An10g00290	alpha-L-rhamnosidase-like protein	GH78	1.072	1.365	1.259	#N/A	#N/A	#N/A
NRRL3_07520	An04g09070	alpha-L-rhamnosidase	GH78	2.228	2.178	1.749	#N/A	#N/A	#N/A

Table 1. Transcriptome and exoproteome analyses of predictive pectinases, GA transporters and catabolic pathway genes, *gaaR* and *gaaX* in the reference, *OEgaaR* and *ΔgaaX* strains. Transcript and extracellular protein levels are represented as TPM and normalized protein (EIC) area values, respectively. Transcript levels with a fold change≥4 and FDR≤0.001 compared to the reference strain are indicated with an asterisk. Genes belonging to the GaaR/GaaX panregulon (Niu et al. 2017) are written in bold. Transcriptome data for the reference and *ΔgaaX* strains were taken from Niu et al. 2017. (Continued)

Gene ID NRRL3	Gene ID CBS 513.88	Description	CAZy family	Transcript level (ave TPM)			Extracellular protein level (ave normalized protein (EIC) area)		
				Reference	*ΔgaaX*	*OEgaaR*	Reference	*ΔgaaX*	*OEgaaR*
NRRL3_11451	An08g09140	alpha-L-rhamnosidase-like protein	GH78	0.085	0.166	0.009	#N/A	#N/A	#N/A
NRRL3_10558	**An18g04800**	**alpha-L-rhamnosidase**	GH78	0.350	3.847*	17.470*	#N/A	#N/A	#N/A
NRRL3_01739	An01g01340	glycoside hydrolase family 88 protein	GH88	0.163	0.219	0.212	#N/A	#N/A	#N/A
NRRL3_00824		sialidase-like protein	GH93	149.296	144.208	132.943	#N/A	#N/A	#N/A
NRRL3_00965	**An14g04370**	**pectin lyase Pel1A**	PL1_4	1.657	3.255	332.933*	2.21	2.57	653.22
NRRL3_08767	An03g00190	pectin lyase Pel1B	PL1_4	5.072	5.031	2.944	#N/A	#N/A	#N/A
NRRL3_04153	An15g07160	pectin lyase Pel1C	PL1_4	18.220	18.640	18.246	#N/A	#N/A	#N/A
NRRL3_06269		pectin lyase	PL1_4	0.015	0.033	0.083	#N/A	#N/A	#N/A
NRRL3_01237	**An19g00270**	**pectin lyase**	PL1_4	0.174	1.475*	13.613*	0.00	0.00	20.51
NRRL3_09811	**An11g04030**	**pectin lyase**	PL1_4	0.000	0.011	0.503*	#N/A	#N/A	#N/A
NRRL3_06359	An10g00870	pectate lyase Ply1A	PL1_7	0.148	0.134	1.366	#N/A	#N/A	#N/A
NRRL3_00684	An14g01130	rhamnogalacturonan lyase	PL4_1	0.021	0.052	0.320	#N/A	#N/A	#N/A
NRRL3_10115	**An11g00390**	**rhamnogalacturonan lyase**	PL4_3	0.648	0.810	11.176	#N/A	#N/A	#N/A
NRRL3_00958	**An14g04280**	**MFS-type sugar/inositol transporter GatA**		3.472	524.952*	125.820*	#N/A	#N/A	#N/A
NRRL3_08663	**An03g01620**	**MFS-type sugar/inositol transporter**		0.274	5.622*	0.577	#N/A	#N/A	#N/A
NRRL3_04281	**An07g00780**	**MFS-type transporter**		3.110	4.596	13.958*	#N/A	#N/A	#N/A
NRRL3_05650	**An02g07710**	**D-galacturonate reductase GaaA**		19.917	1515.440*	135.009*	#N/A	#N/A	#N/A
NRRL3_06890	**An16g05390**	**L-galactonate dehydratase GaaB**		47.765	6256.695*	782.361*	#N/A	3.78	#N/A
NRRL3_05649	**An02g07720**	**2-keto-3-deoxy-L-galactonate aldolase GaaC**		12.536	2283.765*	240.798*	#N/A	2.23	#N/A
NRRL3_10050	**An11g01120**	**L-glyceraldehyde reductase GaaD**		256.409	2732.370*	570.508	#N/A	6.00	0.24

Table 1. Transcriptome and exoproteome analyses of predictive pectinases, GA transporters and catabolic pathway genes, *gaaR* and *gaaX* in the reference, *OEgaaR* and *ΔgaaX* strains. Transcript and extracellular protein levels are represented as TPM and normalized protein (EIC) area values, respectively. Transcript levels with a fold change≥4 and FDR≤0.001 compared to the reference strain are indicated with an asterisk. Genes belonging to the GaaR/GaaX panregulon (Niu et al. 2017) are written in bold. Transcriptome data for the reference and *ΔgaaX* strains were taken from Niu et al. 2017. (Continued)

Gene ID NRRL3	**Gene ID CBS 513.88**	**Description**	**CAZy family**	**Transcript level (ave TPM)**			**Extracellular protein level (ave normalized protein (EIC) area)**		
				Reference	***ΔgaaX***	***OEgaaR***	**Reference**	***ΔgaaX***	***OEgaaR***
NRRL3_08195	An04g00780	D-galacturonic acid responsive transcription factor GaaR		14.444	18.512	850.848*	#N/A	#N/A	#N/A
NRRL3_08194	**An04g00790**	**Repressor of D-galacturonic acid utilization GaaX**		15.968	0.000	352.560*	#N/A	#N/A	#N/A

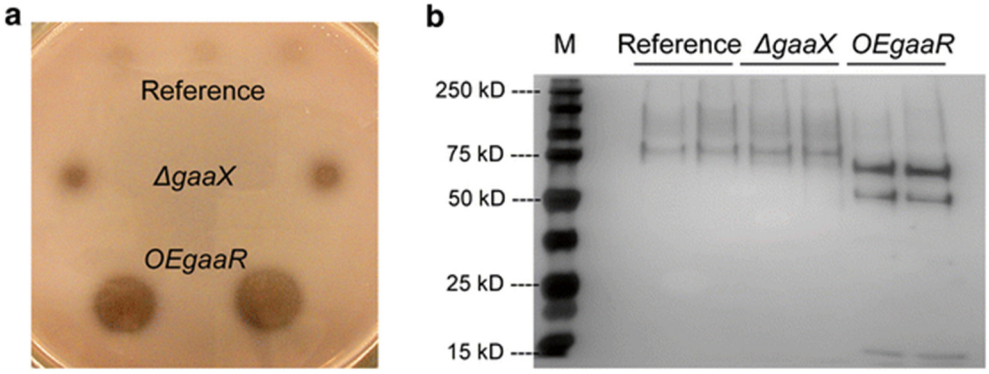

Fig. 1. Enzymatic analysis and secretome profiles of the *A. niger* reference (MA234.1), *ΔgaaX* (JN123.1) and *OEgaaR* (EA21.6) strains grown in bioreactors on 0.75% fructose. **a** PGA plate assay. Supernatant from each bioreactor culture at the sample point following the RNA-seq sample point was spotted on a PGA plate. **b** Silver stained SDS-PAGE patterns of secretomes from a bioreactor culture of each strain at two successive sample points following the RNA-seq sample point. Three micrograms of total protein were loaded in each lane. Marker (M) molecular weight in kD is indicated.

of 24.1 (Online Resource 4 and Table 1). Apart from the aforementioned genes, several genes with unknown relation to GA utilization were also upregulated in the *OEgaaR* strain. These include genes encoding hypothetical/uncharacterized proteins, proteins involved in diverse processes such as dehydrogenases and non-ribosomal peptide synthetases, MFS-type transporters and a Zn_2Cys_6 type transcription factor (NRRL3_11827) (Online Resource 4).

Exoproteome analysis of the *OEgaaR* strain

To support the observed transcriptional upregulation of CAZymes in the *OEgaaR* strain, we analyzed the exoproteome of the *OEgaaR* strain and compared it to the exoproteome of the reference and *ΔgaaX* strains grown in bioreactors on fructose (Online Resource 5). Mass spectrometric analysis revealed 18 pectinases in the exoproteome of the *OEgaaR* strain. Seventeen of them were secreted at higher levels compared to the reference strain and the *ΔgaaX* strain. The protein level of the putative pectin methylesterase NRRL3_05252 was higher in *ΔgaaX* than in the *OEgaaR* strain similar to observed higher mRNA level of this gene in *ΔgaaX* (Table 1). Fifteen of these detected pectinases were also transcriptionally upregulated in the *OEgaaR* strain. In addition, eight genes encoding CAZymes that were expressed at higher levels in the *OEgaaR* strain compared to the reference strain and *ΔgaaX* were found to accumulate at higher levels in the culture media of the *OEgaaR* strain (Table 2).With regard to the degradation of pectin, there is a good correspondence between the upregulated expression of genes and the increased extracellular accumulation of their encoding pectinolytic enzymes in the *OEgaaR* strain, for example the pectinases Pel1A, Pga28E, and Pga28C (Table 1). The distinct SDS-PAGE profile of the *OEgaaR* strain compared to the reference strain and *ΔgaaX* might represent the differences in abundance of the aforementioned extracellular proteins, such as

Table 2. Transcriptome and exoproteome analyses of 20 additional CAZymes upregulated in the *OEgaaR* strain compared to the reference strain. Transcript and extracellular protein levels are represented as TPM and normalized protein (EIC) area values, respectively. Transcript levels with a fold change≥4 and FDR≤0.001 compared to the reference strain are indicated with an asterisk. Genes belonging to the GaaR/GaaX panregulon (Niu et al. 2017) are written in bold. Transcriptome data for the reference and *ΔgaaX* strains were taken from Niu et al. 2017.

Gene ID NRRL3	Gene ID CBS 513.88	Description	CAZy family	Transcript level (ave TPM)			Extracellular protein level (ave normalized protein (EIC) area)		
				Reference	*ΔgaaX*	*OEgaaR*	Reference	*ΔgaaX*	*OEgaaR*
NRRL3_00814	An14g02670	lytic polysaccharide monooxygenase	AA9	1.053	2.625	20.043*	#N/A	#N/A	#N/A
NRRL3_06379		acetylesterase	CE16	0.042	0.142	24.511*	0.00	0.00	0.84
NRRL3_06053	**An02g02540**	**carbohydrate esterase family 16 protein**	CE16	2.055	17.617*	105.698*	0.00	12.71	71.49
NRRL3_01918	An01g03340	xyloglucanase Xeg12A	GH12	9.203	11.484	911.087*	#N/A	#N/A	#N/A
NRRL3_02746	An01g13610	glucan 1,4-alpha-maltohexaosidase	GH13_5	2.813	5.088	15.211*	#N/A	#N/A	#N/A
NRRL3_01212		chitinase-like protein	GH18	8.261	25.549	43.967*	0.00	0.00	0.34
NRRL3_06419	An17g00300	glycoside hydrolase family 3 protein	GH3	2.992	4.962	30.714*	0.00	0.14	3.32
NRRL3_00268	An09g03300	alpha-xylosidase Axl31A	GH31	1.211	3.397	8.676*	0.00	0.00	0.00
NRRL3_02630	An01g12150	beta-galactosidase Lac35B	GH35	1.589	2.203	33.505*	#N/A	#N/A	#N/A
NRRL3_02479	**An01g10350**	**exo-beta-1,4-galactanase**	GH35	3.480	22.294*	24.938*	0.00	1.39	0.00
NRRL3_06244	**An02g00140**	**glycoside hydrolase family 43 protein**	GH43	0.812	22.185*	22.908*	#N/A	#N/A	#N/A
NRRL3_04608	An07g04930	glycoside hydrolase family 43 protein	GH43	1.269	2.703	5.432*	#N/A	#N/A	#N/A
NRRL3_06791	An16g06800	glycoside hydrolase family 5 protein	GH5_5	22.647	68.809	665.632*	24.31	128.89	700.71
NRRL3_10865	**An08g01710**	**alpha-arabinofuranosidase**	GH51	0.795	11.209*	10.280*	#N/A	#N/A	#N/A
NRRL3_03768	An15g02300	alpha-arabinofuranosidase Abf54B	GH54	0.807	3.673	10.070*	0.38	6.53	2.49
NRRL3_01787	An01g01870	xyloglucanase Xeg74C	GH74	0.055	0.174	4.919*	#N/A	#N/A	#N/A
NRRL3_00701	An14g01330	glycoside hydrolase family 79 protein	GH79	0.070	0.460	8.260*	#N/A	#N/A	#N/A

Table 2. Transcriptome and exoproteome analyses of 20 additional CAZymes upregulated in the *OEgaaR* strain compared to the reference strain. Transcript and extracellular protein levels are represented as TPM and normalized protein (EIC) area values, respectively. Transcript levels with a fold change≥4 and FDR≤0.001 compared to the reference strain are indicated with an asterisk. Genes belonging to the GaaR/GaaX panregulon (Niu et al. 2017) are written in bold. Transcriptome data for the reference and *ΔgaaX* strains were taken from Niu et al. 2017. (Continued)

Gene ID NRRL3	Gene ID CBS 513.88	Description	CAZy family	Transcript level (ave TPM)			Extracellular protein level (ave normalized protein (EIC) area)		
				Reference	*ΔgaaX*	*OEgaaR*	Reference	*ΔgaaX*	*OEgaaR*
NRRL3_05305	An02g11890	beta-glucuronidase Gus79A	GH79	0.209	0.447	2.675*	0.00	0.00	0.38
NRRL3_07382	**An16g00540**	**alpha-L-fucosidase-like protein**	GH95	0.036	0.667*	2.139*	0.00	0.00	0.32
NRRL3_07089	An16g02760	alpha-L-fucosidase	GH95	1.619	2.893	11.179*	0.00	0.00	0.61

NRRL3_06791 and Pel1A with predicted molecular weights of unglycosylated proteins of 54.8 and 39.7 kDa, respectively (Fig. 1b).

Nuclear concentration of GaaR is increased in the *OEgaaR* strain

Strains expressing an eGFP-tagged *gaaR*, directed by either the endogenous *gaaR* promoter (*eGFP-gaaR*) or the strong constitutive *A. nidulans gpdA* promoter (*OEeGFP-gaaR*), were constructed to investigate the subcellular localization of GaaR in the reference and *OEgaaR* strains, respectively. Expression of eGFP-tagged *gaaR* in a *ΔgaaR* background resulted in partial complementation of growth on GA and full complementation of growth on PGA and AP, in both *eGFP-gaaR* (EA19.2) and *OEeGFP-gaaR* (EA20.10) strains (Fig. ESM_3.2 a). The polygalacturonase activity in the culture supernatant of the *OEeGFP-gaaR* strain grown in fructose was higher compared to the *eGFP-gaaR* strain, resembling the polygalacturonase production capacities of the *OEgaaR* and reference strains, respectively (Fig. ESM_3.3 c). These results indicate that the eGFP-tagged GaaR is able to activate the transcription of the GA-responsive genes required for growth on GA-containing carbon sources, and that overexpression of the eGFP-tagged *gaaR* results in an increased accumulation of pectinases as in the overexpression of the untagged *gaaR*.

The subcellular localization of GaaR and GaaX was analyzed qualitatively using confocal laser scanning microscopy. As a nuclear marker suitable for co-localization experiments, the SYTO59 dye was used. The nuclear localization of the SYTO59 dye was confirmed in an *A. niger* strain harboring the eGFP-tagged H2B protein (MA26.1) (Fig. ESM_3.4 a). The *eGFP-gaaR*, *OEeGFP-gaaR*, and *gaaX-eGFP* (JN126.2) strains were grown in GA and nuclei were stained with SYTO59 (Fig. ESM_3.4 b). Both eGFP-GaaR and GaaX-eGFP were found to be present in the cytoplasm and nucleus, although we cannot exclude the possibility that eGFP was cleaved off from the fusion proteins and resulted in cytoplasmic or nuclear fluorescence signal. The co-localization experiment showed that eGFP-GaaR was mainly localized in the nucleus in both *eGFP-gaaR* and *OEeGFP-gaaR* strains. In the *gaaX-eGFP* strain grown in GA, GaaX-eGFP was present in both the cytoplasm and the nuclei at roughly the same level.

We next quantified the cytoplasmic and nuclear eGFP-GaaR intensity in the *eGFP-gaaR* and *OEeGFP-gaaR* strains grown in GA or fructose. As shown in Fig. 2, nuclear eGFP-GaaR fluorescence was higher than the cytoplasmic intensity regardless of the presence of an inducing carbon source or the promoter used to overexpress eGFP-tagged *gaaR*. The GFP signals in the *eGFP-gaaR* strain were low after growth in GA or fructose, confirming previous findings that *gaaR* is expressed at low levels on both GA and fructose (Alazi et al. 2016 and Table 1). Overexpression of *eGFP-gaaR* resulted in a much higher nuclear eGFP-GaaR concentration in the *OEeGFP-gaaR* strain than in the *eGFP-gaaR* strain, while only a slight increase in the cytoplasmic concentration was observed. This indicates that the excess eGFP-GaaR produced in the *OEeGFP-gaaR* strain localizes mainly in the nucleus. This result is in line with the observation that in a *Botrytis cinerea* strain overexpressing *BcgaaR-eGFP*, BcGaaR-eGFP mainly localizes in the nucleus under inducing or non-inducing conditions (Zhang et al. 2016).

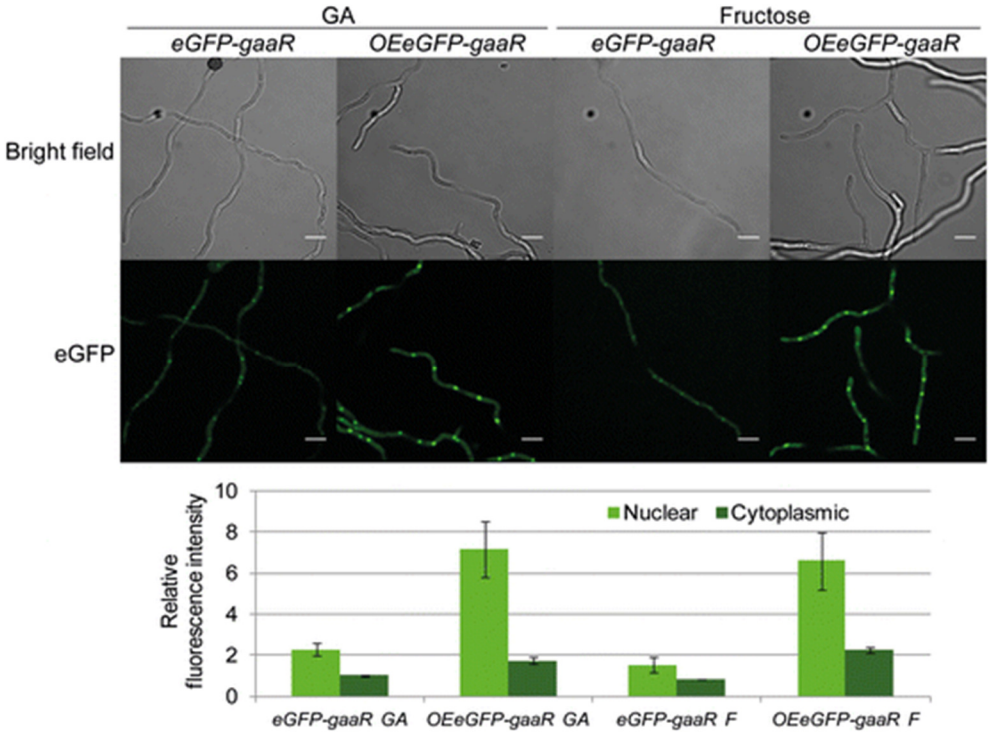

Fig. 2. Nuclear and cytoplasmic fluorescence intensity of the eGFP-tagged GaaR protein. The *eGFP-gaaR* (EA19.2) and *OEeGFP-gaaR* (EA20.10) strains were grown in MM containing 50 mM GA and 1mM fructose, or 50 mM fructose (F) for 17.5 h. Example micrographs representing each condition are shown. Bars represent averages of two biological replicates and standard deviation is shown. Data is represented relative to the cytoplasmic fluorescence intensity in the *eGFP-gaaR* strain on GA. Scale bar 10 µm.

Deletion of *creA* in the *OEgaaR* strain results in elevated production of pectinases

Carbon catabolite repression in the presence of glucose on several GA-responsive genes encoding exo-polygalacturonases, e.g., NRRL3_03144 (*pgaX*) and *pgx28B* (*pgxB*), was previously shown to be CreA-mediated (Niu et al. 2015). The *OEgaaR* strains produced more polygalacturonases during growth in fructose compared to glucose, indicating that fructose exerts less repression than glucose on pectinase gene expression (Fig. ESM_3.2 a). To investigate to which extent the presence of fructose affects CreA-mediated carbon catabolite repression on pectinase gene expression, we used promoter-reporter strains PNRRL3_03144-*amdS* and P*pgx28B-amdS*, which are able to grow on acetamide as the sole nitrogen source only when the *amdS* gene encoding the acetamidase enzyme is expressed via the GA-responsive promoters of the pectinase genes NRRL3_03144 and *pgx28B*, respectively (Niu et al. 2015). Growth of the promoter-reporter strains on plates containing acetamide and GA decreased as the fructose concentration in the growth media increased (Fig. 3a), indicating that fructose also represses the expression of genes encoding those pectinases. In addition, we directly compared the repression power of

glucose, fructose, sorbitol, and acetate on NRRL3_03144 expression in a single experiment (Fig.ESM_3.5). Radial growth assay confirmed that the expression NRRL3_03144 is repressed strongly by glucose and mildly by fructose. Sorbitol and acetate exerted negligible repression on NRRL3_03144 (Fig. ESM_3.5). Deletion of *creA* restored the growth of the promoter-reporter strains on fructose, showing that fructose-imposed carbon catabolite repression on pectinase gene expression is also mediated by CreA (Fig. 3a).

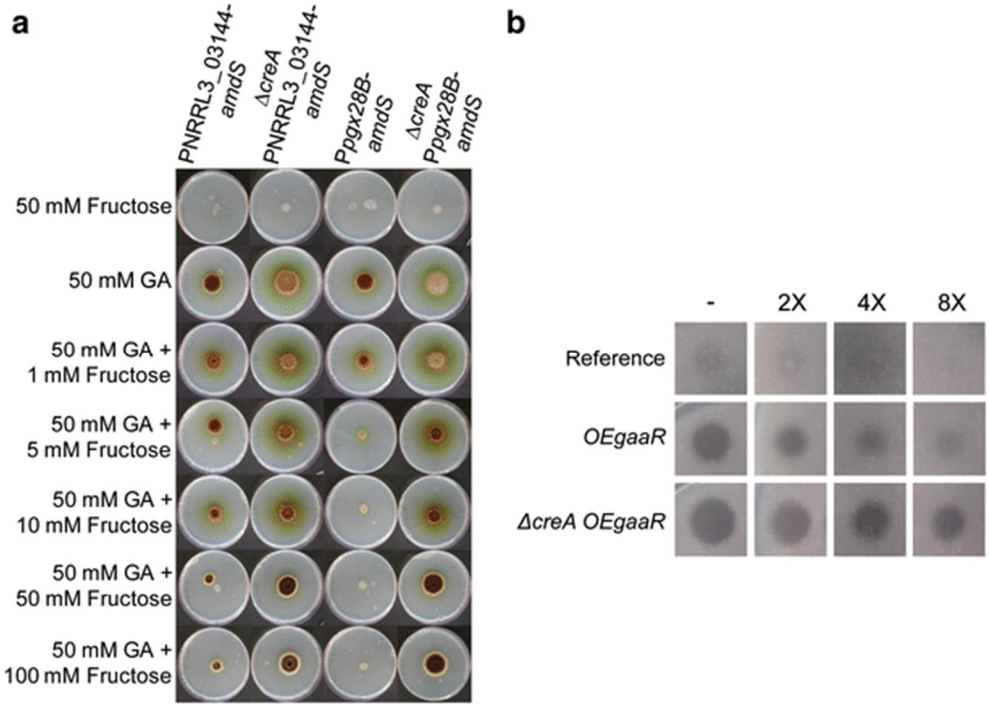

Fig. 3. Analysis of CreA-mediated carbon catabolite repression on pectinase genes. **a** Growth phenotype of the PNRRL3_03144-*amdS* (JC1.5), *ΔcreA* PNRRL3_03144-*amdS* (JN29.2), P*gx28B-amdS* (JC3.6) and *ΔcreA* P*gx28B-amdS* (JN31.3) strains on solid MM containing 50 mM fructose, 50 mM GA, or 50 mM GA with increasing amounts of fructose after 7 days at 30 °C. All plates contain 10 mM acetamide as the sole nitrogen source. **b** PGA plate assay. The reference (MA234.1), *OEgaaR* (EA21.6), and *ΔcreA OEgaaR* (TK2.1) strains were grown in MM containing 50 mM fructose for 36 h, and serial dilutions of culture supernatants were spotted on PGA plates. Dilution factors are indicated.

All 124 genes that were upregulated in the *OEgaaR* strain in fructose and the promoter regions of which could be screened for the presence of transcription factor binding sites, contain at least one CreA binding motif (SYGGRG) (Cubero and Scazzocchio 1994) in their 1 kb upstream sequences (Online Resource 4), suggesting that carbon repression has a major effect on the expression on these GaaR target genes. Because the presence of fructose has a repressing effect on the expression of pectinase genes such as exopolygalacturonases, NRRL3_03144 and *pgx28B*, we hypothesized that deletion of *creA*

in the *OEgaaR* background would result in an elevated expression of pectinase genes on fructose. Therefore, the *ΔcreA OEgaaR* strain (TK2.1) was created in the EA21.6 background to allow a direct comparison. Growth analysis on plates showed a reduced growth of the *ΔcreA OEgaaR* strain on glucose, fructose, and AP, which was also observed in a control *ΔcreA* strain (MA342.2) indicating that the reduced growth is caused by the *creA* deletion and not by *gaaR* overexpression (ESM_3.2 a).To assess the effect of *creA* deletion on pectinase production in combination with *gaaR* overexpression, the reference strain (MA234.1), *OEgaaR* (EA21.6) and *ΔcreA OEgaaR* (TK2.1) strains were grown in fructose and the polygalacturonase activity in the culture supernatants was analyzed via a PGA plate assay. The culture supernatant of the *ΔcreA OEgaaR* strain displayed the highest polygalacturonase activity, thereby providing additional evidence that fructose exerts repression on pectinase gene expression through CreA (Fig. 3b).

DISCUSSION

The Zn_2Cys_6 transcriptional activator GaaR (Alazi et al. 2016) and the repressor protein GaaX (Niu et al. 2017) are the two important players in the transcriptional regulation of the GA-responsive genes in *A. niger*. Both GaaR and GaaX are highly conserved in filamentous fungi of the phylum ascomycetes. Therefore, the GaaR/GaaX module is expected to be the main regulatory mechanism in controlling GA-induced gene expression in filamentous fungi of ascomycetes. The combination of an activator (GaaR) and repressor (GaaX) protein to control gene expression represents a conserved mechanism which shows striking similarities with the regulation of genes involved in quinic acid utilization in *A. nidulans* (Lamb et al. 1996). In both the regulation of GA-responsive genes as well as in quinic acid-responsive genes, loss of function of the respective repressor proteins results in constitutive and inducer-independent expression of target genes. Importantly, the induced expression still requires the corresponding transcriptional activator (Grant et al. 1988; Niu et al. 2017). These observations suggest a model in which the transcriptional activator is kept inactive by its corresponding repressor protein under non-inducing conditions. Upon inducing conditions, an inducer molecule is expected to bind to the repressor thereby causing its dissociation from the transcriptional activator. Non-repressor bound activator is expected to be active as a transcription factor to induce the expression of target genes (Lamb et al. 1996; Niu et al. 2017).

In this study, we constructed several *A. niger* strains that overexpress *gaaR* via the *A. nidulans gpdA* promoter. The *OEgaaR* strains, carrying different copy numbers of the ectopically integrated *gaaR* overexpression construct, showed partial and different levels of complementation of growth on GA-containing carbon sources, whereas their growth on glucose or fructose was similar to the reference strain. While in the wild type high levels of pectinases are produced only under inducing conditions, the *OEgaaR* strains secreted high levels of polygalacturonases under both inducing and noninducing conditions. These results imply that the *OEgaaR* strains possess a functional GaaR that is able to activate the expression of genes required for growth on GA and genes encoding polygalacturonases. Among all *OEgaaR* strains, EA21.6 and TK1.1 displayed the most impaired growth on GA, PGA, and AP and produced the highest levels of polygalacturonases in fructose. This might indicate a possible cofactor imbalance due to increased amounts of NAD(P)H-dependent

GaaA and NADPH-dependent GaaD enzymes, or accumulation of a toxic GA catabolic pathway intermediate, when *OEgaaR* strains grow on GA-rich carbon sources.

The GA-responsive genes are transcriptionally induced by GaaR under inducing conditions (Alazi et al. 2016; Martens-Uzunova and Schaap 2008) and the transcriptional activity of GaaR is suggested to be controlled by GaaX, possibly via a protein-protein interaction, under non-inducing conditions (Niu et al. 2017). We showed that eGFP-GaaR is localized mainly in the nucleus under both inducing and non-inducing conditions, indicating that the transcriptional activity of GaaR is not regulated through nuclear translocation upon induction and the mechanism which keeps GaaR inactive under noninducing conditions is likely to occur in the nucleus. Prediction of nuclear localization signals using the prediction tool NucPred (Brameier et al. 2007) indicated that GaaR likely localizes in the nucleus (score of 0.90) whereas GaaX (score of 0.27) is expected to spend less time in the nucleus. Nevertheless, GaaX-eGFP was found to be present in both cytoplasm and nucleus under inducing conditions, showing that it can enter the nucleus (Fig. ESM_3.4 b). These results imply that GaaX might inhibit the transcriptional activity of GaaR in the nucleus under non-inducing conditions.

Ectopic integration of *gaaR* in a *ΔgaaR* strain was previously shown to result in a full complementation of growth on GA (Alazi et al. 2016), whereas the *eGFP-gaaR* strain EA19.2 that was derived from a *ΔgaaR* strain and expresses N-terminally eGFP-tagged *gaaR* displayed a slightly reduced growth on GA compared to the reference strain. N-terminal eGFP-tagging might result in a minor decrease in GaaR transcription factor activity and therefore partial restoration of growth. As GaaR is expected to interact with GaaX, it was also assessed whether the N-terminal tagging of GaaR influenced its interaction with GaaX. The expression of *eGFP-gaaR* via the endogenous *gaaR* promoter did not result in a constitutive expression of the genes encoding polygalacturonases (Fig. ESM_3.3 c), indicating that eGFP-GaaR activity is properly controlled by GaaX under non-inducing conditions.

Overexpression of eGFP-GaaR driven by the *A. nidulans gpdA* promoter leads to a much higher nuclear eGFP-GaaR concentration under both inducing and non-inducing conditions compared to expression driven by the endogenous *gaaR* promoter. The increase in nuclear GaaR concentration was accompanied by transcriptional upregulation of GA-responsive genes and the increased accumulation of pectinases in the extracellular medium. The transcriptional activation of GA-responsive genes in the *OEgaaR* strain under non-inducing conditions can be explained by the possibility that the excess of GaaR titrates out the concentration of GaaX and escapes GaaX inhibition, even though *gaaX* is induced upon GaaR overexpression.

Genome-wide gene expression analysis in the reference strain grown in GA has been previously performed (Alazi et al. 2016). Direct comparison of the gene expression values between the study of Alazi et al. (2016), and this study needs careful interpretation due to different experimental setups (growth in shake flasks vs bioreactors) and representation of transcript levels (FPKM vs TPM). Notwithstanding, it can be observed that the expression level of the genes encoding pectinases are generally comparable between the reference and *OEgaaR* strains grown under inducing and noninducing conditions, respectively. However, drastically higher expression of NRRL3_05252, NRRL3_03144,

pgx28B, and *gan53A* in the reference strain and *pga28C*, *pga28E*, and *pel1A* in the *OEgaaR* strain was also observed.

Elimination of the repressing activity of GaaX by deleting *gaaX* was previously shown to be another way to activate the expression of GA-responsive genes under non-inducing conditions (Niu et al. 2017). The concentration of the nuclear GaaR in *ΔgaaX* is expected to be similar to wild type and much less compared to the *OEgaaR* strain. Only nine out of 48 genes encoding pectinases were upregulated in the *ΔgaaX* strain in fructose compared to the reference strain, and the transcript and extracellular protein levels of these pectinases were generally lower compared to the *OEgaaR* strain. This indicates that the nuclear concentration of active GaaR is indeed an important factor for transcriptional activation of GA-responsive genes. On the other hand, the genes encoding the (putative) GA transporters and catabolic pathway enzymes were expressed at higher levels in *ΔgaaX* compared to the *OEgaaR* strain, indicating that factors other than GaaR concentration might play a role in the regulation of these genes.

RNA-seq analysis showed that besides the genes encoding pectinases, 20 genes predicted to encode CAZymes involved in the degradation of multiple substrates or specifically of cellulose, starch, or xyloglucan were upregulated in the *OEgaaR* strain in fructose. This indicates that these enzymes might be involved or assist in enabling the degradation of pectin. Five of these CAZymes were shown to be upregulated in *ΔgaaX*, and therefore designated as part of the GaaR/GaaX panregulon (Niu et al. 2017). In addition, many of the 20 additional CAZymes were reported to be potentially regulated by transcription factors AraR and/or XlnR (Gruben et al. personal communication). However, the expression of the genes encoding AraR or XlnR were not significantly changed in *OEgaaR*, discounting the possibility that overexpression of *gaaR* caused transcriptional upregulation of the genes encoding CAZymes via their specific transcription factors.

Fructose was found to exert CreA-mediated repression of gene expression in case of the genes NRRL3_03144 (*pgaX*) and *pgx28B* (*pgxB*) encoding exo-polygalacturonases, which were previously shown to be strongly repressed in the presence of glucose (Niu et al. 2015). The repression power of fructose was lower than that of glucose (Fig. ESM_3.5). As shown by Niu et al. (2017), deletion of *creA* is not sufficient for an increased production of polygalacturonases under non-inducing conditions, showing that GA-responsive gene expression requires the presence of active GaaR relieved from GaaX inhibition. A similar phenomenon was previously observed in *T. reesei*, where high expression of the genes encoding cellulases in a Cre1-disrupted strain required the presence of the transcriptional activator Xyr1 under non-inducing conditions (Wang et al. 2013). The strain that lacks *creA* and overexpresses *gaaR* (TK2.1) secreted higher levels of polygalacturonases compared to the reference and *OEgaaR* strains, indicating that CreA substantially represses the expression of GA-responsive genes in the presence of fructose even when GaaR is abundant.

To conclude, genetic evidence suggests that the activity of the GaaR transcription factor is negated by the action of the GaaX repressor protein. Either deletion of GaaX or overexpression of GaaR results in a constitutive expression of GaaR/GaaX target genes. The simplest interpretation of these observations is that GaaX mediates its repressing activity by a direct interaction with GaaR. Loss of function of GaaX or overexpression of

GaaR will affect the stoichiometry of GaaR-GaaX and lead to high levels of "repressor free" GaaR which is expected to act as an active transcription factor to induce expression of GA-responsive genes. We have shown that overexpression of GaaR leads to an increased level of pectinase production under non-inducing conditions, and that deletion of *creA* further increases the pectinase production capacity of *A. niger*. The *ΔcreA OEgaaR* strain represents an interesting strain for applications in industry with its high pectinase production capacity in the absence of an inducing carbon source and in the presence of a repressing carbon source.

NOTE AFTER PUBLICATION

In RNA-seq experiments, two or three biological replicates representing each condition were used, and differential gene expression was identified using DESeq2 (FDR≤0.001 and FC≥4). This approach allows addressing only major differences in gene expression.

FUNDING

EA was supported by a grant from BE-Basic (Flagship 10). This work was in part supported by the NSERC IBN Strategic Network and Genome Canada.

ELECTRONIC SUPPLEMENTARY MATERIAL

The online version of this article (https://doi.org/10.1007/s00253-018-8753-7) contains supplementary material, which is available to authorized users.

REFERENCES

Abramoff MD, Magalhaes PJ, Ram SJ (2004) Image processing with ImageJ. Biophoton Int 11(7):36–42

Alazi E, Niu J, Kowalczyk JE, Peng M, Aguilar Pontes MV, van Kan JAL, Visser J, de Vries RP, Ram AFJ (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

Alazi E, Khosravi C, Homan TG, du Pré S, Arentshorst M, Di Falco M, Pham TTM, Peng M, Aguilar-Pontes MV, Visser J, Tsang A, de Vries RP, Ram AFJ (2017) The pathway intermediate 2-keto-3-deoxy-L-galactonate mediates the induction of genes involved in D-galacturonic acid utilization in *Aspergillus niger*. FEBS Lett 591:1408–1418

Arentshorst M, Ram AFJ, Meyer V (2012) Using non-homologous end-joining- deficient strains for functional gene analyses in filamentous fungi. Methods Mol Biol 835:133–150

Arentshorst M, Niu J, Ram AFJ (2015) Efficient generation of *Aspergillus niger* knock out strains by combining NHEJ mutants and a split marker approach. In: van den Berg MA, Maruthachalam K (eds) Genetic transformation systems in fungi, vol 1. Springer International Publishing, Cham, pp 263–272

Bos CJ, Debets AJ, Swart K, Huybers A, Kobus G, Slakhorst SM (1988) Genetic analysis and the construction of master strains for assignment of genes to six linkage groups in *Aspergillus niger*. Curr Genet 14:437–443

Brameier M, Krings A, MacCallum AM (2007) NucPred—predicting nuclear localization of proteins. Bioinformatics 23(9):1159–1160

Caffall KH, Mohnen D (2009) The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. Carbohydr Res 344:1879– 1900

Chua G, Morris QD, Sopko R, Robinson MD, Ryan O, Chan ET, Frey BJ, Andrews BJ, Boone C, Hughes TR (2006) Identifying transcription factor functions and targets by phenotypic activation. Proc Natl AcadSci U S A 103(32):12045–12050

Coutinho PM, Andersen MR, Kolenova K, Vankuyk PA, Benoit I, Gruben BS, Trejo-Aguilar B, Visser H, van Solingen P, Pakula T, Seiboth B, Battaglia E, Aguilar-Osorio G, de Jong JF, Ohm RA, Aguilar M, Henrissat B, Nielsen J, Stalbrand H, de Vries RP (2009) Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. Fungal Genet Biol 46(1):S161–S169

Cubero B, Scazzocchio C (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(2):407–415

de Vries RP, Riley R, Wiebenga A, Aguilar-Osorio G, Amillis S, Uchima CA, Anderluh G, Asadollahi M, Askin M, Barry K, Battaglia E, Bayram Ö, Benocci T, Braus-Stromeyer SA, Caldana C, Cánovas D, Cerqueira GC, Chen F, Chen W, Choi C, Clum A, Dos Santos RA, Damásio AR, Diallinas G, Emri T, Fekete E, Flipphi M, Freyberg S, Gallo A, Gournas C, Habgood R, Hainaut M, Harispe ML, Henrissat B, Hildén KS, Hope R, Hossain A, Karabika E, Karaffa L, Karányi Z, Kraševec N, Kuo A, Kusch H, LaButti K, Lagendijk EL, Lapidus A, Levasseur A, Lindquist E, Lipzen A, Logrieco AF, MacCabe A, Mäkelä MR, Malavazi I, Melin P, Meyer V, Mielnichuk N, Miskei M, Molnár ÁP, Mulé G, Ngan CY, Orejas M, Orosz E, Ouedraogo JP, Overkamp KM, Park HS, Perrone G, Piumi F, Punt PJ, Ram AF, Ramón A, Rauscher S, Record E, Riaño-Pachón DM, Robert V, Röhrig J, Ruller R, Salamov A, Salih NS, Samson RA, Sándor E, Sanguinetti M, Schütze T, Sepčić K, Shelest E, Sherlock G, Sophianopoulou V, Squina FM, Sun H, Susca A, Todd RB, Tsang A, Unkles SE, van de Wiele N, van Rossen-Uffink D, Oliveira JV, Vesth TC, Visser J, Yu JH, Zhou M, Andersen MR, Archer DB, Baker SE, Benoit I, Brakhage AA, Braus GH, Fischer R, Frisvad JC, Goldman GH, Houbraken J, Oakley B, Pócsi I, Scazzocchio C, Seiboth B, vanKuyk PA, Wortman J, Dyer PS, Grigoriev IV (2017) Comparative genomics reveals high biological diversity and specific adaptations in the industrially and medically important fungal genus *Aspergillus*. Genome Biol 18(1):28

Edwards MC, Doran-Peterson J (2012) Pectin-rich biomass as feedstock for fuel ethanol production. Appl Microbiol Biotechnol 95:565–575

Grant S, Roberts CF, Lamb H, Stout M, Hawkins AR (1988) Genetic regulation of the quinic acid utilization (QUT) gene cluster in *Aspergillus nidulans*. J Gen Microbiol 134:347–358

Hynes MJ, Corrick CM, King JA (1983) Isolation of genomic clones containing the *amdS* gene of *Aspergillus nidulans* and their use in the analysis of structural and regulatory mutations. Mol Cell Biol 3(8):1430–1439

Jiang Y, Duarte AV, van den Brink J, Wiebenga A, Zou G, Wang C, de Vries RP, Zhou Z, Benoit I (2016) Enhancing saccharification of wheat straw by mixing enzymes from genetically-modified *Trichoderma reesei* and *Aspergillus niger*. Biotechnol Lett 38(1):65–70

Kashyap D, Vohra P, Chopra S, Tewari R (2001) Applications of pectinases in the commercial sector: a review. Bioresour Technol 77:215–227

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

Lamb HK, Newton GH, Levett LJ, Cairns E, Roberts CF, Hawkins AR (1996) The QUTA activator and QUTR repressor proteins of *Aspergillus nidulans* interact to regulate transcription of the quinate utilization pathway genes. Microbiology 142:1477–1490

Lv X, Zheng F, Li C, Zhang W, Chen G, Liu W (2015) Characterization of a copper responsive promoter and its mediated overexpression of the xylanase regulator 1 results in an induction-independent production of cellulases in *Trichoderma reesei*. Biotechnol Biofuels 8:67

Martens-Uzunova ES, Schaap PJ (2008) An evolutionary conserved D-galacturonic acid metabolic pathway operates across filamentous fungi capable of pectin degradation. Fungal Genet Biol 45(11): 1449–1457

Martens-Uzunova ES, Schaap PJ (2009) Assessment of the pectin degrading enzyme network of *Aspergillus niger* by functional genomics. Fungal Genet Biol 46:S170–S179

Meyer V, Arentshorst M, Flitter SJ, Nitsche BM, Kwon MJ, Reynaga-Peña CG, Bartnicki-Garcia S, van den Hondel CAMJJ, Ram AFJ (2009) Reconstruction of signaling networks regulating fungal morphogenesis by transcriptomics. Eukaryot Cell 8(11):1677–1691

Nitsche BM, Crabtree J, Cerqueira GC, Meyer V, Ram AFJ, Wortman JR (2011) New resources for functional analysis of omics data for the genus *Aspergillus*. BMC Genomics 12(1):486

Niu J, Homan TG, Arentshorst M, de Vries RP, Visser J, Ram AF (2015) The interaction of induction and repression mechanisms in the regulation of galacturonic acid-induced genes in *Aspergillus niger*. Fungal Genet Biol 82:32–42

Niu J, Alazi E, Reid ID, Arentshorst M, Punt PJ, Visser J, Tsang A, Ram AFJ (2017) An evolutionarily conserved transcriptional activator-repressor module controls expression of genes for D-galacturonic acid utilization in *Aspergillus niger*. Genetics 205:169–183

Noguchi Y, Sano M, Kanamaru K, Ko T, Takeuchi M, Kato M, Kobayashi T (2009) Genes regulated by AoXlnR, the xylanolytic and cellulolytic transcriptional regulator, in *Aspergillus oryzae*. Appl Microbiol Biotechnol 85(1):141–154

Ogawa M, Kobayashi T, Koyama Y (2012) ManR, a novel $Zn(II)_2Cys_6$ transcriptional activator, controls the *β*-mannan utilization system in *Aspergillus oryzae*. Fungal Genet Biol 49:987–995

Pedrolli DB, Monteiro AC, Gomes E, Carmona EC (2009) Pectin and pectinases: production, characterization and industrial application o microbial pectinolytic enzymes. Open Biotechnol J 3:9–18

Punt PJ, Oliver RP, Dingemanse MA, Pouwels PH, van den Hondel CA (1987) Transformation of *Aspergillus* based on the hygromycin B resistance marker from *Escherichia coli.* Gene 56(1):117–124

Sloothaak J, Schilders M, Schaap PJ, de Graaff LH (2014) Overexpression of the *Aspergillus niger* GatA transporter leads to preferential use of D-galacturonic acid over D-xylose. AMB Express 4(66):1–9

Toushik SH, Lee K, Lee J, Kim K (2017) Functional applications of lignocellulolytic enzymes in the fruit and vegetable processing industries. J Food Sci 88(3):583–595

Vinck A (2007) Hyphal differentiation in the fungal mycelium. Dissertation, Utrecht University

Wang S, Liu G, Wang J, Yu J, Huang B, Xing M (2013) Enhancing cellulase production in *Trichoderma reesei* RUT C30 through combined manipulation of activating and repressing genes. J Ind Microbiol Biotechnol 40(6):633–641

Yuan XL, Goosen C, Kools H, van den Maarel MJEC, van den Hondel CAMJJ, Dijkhuizen L, Ram AFJ (2006) Database mining and transcriptional analysis of genes encoding inulin-modifying enzymes of *Aspergillus niger*. Microbiology 152(10):3061–3073

Zhang L, Lubbers RJM, Simon A, Stassen JHM, Ribera PRV, Viaud M, van Kan JAL (2016) A novel Zn_2Cys_6 transcription factor BcGaaR regulates D-galacturonic acid utilization in *Botrytis cinerea*. Mol Microbiol 100(2):247–262

CHAPTER 7

W361R mutation in GaaR, the regulator of D-galacturonic acid responsive genes, leads to constitutive expression of pectinases in *Aspergillus niger*

Alazi E, Niu J, Otto SB, Pham TTM, Tsang A, Ram AFJ

Submitted to *MicrobiologyOpen*

ABSTRACT

Polysaccharides present in plant biomass, such as pectin, are the main carbon source for filamentous fungi. *Aspergillus niger* naturally secretes pectinases to degrade pectin and utilize the released monomers, mainly D-galacturonic acid. The transcriptional activator GaaR, the repressor of D-galacturonic acid-utilization GaaX, and the physiological inducer 2-keto-3-deoxy-L-galactonate play important roles in the transcriptional regulation of D-galacturonic acid-responsive genes, which include the genes encoding pectinases. In this study, we described the mutations found in *gaaX* and *gaaR* that enabled constitutive (i.e. inducer-independent) expression of pectinases by *A. niger*. Via gene expression analyses using promoter-reporter strains (P*pgaX-amdS*) and PGA-plate assay, we showed that W361R mutation in GaaR results in constitutive expression of pectinase genes. Analysis of subcellular localization of C-terminally eGFP-tagged GaaR/GaaR^{W361R} revealed important differences in nuclear accumulation of N- versus C-terminal eGFP-tagged GaaR.

INTRODUCTION

Polysaccharides present in plant biomass, such as pectin, are the principle carbon source for filamentous fungi (Kowalczyk *et al.*, 2014). All of the four pectin substructures, i.e. polygalacturonic acid (PGA), rhamnogalacturonan I, rhamnogalacturonan II and xylogalacturonan, contain D-galacturonic acid (GA) in their backbones, with various monomers or polymers attached to their backbones. PGA is the major polysaccharide present in pectin, and consists of α-1,4-linked GA subunits (Caffall and Mohnen, 2009).

When *Aspergillus niger* encounters pectin in its environment, it secretes pectinases to degrade pectin, transports the released monosaccharides into the cell and catabolizes them to generate energy. The catabolic intermediate 2-keto-3-deoxy-L-galactonate produced from GA was shown to induce the production of pectinases by *A. niger* (Alazi *et al.*, 2017). The regulated production of pectinases, especially polygalacturonases, under inducing conditions requires the Zn_2Cys_6 type transcription factor GaaR (Alazi *et al.*, 2016), the action of which is suggested to be inhibited by the repressor protein GaaX (Niu *et al.*, 2017). Additionally, the expression of pectinase genes is subject to carbon catabolite repression via the repressor protein CreA in the presence of energetically favorable carbon sources, such as glucose or D-fructose (Niu *et al.*, 2015; Alazi *et al.*, 2018).

Pectinases produced by *A. niger* are used industrially (Kashyap *et al.,* 2001; Toushik *et al.*, 2017; Khan *et al.*, 2013; Edwards and Doran-Peterson, 2012), especially in food industry and in hydrolysis of plant biomass for the subsequent production of biofuel and high-value biopolymers. Inducer-independent production of pectinases by *A. niger* would enable the use of waste and residues from agriculture, forestry and food industries as a cheap and sustainable feedstock for fungal cultivations. In this study, we describe mutations in *gaaX* and *gaaR* that result in constitutive expression of pectinase genes by *A. niger*. A mutation in *gaaR* causing an amino acid change from tryptophan to arginine at position 361 (W361R) is shown to lead to a constitutively active form of the GaaR transcription factor, leading to inducer-independent production of polygalacturonases.

MATERIALS AND METHODS

Strains, media and growth conditions

All strains used in this study are listed in Table S1. Media were prepared as described in Arentshorst *et al.* (2012). Radial growth phenotype of the strains were analysed on minimal medium (MM) (pH 5.8) containing 1.5% (w/v) agar (Scharlau, Barcelona, Spain) and various carbon sources: 50 mM D-glucose (VWR International, Amsterdam, The Netherlands), D-fructose (Sigma-Aldrich, Zwijndrecht, The Netherlands) or GA (Chemodex, St Gallen, Switzerland); or 1% (w/v) PGA (Sigma, Zwijndrecht, The Netherlands) or apple pectin (AP) (Sigma-Aldrich, Zwijndrecht, The Netherlands). Media of the uridine auxotrophic strains were supplemented with 10 mM uridine. Plates were inoculated with 5 µL 0.9% NaCl containing $1x10^4$ or $5x10^4$ freshly harvested spores, and cultivated at 30 °C for 7 or 5 days, respectively. MM (pH 5.8) containing 1.5% (w/v) agar, 10 mM acetamide (Sigma-Aldrich, Steinheim, Germany) as the sole nitrogen source, and glucose, D-fructose, GA or 2-keto-3-deoxy-L-galactonate as the carbon source was prepared as described previously (Arentshorst *et al.*, 2012). Plates were inoculated with 5 µL 0.9% NaCl

containing 10^4 freshly harvested spores and cultivated at 30 °C for 7 or 10 days. Filter sterilized carbon source solutions were added after autoclaving MM containing agar. PGA and AP were autoclaved together with the medium. All growth experiments were performed in duplicate.

For enzymatic analysis, 300 mL shake flasks that include 100 mL MM (pH 5.8) containing 50 mM D-fructose and 0.003% yeast extract were inoculated with $7.5x10^7$ freshly harvested spores and cultivated for 36h in a rotary shaker at 30 °C and 250 rpm. Experiments were performed in duplicate.

For microscopic analysis of the localization of the eGFP-tagged GaaR or GaaR^{W361R} proteins, $4x10^5$ freshly harvested spores were inoculated on cover slips in Petri dishes that include 4 mL MM containing 0.003% yeast extract and 10 mM D-fructose or 2-keto-3-deoxy-L-galactonate, and grown at 30 °C for approximately 21 h. For each condition, two biological replicates were performed.

Sequencing *gaaX* and *gaaR* genes in the mutants with constitutive expression of genes involved in PGA utilization

Sixty five *trans*-acting mutants, spontaneous or obtained after mild UV mutagenesis with constitutive expression of genes involved in PGA utilization were obtained as previously described (Niu *et al*., 2017) on solid MM containing 50 mM glucose as the carbon source and 10 mM acetamide as the sole nitrogen source. Genomic DNA of 11 spontaneous and 53 UV mutants was extracted as described by Arentshorst *et al.* (2012), and *gaaX* and/or *gaaR* genes were PCR-amplified using the primers gaaX_P5f and gaaX_GSP9r, or gaaRP7f and gaaRP8r, respectively (Table S2). The PCR fragments were sequenced in both directions using *gaaX* or *gaaR* sequencing primers, respectively (Table S2).

Construction of the promoter-reporter strains expressing *gaaR^{W361R}*

Protoplast-mediated transformation of *A. niger* and purification of the transformants were performed as described by Arentshorst *et al.* (2012).

The *gaaR^{W361R}* gene together with its 964-bp promoter and 992-bp terminator regions was amplified by PCR using the primers gaaRP5f and gaaRP6r (Table S2) with JN103.1 genomic DNA as template. The PCR product was transformed into strains JC1.5 and JN29.2, yielding in the strains JN130.4 and JN129.1, respectively. Transformants were selected on transformation plates (Arentshorst *et al.*, 2012) containing 10 mM acetamide as the sole nitrogen source. Correct gene replacements in strains JN130.4 and JN129.1 were verified by Southern blot and sequencing analyses. For sequence analysis, the *gaaR* locus was amplified using the primers gaaRP7f and gaaRP8r (Table S2) and JN130.4 and JN129.1 genomic DNA as template, and ligated into pJET1.2/blunt cloning vector (Thermo Fisher Scientific, Carlsbad, California). The resulting plasmids were amplified in *Escherichia coli* and sequenced using *gaaR* sequencing primers. Integration of the *gaaR^{W361R}* into the endogenous *gaaR* locus was confirmed via Southern blot analysis. Genomic DNA was restricted overnight with *Nco*I or *Hin*dIII restriction enzymes. A 501-bp fragment containing the *gaaR* gene was PCR-amplified using the primer pairs listed in Table S2 with

N402 genomic DNA as template, and was used as a probe.

Construction of strains expressing *gaaR-eGFP or gaaR-(GA)$_4$-eGFP*

The plasmid pEA9 containing the P*gaaR-gaaR-eGFP*-T*gaaR* construct was created as follows: The *eGFP* gene and the 715-bp terminator of the *gaaR* gene were amplified by PCR using the primer pairs listed in Table S2 with the plasmid pFG029 (unpublished vector, containing P*gpdA-eGFP*-T*trpC*) and N402 genomic DNA as template, respectively. *eGFP* and T*gaaR* were combined by fusion PCR using primers GaaR_GFP1F and GaaR_GFP3R. The *gaaR* gene without the stop codon and together with the 731-bp promoter region was PCR amplified using the primer pairs listed in Table S2. P*gaaR-gaaR* and *eGFP*-T*gaaR* were combined by fusion PCR using primers eGFP-gaaR-For and eGFP-gaaR-Rev, ligated into pJET1.2/blunt cloning vector (Thermo Fisher Scientific, Carlsbad, California) and amplified in *E. coli*. Sequencing of pEA9 showed that a PCR error (T to C) has occurred in T*gaaR* at a distance of 519 bp downstream of the stop codon of the *eGFP* gene.

The plasmid pEA10 containing the P*gaaR-gaaR-(GA)$_4$-eGFP*-T*gaaR* construct was created in a similar way to pEA9, except that P*gaaR-gaaR-(GA)$_4$* was amplified using the primer GaaR_GFP6R adding a linker containing four repeats of glycine and alanine residues between GaaR and eGFP. Sequencing of pEA10 showed that a PCR error (C to T) has occurred in T*gaaR* at a distance of 430 bp downstream of the stop codon of the *eGFP* gene.

To create the strain EA29.14, pEA9 was co-transformed into strain JN36.1 (*ΔgaaR*) together with the plasmid pMA357 containing the *A. nidulans amdS* gene behind the *A. nidulans gpdA* promoter (Alazi *et al.*, 2016). EA30.6 was created by co-transformation of pEA10 into strain JN36.1 (*ΔgaaR*) together with the plasmid pMA357. Transformants were selected on transformation plates containing acetamide as the sole nitrogen source.

Construction of the *gaaR::AOpyrG* deletion strain

SO1.1 (*ΔgaaR::AOpyrG*) was created using the split marker approach (Arentshorst *et al.*, 2015). 5' and 3' flanks of *gaaR* gene were PCR-amplified using the primer pairs listed in Table S2 with N402 genomic DNA as template. The *A. oryzae pyrG* gene was PCR-amplified as two fragments using the primer pairs listed in Table S2 and the plasmid pAO4-13 (de Ruiter-Jacobs *et al.*, 1989) as template. Split marker fragments with the *AOpyrG* selection marker were created by fusion PCR and used to transform the strain MA169.4, resulting in the strain SO1.1. Proper deletion of *gaaR* was confirmed by diagnostic PCR (data not shown) and via Southern blot analysis.

Construction of strains expressing *gaaR^{W361R}, gaaR^{W361R}-eGFP or gaaR-eGFP* from the endogenous *gaaR* locus

To construct plasmid pSO1.2, the P*gaaR-gaaR^{W361R}*-T*gaaR* allele was amplified by PCR using the primers gaaRP5f and gaaRP6r (Table S2) with JN103.1 genomic DNA as template and ligated into pJET1.2/blunt cloning vector (Thermo Fisher Scientific, Carlsbad, California). The resulting plasmid pSO1.2 was amplified in *E. coli*. pSO2.1 (containing the

P*gaaR-gaaR*^{W361R}*-eGFP*-T*gaaR* construct) was created by digesting the plasmids pSO1.2 and pEA9 (containing the P*gaaR-gaaR-eGFP*-T*gaaR* construct) with the restriction enzymes *Bcu*I and *Bam*HI, and by ligating the 2408-bp *Bcu*I-*Bam*HI fragment from pSO1.2 (containing the *gaaR*^{W361R} mutation) with the *Bcu*I and *Bam*HI cut opened pEA9. pSO1.2 and pSO2.1 were sequenced to ensure no PCR errors have occurred and proper ligation and orientation of the fragments.

Repair fragments for CRISPR-Cas9 mediated targeted integration (see below) were obtained as follows: The P*gaaR-gaaR*^{W361R}-T*gaaR* construct containing the 964-bp P*gaaR* and 992-bp T*gaaR* was excised from pSO1.2 using the restriction enzymes *Not*I and *Xba*I. The P*gaaR-gaaR-eGFP*-T*gaaR* and P*gaaR-gaaR*^{W361R}*-eGFP*-T*gaaR* constructs containing the 731-bp P*gaaR* and 419-bp T*gaaR* were excised from pEA9 and pSO2.1, respectively, using the restriction enzyme *Bgl*II.

The strains SO2.1 (*gaaR*^{W361R}), EA31.1 (*gaaR-eGFP*) and EA32.1 (*gaaR*^{W361R}*-eGFP*) were created using the CRISPR-Cas9 technique (Nødvig *et al.*, 2015; Song *et al.*, manuscript in preparation). A guide RNA protospacer sequence GTGGATACGTACTCCTTTTA targeting the *AOpyrG* gene was designed using E-CRISP (Heigwer *et al.*, 2014).

To create SO2.1, the DNA template for the *in vitro* synthesis of the guide RNA via the T7 promoter was amplified by PCR using the primers sgRNAP1 and OTL19 (Table S2), and the plasmid p426-SNR52p-gRNA.CAN1.Y-SUP4t (Addgene, USA) (DiCarlo *et al.*, 2013) as template. The PCR consisted of the following reactions: Initial denaturation at 98 °C for 30 sec; 30 cycles of denaturation at 98 °C 5 sec, annealing at 77 °C 5 sec, extension at 72 °C 2 sec; and final extension at 72 °C 30 sec. The amplified DNA template was transcribed *in vitro* using the MEGAscript T7 Kit (Thermo Fisher Scientific, Carlsbad, California). P*gaaR-gaaR*^{W361R}-T*gaaR* repair fragment was co-transformed into strain SO1.1 (*ΔgaaR::AOpyrG*) together with the *in vitro* transcribed guide RNA and the plasmid pFC332 containing the *cas9* gene and the hygromycin selection marker (Nødvig *et al.*, 2015), yielding to SO2.1. Transformants were selected on transformation plates containing hygromycin and uridine, and purified twice on MM containing GA, 5'-fluoroorotic acid, uridine and hygromycin. pFC332 was cured by further purifying transformants strains twice without hygromycin selection on MM containing glucose and uridine. Transformants were not able to grow on MM containing glucose, uridine and hygromycin after the first round of purification, indicating that pFC332 was successfully cured.

The DNA template for the *in vivo* synthesis of the guide RNA via the RNA polymerase III promoter was amplified in two parts by PCR using the primers Fw_LIC2 and Rev_P1, and Fw_P1 and Rev_LIC2 (Table S2) using the plasmid ANEp8_Cas9-gRNAalbA (Song *et al.*, manuscript in preparation) as template. The two PCR products were combined by fusion PCR using the primers Fw_LIC2 and Rev_LIC2. The fusion PCR product was amplified by PCR using the primers for_pTE1 and rev_pTE1 to introduce *Pac*I restriction sites at both ends, digested with *Pac*I, and ligated into *Pac*I cut opened pFC332, yielding to plasmid pTE1. pTE1 was cotransformed together with P*gaaR-gaaR-eGFP*-T*gaaR* or P*gaaR-gaaR*^{W361R}*-eGFP*-T*gaaR* repair fragments into the strain SO1.1 (*ΔgaaR::AOpyrG*) to create EA31.1 or EA32.1, respectively. Transformants were selected on transformation plates containing hygromycin and uridine, and purified twice on MM containing GA, 5'-fluoroorotic acid and uridine. Transformants were not able to grow on MM containing GA,

5'-fluoroorotic acid, uridine and hygromycin after the first round of purification, indicating that pTE1 was successfully cured.

The *gaaR* locus in SO2.1 and EA31.1 was amplified using the primers gaaRP7f and gaaRP8r, and the one in EA32.1 using the primers PgaaR_seq_P2 and TgaaR_seq_P2, and ligated into pJET1.2/blunt cloning vector (Thermo Fisher Scientific, Carlsbad, California). The resulting plasmids were amplified in *E. coli* and sequenced using *gaaR* sequencing primers to confirm the presence of the mutated *gaaR* allele in SO2.1 and EA32.1, and the wild type *gaaR* allele in EA31.1. Integration of the repair fragments into the endogenous *gaaR* locus in SO2.1, EA31.1 and EA32.1 was confirmed via Southern blot analysis (Figure S3). SO2.1, EA31.1 and EA32.1 were made uridine prototroph by transforming with the plasmid pAB4.1 (van Hartingsveldt *et al.*, 1987), yielding to strains EA34.1, EA36.1 and EA39.1, respectively.

Enzymatic analysis

Supernatants from shake flask cultures were obtained by filtration through glass microfiber filters (Whatman, Buckinghamshire, UK), and the filtrate was stored at -80 °C. PGA plate assays were performed as described by Niu *et al.* (2017). Twenty-five microlitres of supernatant from each culture was spotted on plates containing 0.2% PGA, and plates were incubated at 37 °C for 24 h before staining with 0.1% Congo Red (Sigma-Aldrich, Zwijndrecht, The Netherlands) solution.

Microscopy

The cover slips with adherent germlings were placed upside down on glass slides. The GFP fluorescence was excited using 488 nm laser line in a Zeiss Observer confocal laser scanning microscope (Zeiss, Jena, Germany). The images were analysed with the ImageJ software (Abramoff *et al.*, 2004). On each image, the exact same brightness and contrast adjustments were applied.

RESULTS

Genetic characterization of *A. niger* mutants showing constitutive expression of pectinases

Mutants showing constitutive expression of pectinases were previously obtained via a forward genetic screen using the promoter-reporter strain P*pgaX-amdS ΔcreA* (JN29.2) (Niu *et al.*, 2017). This strain contains the P*pgaX-amdS* reporter construct and is defective in carbon catabolite repression due to disruption of the *creA* gene (*ΔcreA*). P*pgaX-amdS ΔcreA* (JN29.2) is able to grow on acetamide as the sole nitrogen source only when the GA-responsive promoter of the *pgaX* (NRRL3_03144) gene gets activated to express the *amdS* gene encoding the acetamidase enzyme (Niu *et al.*, 2015). In total, 65 mutants showing constitutive expression of pectinases were isolated on plates containing MM with glucose as the carbon source and acetamide as the sole nitrogen source. Eleven of these mutants were analysed for mutations in the GA-responsive transcription factor *gaaR*

(NRRL3_08195) but no mutations were found. The genomes of five mutants were sequenced and it was found that they are all mutated in a common gene, which was named the repressor of D-galacturonic acid utilization, *gaaX* (NRRL3_08194) (Niu *et al*., 2017). It was also shown that deletion of *gaaX* results in constitutive expression of pectinases (Niu *et al*., 2017).

We sequenced the *gaaX* gene in the other 59 mutants showing constitutive expression of pectinases (Table S3, Figure S1). In total, 28 and nine mutants were found to carry nonsense and frameshift mutations, respectively. The nonsense or frameshift mutation closest to the C-terminus was the frameshift mutation V653G, indicating the importance of the last 45 aa residues for a functional GaaX. Twenty-three mutants were found to have missense mutations in the *gaaX* gene. These mutations were found throughout *gaaX*, indicating the importance of the different domains of GaaX for proper functioning. Four mutants did not carry any mutations in *gaaX*. Sequencing the *gaaR* gene in these mutants resulted in the identification of a mutation in *gaaR* in one of the mutants.

The UV mutant P*pgaX-amdS ΔcreA gaaR*$^{W361R\text{-UV}}$ (JN103.1) carries a missense mutation (W 361 to R) caused by codon change from TGG to CGG at a distance of 1285 bp from the start codon of *gaaR*. The tryptophan361 residue lies in the fungal-specific transcription factor domain that spans the residues 139-518 in the 740 aa long GaaR (Figure 1A), and is 100 % conserved in the homologous GaaR sequences present in *Aspergillus* species (Figure 1B) and in other Ascomycetes analysed belonging to the Pezizomycotina subdivision (Figure 1C). Growth of the P*pgaX-amdS ΔcreA gaaR*$^{W361R\text{-UV}}$ (JN103.1) strain was similar to the growth of the *ΔcreA* (XY1.1) strain on MM containing glucose, D-fructose, GA, PGA or AP (Figure S2), both displaying a smaller radial growth compared to the reference (MA234.1) strain due to *creA* deletion.

Analyses of gene expression in *gaaR*^{W361R}

The UV mutant P*pgaX-amdS ΔcreA gaaR*$^{W361R\text{-UV}}$ (JN103.1) possibly carries multiple mutations in its genome. To create a strain that carries only the W361R mutation in *gaaR*, the *gaaR*^{W361R} gene was transformed into a *ΔgaaR* (SO1.1) strain. Southern blot analysis indicated that the P*gaaR-gaaR*^{W361R}-T*gaaR* construct was successfully integrated into the endogenous *gaaR* locus in the genome of the *gaaR*^{W361R} (SO2.1) strain (Figure S3). Growth of *gaaR*^{W361R} (SO2.1) was analysed on different monomeric and polymeric carbon sources (Figure S2). *ΔgaaR* (SO1.1) displayed a strongly impaired growth on GA, PGA and AP, as it was shown previously (Alazi *et al.*, 2016). Introduction of *gaaR*^{W361R} gene resulted in a full complementation of growth on GA containing carbon sources, indicating the presence of a functional GaaR. Growth of *gaaR*^{W361R} (SO2.1) was similar to the growth of the reference (MA234.1) strain on all carbon sourcestested.

The pectinase production capacity of *gaaR*^{W361R} (EA34.1), a uridine prototroph derived from *gaaR*^{W361R} (SO2.1), was assessed via PGA plate assay (Figure 2A). Strains were grown in liquid medium containing D-fructose, a carbon source that does not induce the expression of GA-responsive genes, and the culture supernatants were spotted on PGA plates. Comparison of hydrolysis zones on PGA plates showed that *gaaR*^{W361R} (EA34.1) constitutively produces pectinases involved in PGA degradation, as the UV mutant P*pgaX-*

amdS ΔcreA gaaR$^{W361R\text{-UV}}$ (JN103.1) does (Figure 2A). This indicates that the W361R mutation in *gaaR*, and no other additional mutation(s) induced by UV mutagenesis in P*pgaX-amdS ΔcreA gaaR*$^{W361R\text{-UV}}$ (JN103.1), is responsible for the constitutive expression of the genes encoding pectinases. The reference (MA234.1) strain did not display any polygalacturonase activity in its supernatant under non-inducing conditions, while the P*pgaX-amdS ΔgaaX* (JN123.1) strain did, as it was shown previously (Niu *et al.*, 2017).

Figure 1. Schematic representation of the domains present in GaaR **(A)** and conservation of GaaR^{W361} in *Aspergillus* species **(B)** and other Ascomycetes **(C)**. Domains in GaaR (NRRL3_08195) were identified using NCBI's CDD (Marchler-Bauer *et al.*, 2015). Protein sequences homologous to the *A. niger* GaaR were retrieved using the blastp algorithm from NCBI (Altschul *et al.*, 1990) against the nonredundant protein sequences database, and were aligned using Clustal Omega (Sievers *et al.*, 2011).

The effect of the W361R mutation in *gaaR* on the expression of pectinase genes was also analysed using the promoter-reporter strains expressing the *amdS* gene via the *pgaX*

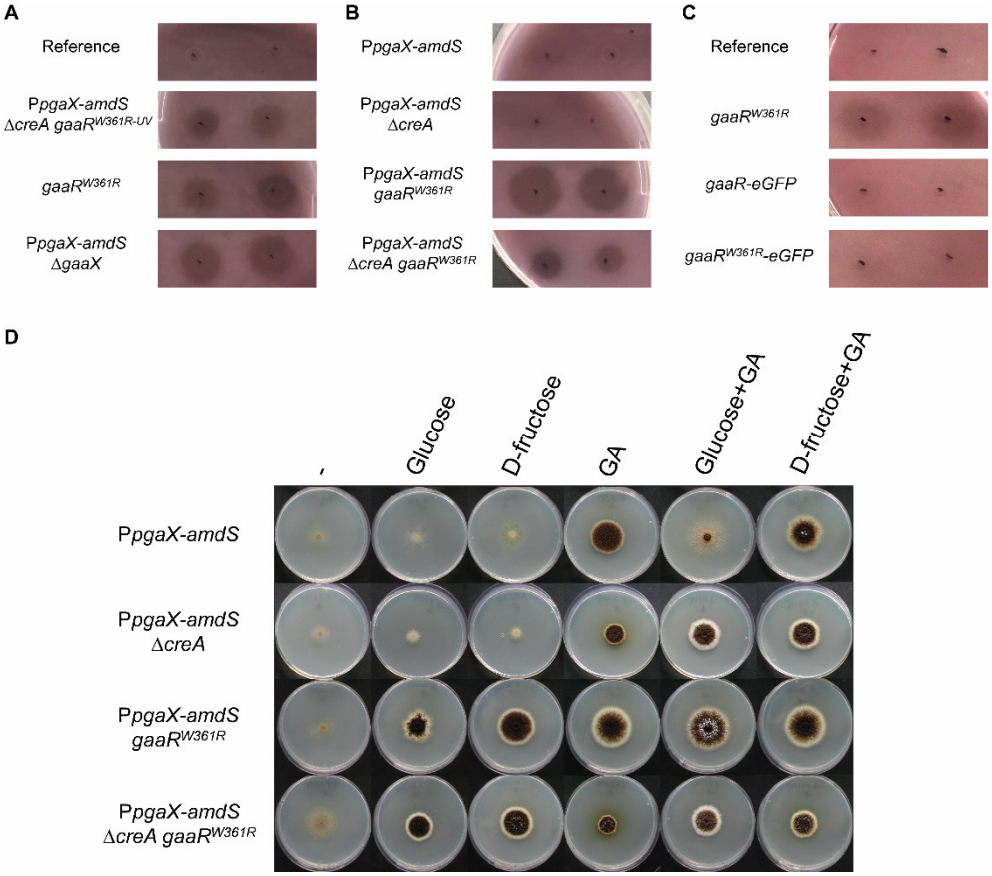

Figure 2. Effect of the W361R mutation in *gaaR* on the expression of pectinase genes. **A-C** PGA plate assay. Strains were grown in MM containing 50 mM D-fructose for 36 h, and culture supernatants were spotted PGA plates. Enzymatic activities in the supernatants from duplicate cultures are shown: **(A)** the reference (MA234.1), P*pgaX-amdS ΔcreA gaaR*$^{W361R\text{-}UV}$ (JN103.1), *gaaR*^{W361R} (EA34.1), P*pgaX-amdS ΔgaaX* (JN123.1); **(B)** P*pgaX-amdS* (JC1.5), P*pgaX-amdS ΔcreA* (JN29.2), P*pgaX-amdS gaaR*^{W361R} (JN130.4), P*pgaX-amdS ΔcreA gaaR*^{W361R} (JN129.1); and **(C)** reference (MA234.1), *gaaR*^{W361R} (EA34.1), *gaaR-eGFP* (EA36.1) and *gaaR*^{W361R}*-eGFP* (EA39.1). **D** Growth phenotype of the P*pgaX-amdS* (JC1.5), P*pgaX-amdS ΔcreA* (JN29.2), P*pgaX-amdS gaaR*^{W361R} (JN130.4) and P*pgaX-amdS ΔcreA gaaR*^{W361R} (JN129.1) strains after 7 days at 30 °C on solid MM containing no carbon source (-); 50 mM glucose, D-fructose or GA as the sole carbon source; or 50 mM glucose or D-fructose together with 50 mM GA. All plates contain 10 mM acetamide as the sole nitrogen source.

promoter. The *gaaR*^{W361R} gene was transformed into the P*pgaX-amdS* (JC1.5) and P*pgaX-amdS ΔcreA* (JN29.2) strains. Transformants were selected on medium containing a non-inducing carbon source and acetamide as the nitrogen source, and only transformants in which the *gaaR*^{W361R} was integrated were expected grow on the transformation plate because of constitutive activation of the P*pgaX-amdS* reporter. Southern blot analysis and sequencing of the *gaaR* locus showed that the endogenous *gaaR* gene was replaced with *gaaR*^{W361R} in the resulting strains P*pgaX-amdS gaaR*^{W361R} (JN130.4) and P*pgaX-amdS ΔcreA gaaR*^{W361R} (JN129.1), respectively (Figure S3). Growth of P*pgaX-amdS gaaR*^{W361R} (JN130.4)

and P*pgaX-amdS ΔcreA gaaR*^{W361R} (JN129.1) was similar to the growth of their parental strains on MM containing glucose, D-fructose, GA, PGA or AP (Figure S2).

The PGA plate assay showed that the parental strains P*pgaX-amdS* (JC1.5) and P*pgaX-amdS ΔcreA* (JN29.2) did not produce any polygalacturonases in D-fructose, while P*pgaX-amdS gaaR*^{W361R} (JN130.4) and P*pgaX-amdS ΔcreA gaaR*^{W361R} (JN129.1) showed polygalacturonase activity in their culture supernatants (Figure 2B). The expression of *pgaX* was previously shown to be induced only in the presence of an inducing carbon source, such as GA (Niu *et al.*, 2015). Radial growth assays on plates containing acetamide showed that P*pgaX-amdS gaaR*^{W361R} (JN130.4) and P*pgaX-amdS ΔcreA gaaR*^{W361R} (JN129.1) can grow on glucose or D-fructose, while their parental strains cannot (Figure 2D). This indicates that GaaR^{W361R} can activate the expression of *amdS* via the *pgaX* promoter in an inducer-independent way, confirming the results obtained by PGA plate assays.

The expression of *pgaX* is known to be repressed strongly on glucose and mildly on D-fructose via CreA (Niu *et al.*, 2015; Alazi *et al.*, 2018). As shown in Figure 2D, growth of the promoter-reporter strains on plates containing acetamide and GA decreased when glucose or D-fructose was added to the growth media, and this growth reduction was restored when *creA* is deleted. Although P*pgaX-amdS gaaR*^{W361R} (JN130.4) could grow both on glucose and D-fructose as the sole carbon source, its growth on glucose was reduced compared to on D-fructose (Figure 2D). This indicates that CreA-mediated repression on *pgaX* plays an important role even in the presence of the constitutively active GaaR^{W361R}.

Subcellular localization of GaaR-eGFP and GaaR^{W361R}-eGFP

We have previously shown that *eGFP-gaaR* (EA19.2), a strain expressing the N-terminally eGFP-tagged *gaaR*, shows a slightly reduced growth on GA compared to the reference strain (Alazi *et al.*, 2018), indicating that N-terminal GFP-tagging might result in a slight decrease in GaaR activity. To check whether C-terminal GFP-tagging affects GaaR activity, strains expressing C-terminally eGFP-tagged *gaaR*, with or without a linker between GaaR and eGFP, were created in a *ΔgaaR* background. While *ΔgaaR* (JN35.1) could not grow on GA, growth on GA was fully complemented in both *gaaR-eGFP* (EA29.14) and *gaaR-(GA)*$_4$*-eGFP* (EA30.6) (Figure S4). This indicates that C-terminal eGFP-tagging does not affect the transcriptional activity of GaaR, independently of the presence of a linker between GaaR and eGFP.

To analyse the subcellular localization of GaaR-eGFP and GaaR^{W361}-eGFP, strains that express C-terminally eGFP-tagged *gaaR* or *gaaR*^{W361R} from the endogenous *gaaR* locus were created, and integration of the P*gaaR-gaaR-eGFP*-T*gaaR* or P*gaaR-gaaR*^{W361R}*-eGFP*-T*gaaR* constructs into the endogenous *gaaR* locus was verified by Southern blot analysis (Figure S3). Growth of *gaaR-eGFP* (EA31.1) and *gaaR*^{W361R}*-eGFP* (EA32.1) was fully complemented on GA and GA-containing carbon sources (Figure S2), confirming previous results that GaaR-eGFP is fully functional under inducing conditions.

The ability of *gaaR-eGFP* (EA36.1) and *gaaR*^{W361R}*-eGFP* (EA39.1), uridine prototrophs derived from *gaaR-eGFP* (EA31.1) and *gaaR*^{W361R}*-eGFP* (EA32.1), respectively, to constitutively express pectinases was assessed via a PGA plate assay and compared to that

of *gaaR*^{W361R} (EA34.1) (Figure 2C). No polygalacturonase activity was observed in the culture supernatant of the reference (MA234.1) strain or *gaaR-eGFP* (EA36.1) after growth in D-fructose, indicating that C-terminal eGFP-tagging does not affect the inactive state of GaaR under non-inducing conditions, similar to N-terminal eGFP-tagging (Alazi *et al.*, 2018). While the *gaaR*^{W361R} (EA34.1) strain expressing the constitutively active *gaaR*^{W361R} produced pectinases in D-fructose, the *gaaR*^{W361R}*-eGFP* (EA39.1) strain expressing *gaaR*^{W361R}*-eGFP* did not (Figure 2C). This shows that, unlike GaaR^{W361R}, GaaR^{W361R}-eGFP did not activate the expression of pectinases under non-inducing conditions.

The subcellular localization of the C-terminally eGFP-tagged GaaR and GaaR^{W361R} in *gaaR-eGFP* (EA36.1) and *gaaR*^{W361R}*-eGFP* (EA39.1) grown under inducing (i.e. on 2-keto-3-deoxy-L-galactonate) or non-inducing (i.e. on D-fructose) conditions was analysed using confocal laser scanning microscopy (Figure 3). Beforehand, the potency of 2-keto-3-deoxy-L-galactonate, the physiological inducer (Alazi *et al.*, 2017), to induce the expression of the GA-responsive genes was tested using the promoter-reporter strain P*pgaX-amdS ΔcreA* (JN29.2) and was compared to that of GA (Figure S5). P*pgaX-amdS ΔcreA* (JN29.2) showed an impaired growth on plates containing acetamide and 50 mM glucose. Addition of 100 nM 2-keto-3-deoxy-L-galactonate to the growth media resulted in a major increase in growth, whereas addition of a 100 nM GA was not sufficient to induce the growth of P*pgaX-amdS ΔcreA* (JN29.2), indicating that 2-keto-3-deoxy-L-galactonate is more potent than GA in inducing the expression of pectinases.

We previously showed that the partially functional N-terminally eGFP-tagged GaaR in *eGFP-gaaR* (EA19.2) is localized mainly in the nucleus under both inducing (i.e. on GA) and non-inducing (i.e. on D-fructose) conditions (Alazi *et al.*, 2018). Here, we show again that eGFP-GaaR is mainly localized in the nucleus after growth in 2-keto-3-deoxy-L-galactonate or D-fructose (Figure 3). C-terminally eGFP-tagged GaaR and GaaR^{W361R} were also localized mainly in the nucleus after growth in 2-keto-3-deoxy-L-galactonate, while apparent nuclear fluorescence was not observed after growth in D-fructose. This result indicates that on D-fructose (under non-inducing and mildly repressing conditions) nuclear accumulation of C-terminally eGFP-tagged GaaR is restrained, suggesting that the inability of GaaR^{W361R}-eGFP to activate the expression of pectinases under non-inducing conditions might be caused by improper subcellular localization to execute its function.

DISCUSSION

We previously isolated via a forward genetic screen *A. niger* strains displaying constitutive expression of genes encoding pectinases (Niu *et al*., 2017). The majority of the mutant collection (60 strains) was found to contain a loss-of-function mutation in *gaaX*, the repressor of GA utilization. One mutant with no mutations in *gaaX* was found to possess the missense mutation W361R in *gaaR*. Introducing the W361R mutation in the wild type genetic background and subsequent gene expression analyses allowed us to conclude that this mutation results in constitutive activation of GaaR and therefore constitutive expression of pectinase genes.

To our knowledge, XlnR was the only transcription factor involved in plant biomass degradation in which missense mutations leading to constitutive expression of plant

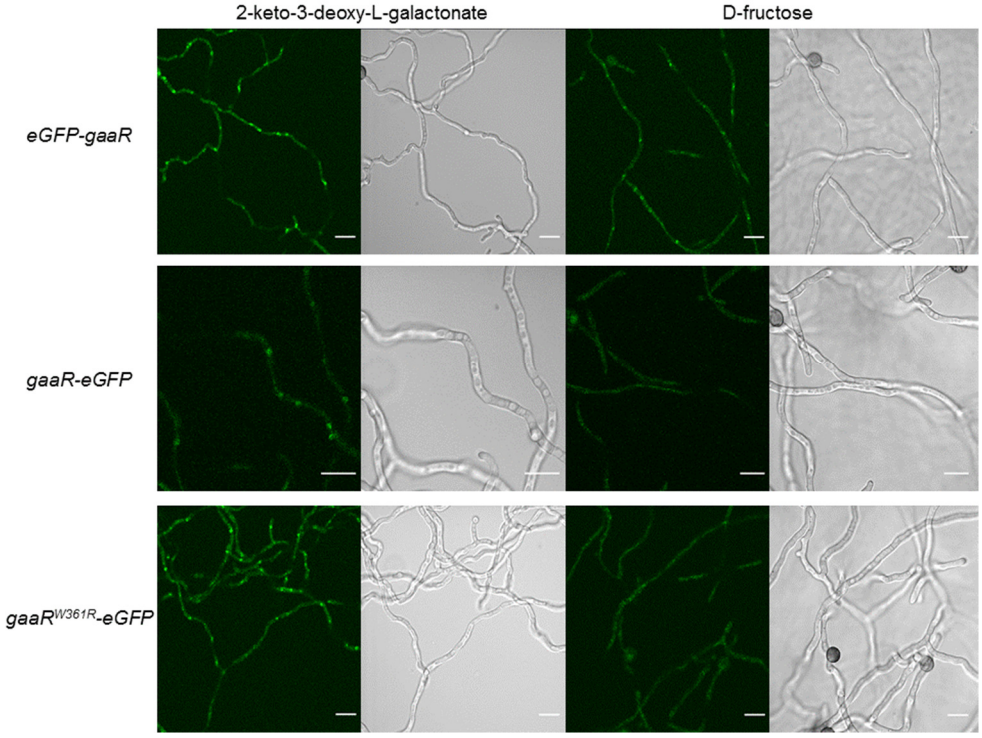

Figure 3. Nuclear and cytoplasmic fluorescence intensities of the eGFP-tagged GaaR and GaaR^{W361R} proteins. The *eGFP-gaaR* (EA19.2), *gaaR-eGFP* (EA36.1) and *gaaR^{W361R}-eGFP* (EA39.1) strains were grown in MM containing 10 mM D-fructose or 2-keto-3-deoxy-L-galactonate for 21 h. Example micrographs representing each condition are shown. Scale bar: 10 µm.

biomass degrading enzymes were reported, such as V756F in *A. niger*, or A824V in *Trichoderma reesei* and orthologous mutations in *Neurospora crassa* and *Penicillium oxalicum* (Hasper, 2004; Hasper *et al.*, 2004; Derntl *et al.*, 2013; Craig *et al.*, 2015; Gao *et al.*, 2017). Identification of the constitutively active GaaR^{W361R} would benefit the industrial production of pectinases in an inducer-independent way.

The activity of GaaR is proposed to be inhibited by GaaX under non-inducing conditions via protein-protein interaction (Niu et al., 2017). Pectinase genes are expressed under inducing conditions (i.e. in the presence of inducer), or under non-inducing conditions (i.e. in the absence of inducer) when *gaaX* is deleted or *gaaR* is overexpressed (Martens-Uzunova and Schaap, 2008; Niu *et al.*, 2017; Alazi *et al.*, 2018). These findings indicate that GA-responsive gene expression only requires the presence of active GaaR relieved/escaped from GaaX inhibition, and does not require the presence of the inducer. This suggests that under inducing conditions, the inducer binds to GaaX resulting in a GaaX-unbound and active form of GaaR. Therefore, we propose that W361R mutation in GaaR disrupts the interaction between GaaX and GaaR under non-inducing conditions. Loss-of-function mutations were found throughout all three domains of *gaaX* (shikimate kinase, 3-dehydroquinate dehydratase and shikimate dehydrogenase) (Figure S1), making it difficult to predict the domain that possibly interacts with GaaR.

N-terminally eGFP-tagged GaaR was previously shown to mainly localize in the nucleus under both inducing and non-inducing conditions, but was unable to fully complement the growth of *ΔgaaR* when grown on inducing carbon sources (Alazi *et al.*, 2018). C-terminally eGFP-tagged GaaR/GaaR^{W361R} accumulated in the nucleus under inducing conditions and complemented the growth of *ΔgaaR* when grown on inducing carbon sources. Based on that we conclude that C-terminally eGFP-tagged GaaR/GaaR^{W361R} is fully functional under inducing conditions. However, unlike the N-terminally eGFP-tagged GaaR, strains expressing the C-terminally eGFP-tagged GaaR/GaaR^{W361R} did not show higher nuclear fluorescence intensity compared to cytoplasmic fluorescence intensity under non-inducing conditions. This result indicates that C-terminal eGFP-tagging restrained nuclear accumulation of GaaR/GaaR^{W361R}, confirming the inability of GaaR^{W361R}-eGFP to constitutively activate the expression of pectinases, as GaaR^{W361R} can.

Although the probability of identifying gain-of-function mutations leading to constitutive activation of GaaR was very low compared to that of loss-of-function of GaaX, constitutively active forms of GaaR can be utilized in protein/metabolite interaction studies and provide insight in regulation of GA-responsive gene expression. The nature of mutations in three UV mutants with no mutations in *gaaX* or *gaaR* remains to be discovered.

ACKNOWLEDGEMENTS

EA and JN were supported by grants from BE-Basic (Flagship 10) and the China Scholarship Council, respectively. This work was in part supported by the Natural Sciences and Engineering Research Council of Canada Strategic Industrial Biocatalysis Network.

REFERENCES

Alazi E, Khosravi C, Homan TG, du Pré S, Arentshorst M, Di Falco M, Pham TTM, Peng M, Aguilar-Pontes MV, Visser J, Tsang A, de Vries RP, Ram AFJ (2017) The pathway intermediate 2-keto-3-deoxy-L-galactonate mediates the induction of genes involved in D-galacturonic acid utilization in *Aspergillus niger*. *FEBS Letters* 591:1408-1418

Alazi E, Knetsch T, Di Falco M, Reid ID, Arentshorst M, Visser J, Tsang A, Ram AFJ (2018) Inducer-independent production of pectinases in *Aspergillus niger* by overexpression of the D-galacturonic acid responsive transcription factor *gaaR*. *Applied Microbiology and Biotechnology* 102:2723-2736

Alazi E, Niu J, Kowalczyk JE, Peng M, Aguilar Pontes MV, van Kan JAL, Visser J, de Vries RP, Ram AFJ (2016) The transcriptional activator GaaR of *Aspergillus niger* is required for release and utilization of D-galacturonic acid from pectin. *FEBS Letters* 590:1804-1815

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology* 215:403-410

Arentshorst M, Niu J, Ram AFJ (2015) Efficient Generation of *Aspergillus niger* Knock Out Strains by Combining NHEJ Mutants and a Split Marker Approach. In van den Berg MA, Maruthachalam K (ed), Genetic Transformation Systems in Fungi, Vol. 1, Springer International Publishing, pp 263-272 Arentshorst M, Ram AFJ, Meyer V (2012) Using Non-homologous End-Joining-Deficient Strains for Functional Gene Analyses in Filamentous Fungi. *Plant Fungal Pathogens Methods in Molecular*

Biology 835:133-150

Caffall KH, Mohnen D (2009) The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydrate Research* 344(14)

Carvalho NDSP, Arentshorst M, Kwon MJ, Meyer V, Ram AFJ (2010) Expanding the *ku70* toolbox for filamentous fungi: establishment of complementation vectors and recipient strains for advanced gene analyses. *Applied Microbiology and Biotechnology* 87:1463-1473

Craig JP, Coradetti ST, Starr TL, Glass NL (2015) Direct target network of the *Neurospora crassa* plant cell wall deconstruction regulators CLR-1, CLR-2, and XLR-1. *mBio* 6(5):1–11

de Ruiter-Jacobs YMJT, Broekhuijsen M, Unkles SE, Campbell EI, Kinghorn JR, Contreras R, Pouwels PH, van den Hondel CAMJJ (1989) A gene transfer system based on the homologous *pyrG* gene and efficient expression of bacterial genes in *Aspergillus oryzae*. *Current Genetics* 16:159-163

Derntl C, Gudynaite-Savitch L, Calixte S, White T, Mach RL, Mach-Aigner AR (2013) Mutation of the Xylanase regulator 1 causes a glucose blind hydrolase expressing phenotype in industrially used *Trichoderma* strains. Biotechnol. *Biofuels* 6:1–11

DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM (2013) Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Research* 41(7):4336-4343

Edwards MC, Doran-Peterson J (2012) Pectin-rich biomass as feedstock for fuel ethanol production. *Applied Microbiology and Biotechnology* 95:565–575

Gao L, Li Z, Xia C, Qu Y, Liu M, Yang P et al. (2017) Combining manipulation of transcription factors and overexpression of the target genes to enhance lignocellulolytic enzyme production in *Penicillium oxalicum*. Biotechnol. *Biofuels* 10:1-16

Hasper AA (2004) Function and mode of regulation of the transcriptional activator XlnR from *Aspergillus*. [dissertation]. [Wageningen (The Netherlands)]: Wageningen University

Hasper AA, Trindade LM, van der Veen D, van Ooyen AJJ, de Graaff LH (2004) Functional analysis of the transcriptional activator XlnR from *Aspergillus niger*. *Microbiology* 150:1367-1375

Heigwer F, Kerr G, Boutros M (2014) E-CRISP: fast CRISPR target site identification. *Nature Methods* 11:122-123

Kashyap D, Vohra P, Chopra S, Tewari R (2001) Applications of pectinases in the commercial sector: a review. *Bioresource Technology* 77:215–227

Khan M, Nakkeeran E, Umesh-Kumar S (2013) Potential application of pectinase in developing functional foods. *Annual Review of Food Science and Technology* 4:21–34

Kowalczyk JE, Benoit I, de Vries RP (2014) Regulation of plant biomass utilization in *Aspergillus*. *Advances in Applied Microbiology* 88: 31-56

Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Research* 43:D222-226 Martens-Uzunova ES, Schaap PJ (2008) An evolutionary conserved D-galacturonic acid metabolic pathway operates across filamentous fungi capable of pectin degradation. *Fungal Genetics and Biology* 45(11):1449-1457

Niu J, Alazi E, Reid ID, Arentshorst M, Punt PJ, Visser J, Tsang A, Ram AFJ (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, Homan TG, Arentshorst M, de Vries RP, Visser J, Ram AF (2015) The interaction of induction and repression mechanisms in the regulation of galacturonic acid-induced genes in *Aspergillus niger*. *Fungal Genetics and Biology* 82:32-42

Nødvig CS, Nielsen JB, Kogle ME, Mortensen UH (2015) A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi. *PLoS One* 10(7):e0133085

Papadopoulos JS, Agarwala R (2007) COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinformatics* 23:1073-1079

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* 7:539

Toushik SH, Lee K, Lee J, Kim K (2017) Functional applications of lignocellulolytic enzymes in the fruit and vegetable processing industries. *Journal of Food Science* 88(3):583–595

van Hartingsveldt W, Mattern IE, van Zeijl CM, Pouwels PH, van den Hondel CA (1987) Development of a homologous transformation system for *Aspergillus niger* based on the *pyrG* gene. *Molecular and General Genetics* 206(1):71-5

Yuan XL, Goosen C, Kools H, van den Maarel MJEC, van den Hondel CAMJJ, Dijkhuizen L, Ram AFJ (2006) Database mining and transcriptional analysis of genes encoding inulin-modifying enzymes of *Aspergillus niger*. *Microbiology* 152:3061-3073

SUPPORTING INFORMATION

Table S1. Strains used in this study.

Strain	Accession Number	Genotype	Description	Reference
MA169.4	FGSC A1279	*cspA1, pyrG^{378}, kusA::DR-amdS-DR*	Transiently silenced *kusA* in AB4.1	Carvalho *et al.*, 2010
MA234.1	CBS 141255	*cspA1, kusA::DR-amdS-DR*	Restored *pyrG* in MA169.4	Alazi *et al.*, 2016
XY1.1	CBS 143276	*cspA1, pyrG^{378}, creA::pyrG*	*ΔcreA* in AB4.1	Yuan *et al.*, 2006
JC1.5	CBS 143271	*cspA1, Δku70,* P*pgaX-amdS* integrated to the *pyrG* locus	P*pgaX-amdS* in MA299.2	Niu *et al.*, 2015
JN29.2	CBS 143272	*cspA1, Δku70,* P*pgaX-amdS* integrated to the *pyrG* locus*, creA::hygB*	*ΔcreA* in JC1.5	Niu *et al.*, 2015
JN123.1	CBS 143275	*cspA1, Δku70,* P*pgaX-amdS* integrated to the *pyrG* locus*, gaaX::hygB*	*ΔgaaX* in JC1.5	Niu *et al.*, 2017
JN130.4		*cspA1, Δku70,* P*pgaX-amdS* integrated to the *pyrG* locus, *gaaR^{W361R}* integrated to the *gaaR* locus	*gaaR^{W361R}* in JC1.5	This study
JN103.1		*cspA1, Δku70,* P*pgaX-amdS* integrated to the *pyrG* locus*, creA::hygB, UV mutagenized yielding to gaaR^{W361R}*	*gaaR$^{W361R\text{-}UV}$* in JN29.2	This study
JN129.1		*cspA1, Δku70,* P*pgaX-amdS* integrated to the *pyrG* locus*, creA::hygB,* *gaaR^{W361R}* integrated to the *gaaR* locus	*gaaR^{W361R}* in JN29.2	This study
SO1.1		*cspA1, pyrG^{378}, kusA::DR-amdS-DR, gaaR::AOpyrG*	*ΔgaaR* in MA169.4	This study
SO2.1		*cspA1, pyrG^{378}, kusA::DR-amdS-DR, gaaR^{W361R}* integrated to the *gaaR* locus	*gaaR^{W361R}* in SO1.1	This study
EA34.1		*cspA1, pyrG^{378}, kusA::DR-amdS-DR, gaaR^{W361R}* integrated to the *gaaR* locus, ectopically integrated pAB4.1	Restored *pyrG* in SO2.1	This study
EA31.1		*cspA1, pyrG^{378}, kusA::DR-amdS-DR, gaaR-eGFP* integrated to the *gaaR* locus	*gaaR-eGFP* in SO1.1	This study
EA36.1		*cspA1, pyrG^{378}, kusA::DR-amdS-DR, gaaR-eGFP* integrated to the *gaaR* locus, ectopically integrated pAB4.1	Restored *pyrG* in EA31.1	This study
EA32.1		*cspA1, pyrG^{378}, kusA::DR-amdS-DR, gaaR^{W361R}-eGFP* integrated to the *gaaR* locus	*gaaR^{W361R}-eGFP* in SO1.1	This study
EA39.1		*cspA1, pyrG^{378}, kusA::DR-amdS-DR, gaaR^{W361R}-eGFP* integrated to the *gaaR* locus, ectopically integrated pAB4.1	Restored *pyrG* in EA32.1	This study
JN35.1	CBS 141256	*cspA1, kusA::DR-amdS-DR, gaaR::hygB*	*ΔgaaR* in MA234.1	Alazi *et al.*, 2016
JN36.1	CBS 143259	*cspA1, gaaR::hygB*	*amdS* loopout in JN35.1	Alazi *et al.*, 2016

Table S1. Strains used in this study. (Continued)

Strain	Accession Number	Genotype	Description	Reference
EA19.2	CBS 143267	*cspA1, gaaR::hygB, amdS⁺, ectopically integrated* P*gaaR-eGFP-gaaR*	*eGFP-gaaR* in JN36.1	Alazi *et al.*, 2018
EA29.14		*cspA1, gaaR::hygB, amdS⁺, ectopically integrated* P*gaaR-gaaR-eGFP*	*gaaR-eGFP* in JN36.1	This study
EA30.6		*cspA1, gaaR::hygB, amdS⁺, ectopically integrated* P*gaaR-gaaR-*$(GA)_4$*-eGFP*	*gaaR-*$(GA)_4$*-eGFP* in JN36.1	This study

Table S2. Primers used in this study. Fusion PCR overlapping regions are written in bold.

Primer name	Sequence (5' to 3')	Used for
gaaX_P5f	CGTTGGTTCGATGTAAATGGG	Amplification and sequencing of *gaaX*
gaaX_GSP10r	TGTATGTGGCATTGATGGAACC	Sequencing of *gaaX*
gaaX_GSP11f	GGGTTGTTTGTCGGTGTACTCG	Sequencing of *gaaX*
gaaX_GSP12r	CGGTTCTTGGTCTTCAGGTCCT	Sequencing of *gaaX*
gaaX_GSP13f	CCTGTACTCTTTGCGCAGTCCT	Sequencing of *gaaX*
gaaX_GSP14r	GCTGTCCCATGATGTGGAGG	Sequencing of *gaaX*
gaaX_GSP15f	TTTCGCCCATCACGCATC	Sequencing of *gaaX*
gaaX_GSP16r	CCATGCATCGACATCTTCAATC	Sequencing of *gaaX*
gaaX_GSP17f	AGTGTACCTGGACTTGACGC	Sequencing of *gaaX*
gaaX_GSP9r	CAAACGTGGAGAAGCCGGT	Amplification and sequencing of *gaaX*
gaaRP7f	TCCTCTTCGGCTTCTGCTTC	Amplification and sequencing of *gaaR*
gaaRP9r	GCATACTCCAGGCTCCCTTG	Sequencing *gaaR*
gaaRP10f	CCTACATGGGAAGCCCTATTGG	Sequencing *gaaR*
gaaRP11r	CGATAGAGCAGACAGTGCACGA	Sequencing *gaaR*
gaaRP12f	CCCTGTTGGTCGACATTACGC	Sequencing *gaaR*
gaaRP13r	AAATCGGGTCGGTCCAGTG	Sequencing *gaaR*
gaaRP14f	AGGACTTCCTGGACCGGTAT	Sequencing *gaaR*
gaaRP8r	CATGTGATCATTTCGTGGCCT	Amplification and sequencing of *gaaR*
gaaRP5f	TGGACGGCGTATGGGATT	Amplification of P*gaaR-gaaR*-T*gaaR*
gaaRP6r	CCCAAAGACGGGTGGTAGAGT	Amplification of P*gaaR-gaaR*-T*gaaR*
gaaR SBfor	CCTCGACGCCATTCCAGTT	Amplification of Southern blot probe

Table S2. Primers used in this study. Fusion PCR overlapping regions are written in bold. (Continued)

Primer name	Sequence (5' to 3')	Used for
gaaR SBrev	GGTCATGGACACCGCATTG	Amplification of Southern blot probe
GaaR_GFP4F	**GCATGGACGAGCTGTACAAG**TAAGAAGCCGAATATACGAGCCAT	Amplification of T*gaaR*
GaaR_GFP3R	GCGTGCATCAAGGCGATT	Amplification of T*gaaR*
GaaR_GFP1F	**ATGGTGAGCAAGGGCGAG**	Amplification of *eGFP*
GaaR_GFP1R	**CTTGTACAGCTCGTCCATG**	Amplification of *eGFP*
GaaR_GFP2F	TGGTGTGATGAAGGCAGTGTG	Amplification of P*gaaR-gaaR(-(GA)$_4$)*
GaaR_GFP5R	**TCCTCGCCCTTGCTCACCA**TAGGATTCTCCACCTCCACC	Amplification of P*gaaR-gaaR*
eGFP-gaaR-For	TGGTGTGATGAAGGCAGTGTG	Fusion PCR to amplify P*gaaR-gaaR(-(GA)$_4$)-eGFP*-T*gaaR*
eGFP-gaaR-Rev	TGGGTGTTTGGTGGGTTGA	Fusion PCR to amplify P*gaaR-gaaR(-(GA)$_4$)-eGFP*-T*gaaR*
GaaR_GFP6R	**TCCTCGCCCTTGCTCACCA**TAGCGCCAGCGCCAGCGCCAGCGCCAGGATTCT CCACCTCCACC	Amplification of P*gaaR-gaaR-(GA)$_4$*
gaaRP1f	GCTGCTGCTGCTGCTTTACA	Amplification of *gaaR* 5' flank to create SO1.1
gaaRP2r	**CAATTCCAGCAGCGGCTT**GGCATTGCCTGTGCATAGGA	Amplification of *gaaR* 5' flank to create SO1.1
gaaRP3f	**ACACGGCACAATTATCCATCG**GAAGCCGAATATACGAGCCA	Amplification of *gaaR* 3' flank to create SO1.1
gaaRP4r	GCGTGCATCAAGGCGATT	Amplification of *gaaR* 3' flank to create SO1.1
AOpyrGP12f	**AAGCCGCTGCTGGAATTG**	Amplification of *AOpyrG* 5' flank to create SO1.1
AOpyrGP15r	CCGGTAGCCAAAGATCCCTT	Amplification of *AOpyrG* 5' flank to create SO1.1
AOpyrGP13r	**CGATGGATAATTGTGCCGTGT**	Amplification of *AOpyrG* 3' flank to create SO1.1
AOpyrGP14f	ATTGACCTACAGCGCACGC	Amplification of *AOpyrG* 3' flank to create SO1.1
sgRNAP1	TAATACGACTCACTATAGGGTGGATACGTACTCCTTTTAGTTTTAGAGCTAG AAATAGCAA	Amplification of the DNA template for guide RNA
OTL19	GCACCACCGACTCGGTGC	Amplification of the DNA template for guide RNA
Fw_LIC2	CAACCTCCAATCCAATTTGACTCCGCCGAACGTACTG	Amplification of the DNA template for guide RNA
Rev_P1	**TAAAAGGAGTACGTATCCAC**GACGAGCTTACTCGTTTCG	Amplification of the DNA template for guide RNA
Fw_P1	**GTGGATACGTACTCCTTTTAGTTTTA**GAGCTAGAAATAGCAAG	Amplification of the DNA template for guide RNA
Rev_LIC2	ACTACTCTACCACTATTTGAAAAGCAAAAAAGGAAGGTACAAAAAAGC	Amplification of the DNA template for guide RNA
for_pTE1	CCTTAATTAAACTCCGCCGAACGTACTG	Introduction of *Pac*I site to the DNA template for guide RNA
rev_pTE1	CCTTAATTAAAAAAGCAAAAAAGGAAGGTACAAAAAAGC	Introduction of *Pac*I site to the DNA template for guide RNA
PgaaR_seq_P2	CATCCATCTTAGTCCACCGGC	Amplification of *gaaR*
TgaaR_seq_P2	TCTAGTATTCACTTGTCCCTAA	Amplification of *gaaR*

Table S3. Mutations in *gaaX* in the UV mutants showing constitutive expression of pectinases. aa and gDNA represents amino acid and genomic DNA, respectively.

Strain	Type of mutation	Mutation	Codon change	aa change	Predicted length of the mutated protein (aa)	Remarks
S1	Frameshift	Insertion of G	GTT-GGT	V653G	664	gDNA sequenced, no mutation found in *gaaR* (Niu *et al.*, 2017)
S3	Missense	T-C	CTG-CCG	L111P		
S5	Nonsense	C-T	CAA-TAA	Q369Stop	368	gDNA sequenced, no mutation found in *gaaR* (Niu *et al.*, 2017)
S6						No mutations found in *gaaX* or *gaaR*
S7	Missense	T-C	CTG-CCG	L676P		gDNA sequenced, no mutation found in *gaaR* (Niu *et al.*, 2017)
S8	Frameshift	Deletion of TG	GTG-GGG	V193-G	197	
S9	Nonsense	C-T	CGA-TGA	R425Stop	424	
S10	Frameshift	Deletion of G	GGA-GAG	G644E	702	
S11	Frameshift	Insertion G	GGAGCC-GGGAGC	GA469GS	664	
S12	Nonsense	C-T	CAA-TAA	Q369Stop	368	
S13	Nonsense	G-T	GAG-TAG	E129Stop	128	
UV 1	Nonsense	G-T	GAG-TAG	E458Stop	457	gDNA sequenced, no mutation found in *gaaR* (Niu *et al.*, 2017)
UV 3	Nonsense	T-A	TTA-TAA	L101Stop	100	No mutation found in *gaaR* (Niu *et al.*, 2017)
UV 4	Nonsense	C-T	CAA-TAA	Q369Stop	368	No mutation found in *gaaR* (Niu *et al.*, 2017)
UV 5	Nonsense	G-A	TGGGGA-TAGGAA	W270Stop	269	No mutation found in *gaaR* (Niu *et al.*, 2017)
UV 6	Missense	GT-AA	GTA-AAA	V684K		No mutation found in *gaaR* (Niu *et al.*, 2017)
UV 8	Missense	G-A	GGA-GAA	G526E		gDNA sequenced, no mutation found in *gaaR* (Niu *et al.*, 2017)
UV12	Missense	A-G	AAT-GAT	N438D		No mutation found in *gaaR* (Niu *et al.*, 2017)
UV13	Frameshift	Insertion of G	GCG-GGC	A308G	438	No mutation found in *gaaR* (Niu *et al.*, 2017)
UV14	Missense	A-G	AAT-GAT	N521D		
UV15	Missense	GT- AA	GTG-AAG	V193K		
UV16	Missense	GAC-ATT	GAC-ATT	D640I		

Table S3. Mutations in *gaaX* in the UV mutants showing constitutive expression of pectinases. aa and gDNA represents amino acid and genomic DNA, respectively. (Continued)

Strain	Type of mutation	Mutation	Codon change	aa change	Predicted length of the mutated protein (aa)	Remarks
UV17	Missense	T-A	GAT-GAA	D496E		
UV18	Nonsense	T-A	TTA-TAA	L548Stop	547	
UV19	Missense	T-A	ATT-AAT	I190N		
UV20	Missense	G-A	GGG-GAC	G194D		
UV21	Nonsense	C-T	CAG-TAG	Q106Stop	105	
UV22	Nonsense	G-A	TGG-TAG	W270Stop	269	
UV23	Nonsense	A-T	AAA-TAA	K435Stop	434	
UV24	Nonsense	G-T	GAG-TAG	E43Stop	42	
UV25	Nonsense	T-A	TTA-TAA	L38Stop	37	
UV26	Missense	A-T	AAT-ATT	N521I		
UV27	Nonsense	A-T	AAG-TAG	K415Stop	414	
UV28	Missense	A-T	AAC-TAC	N378Y		
UV29	Missense	G-C	GCG-CCG	A321P		
UV30	Missense and frameshift	CA-G	TCCACG-TCGCGC	ST426SR	445	
UV31	Missense and nonsense	AC-CT	TTACAA-TTCTAA	LQ101FStop	101	
UV32	Missense	A-C	TAC-TCC	Y638S		
UV33	Missense and nonsense	AC-CT	TTACAA-TTCTAA	LQ101FStop	101	
UV34	Frameshift	Deletion of T	TTT-TTG	F466L	516	
UV35	Missense	A-G	AAT-GAT	N438D		
UV36	Missense	GT-AA	GTC-AAC	V51N		
UV37	Missense and frameshift	CCC-TC	GGCCCG-GGTCGC	GP579GR	593	
UV38	Missense	T-G	CTG-CGG	L607R		
UV39	Missense	G-A	GGG-GAG	G194E		
UV40	Missense	A-G	AAT-GAA	N438E		
UV41	Missense or nonsense	A-T	ATG-TTG	M1L	- (if nonsense)	
UV42	Nonsense	C-A	TCG-TAG	S306Stop	305	
UV43	Nonsense	AC-CT	CAA-TAA	Q102Stop	101	

Table S3. Mutations in *gaaX* in the UV mutants showing constitutive expression of pectinases. aa and gDNA represents amino acid and genomic DNA, respectively. (Continued)

Strain	Type of mutation	Mutation	Codon change	aa change	Predicted length of the mutated protein (aa)	Remarks
UV44	Missense	T-A	CTG-CAG	L565Q		
UV45	Nonsense	C-T	CGA-TGA	R349Stop	348	
UV46	Nonsense	C-T	CAG-TAG	Q269Stop	268	
UV47	Nonsense	T-A	TTA-TAA	L101Stop	100	
UV48	Nonsense	A-T	AAA-TAA	K293Stop	292	
UV49	Frameshift	Deletion of C	CCG-CGC	P513R	516	
UV50	Nonsense	C-T	CAG-TAG	Q269Stop	268	
UV51						No mutations found in *gaaX* or *gaaR*
JN103.1						No mutation found in *gaaX*, mutation found in *gaaR*
UV53	Nonsense	AC-TT	CAA-TAA	Q102Stop	101	
UV54	Nonsense	T-G	TAT-TAG	Y535Stop	534	
UV55						No mutations found in *gaaX* or *gaaR*
UV56	Nonsense	C-T	CAG-TAG	Q269Stop	268	
UV57	Nonsense	C-A	TCG-TAG	S593Stop	592	
UV58	Missense	T-C	CTG-CCG	L345P		

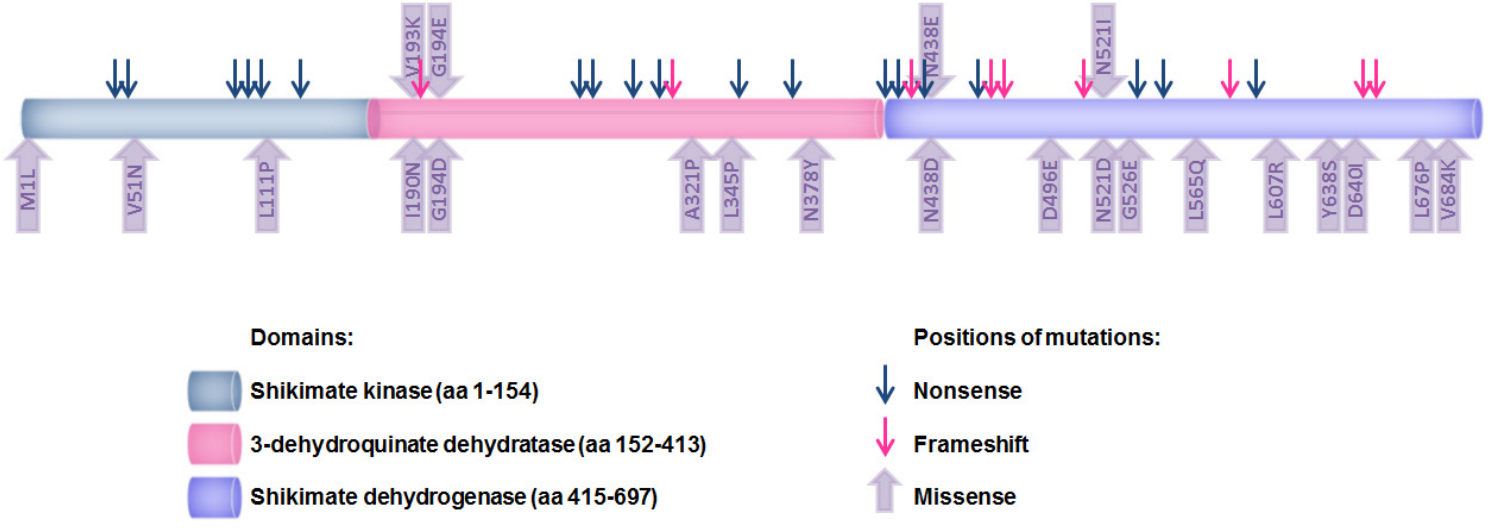

Figure S1. Schematic representation of the domains present in GaaX and mutations in the UV mutants showing constitutive expression of pectinases. Domains in GaaX were assigned based on multiple alignment of protein sequences of *A. niger* AROM (NRRL3_11273) C-terminal part (aa 863-1856) and GaaX (NRRL3_08194), and the *E. coli* shikimate kinase (AroL, NP_414922.1), 3-dehydroquinate dehydratase (AroD, NP_416208.1) and shikimate dehydrogenase (AroE, NP_417740.1), using COBALT (Papadopoulos and Agarwala, 2007).

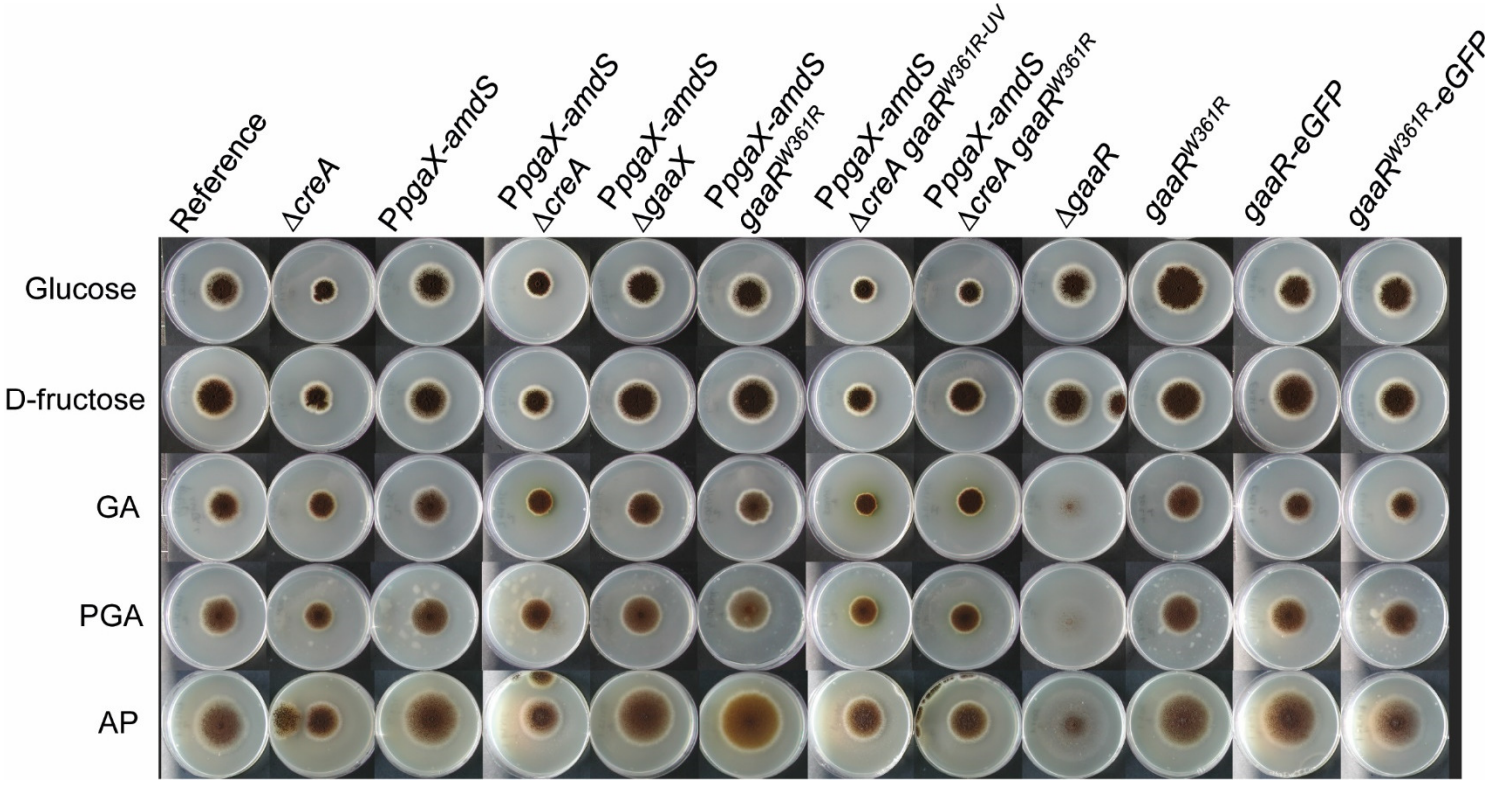

Figure S2. Radial growth assay of the reference (MA234.1), *ΔcreA* (XY1.1), P*pgaX-amdS* (JC1.5), P*pgaX-amdS ΔcreA* (JN29.2), P*pgaX-amdS ΔgaaX* (JN123.1), P*pgaX-amdS gaaR*^{W361R} (JN130.4), P*pgaX-amdS ΔcreA gaaR*$^{W361R\text{-}UV}$ (JN103.1), P*pgaX-amdS ΔcreA gaaR*^{W361R} (JN129.1), *ΔgaaR* (SO1.1), *gaaR*^{W361R} (SO2.1), *gaaR-eGFP* (EA31.1), *gaaR*^{W361R}*-eGFP* (EA32.1), and strains on solid MM containing 50 mM glucose, D-fructose or GA, or 1% PGA or pectin as the carbon source after 7 days at 30 °C.

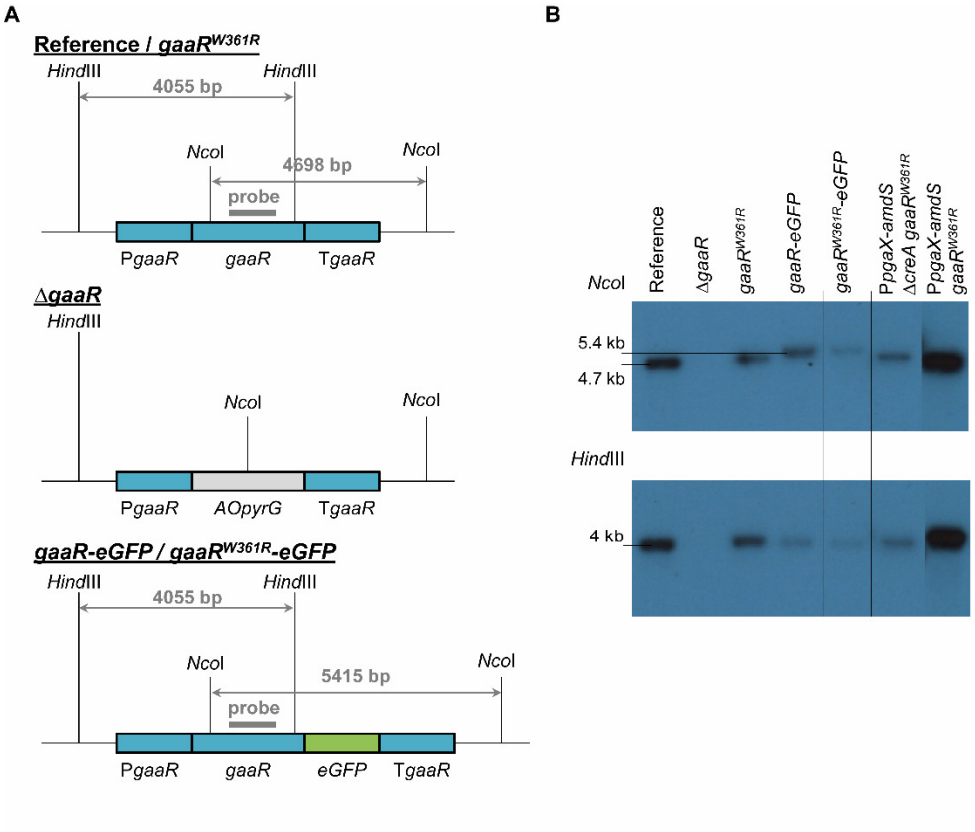

Figure S3. Southern blot analysis of genomic DNA to verify the integration of the *gaaR*^{W361R}, *gaaR-eGFP* and *gaaR*^{W361R}*-eGFP* constructs to the *gaaR* locus in the strains created in this study. **A** Schematic representation of the target gene locus in the reference (MA234.1), *ΔgaaR* (SO1.1), *gaaR*^{W361R} (SO2.1), *gaaR-eGFP* (EA31.1), *gaaR*^{W361R}*-eGFP* (EA32.1), P*pgaX-amdS ΔcreA gaaR*^{W361R} (JN129.1) and P*pgaX-amdS gaaR*^{W361R} (JN130.4) strains. The probe binds to *gaaR* downstream of the *Nco*I restriction site and upstream of the *Hind*III restriction site. Expected band sizes are indicated. **B** Southern blot after hybridization. The lines between genomic DNA samples indicate that the left and right parts of the same blot were combined after removing unnecessary lanes. 4698-bp and 4055-bp bands are visible when the genomic DNA of the reference strain is digested with *Nco*I and *Hind*III, respectively.

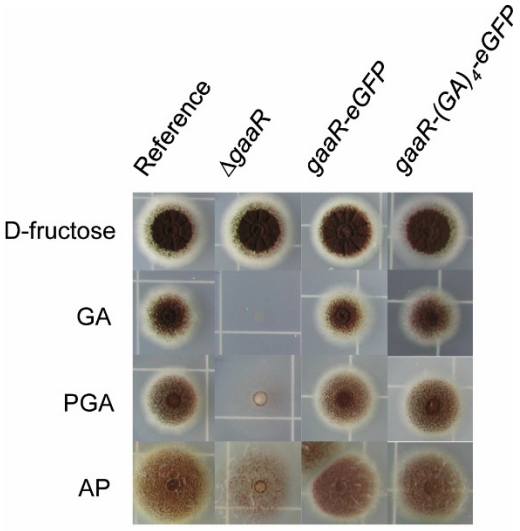

Figure S4. Radial growth assay of the reference (MA234.1), *ΔgaaR* (JN35.1), *gaaR-eGFP* (EA29.14) and *gaaR-(GA)$_4$-eGFP* (EA30.6) on solid MM containing 50 mM D-fructose or GA, or 1% PGA or pectin as the carbon source after 5 days at 30 °C.

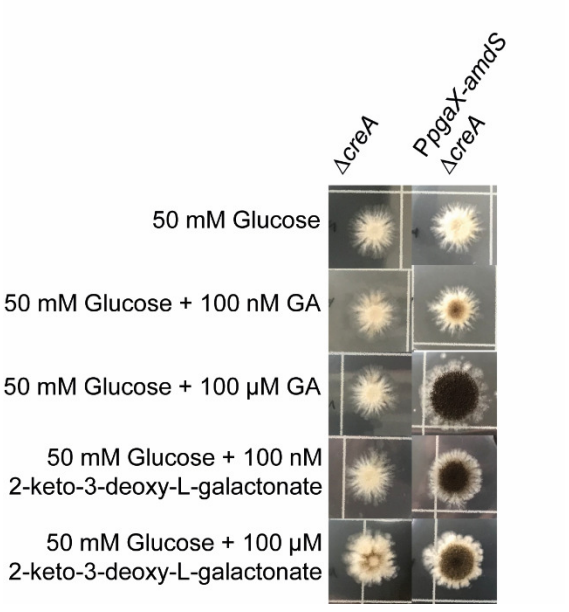

Figure S5. Growth phenotype of the *ΔcreA* (XY1.1) and P*pgaX-amdS ΔcreA* (JN29.2) strains after 10 days at 30 °C on solid MM containing 50 mM glucose or 50 mM glucose with 100 nM or 100 µM GA or 2-keto-3-deoxy-L-galactonate as the carbon source. All plates contain 10 mM acetamide as the sole nitrogen source.

CHAPTER 8

Conclusions and outlook to the future

Pectinases are enzymes naturally secreted by *Aspergillus niger* to degrade the plant cell wall polysaccharide pectin (de Vries et al., 2017). These enzymes are also used in several applications in the food and feed industry, and in the industrial hydrolysis of plant biomass for subsequent production of biofuel and high-value biochemicals. In order to produce large amounts of pectinases to be used industrially, *A. niger* has been used as a work horse (Kashyap et al., 2001; Edwards and Doran-Peterson, 2012; Khan et al., 2013; Toushik et al., 2017). The expression of the genes encoding pectinases is induced in response to the presence of D-galacturonic acid (GA), the main sugar acid in pectin (Martens-Uzunova and Schaap, 2008). The requirement of the presence of the inducer in the fungal cultivation medium hampers the use of cheap and sustainable feedstock. Therefore, it is important to investigate the transcriptional mechanism that regulates the expression of the genes encoding pectinases in *A. niger* and find out ways to constitutively (i.e. independently of the inducer) produce pectinases.

The promoter element GARE required for the transcriptional activation of pectinases was known in both *A. niger* and *Botrytis cinerea*, which allowed the identification of the first fungal GA-responsive Zn_2Cys_6 type transcription factor BcGaaR in *B. cinerea* via the yeast-one-hybrid technique (Niu et al., 2015; Zhang et al., 2016). *A. niger* GaaR was identified via its homology to BcGaaR, and was found to be required for growth on GA and polygalacturonic acid, which is the most abundant pectin substructure. GaaR was found to be essential for the GA-responsive expression of the genes encoding several pectinases, (putative) GA-transporters and GA catabolic pathway enzymes (Chapter 3).

A promoter-reporter strain containing the promoter of the pectinase gene *pgaX* upstream of the reporter gene *amdS* (Niu et al., 2015) was utilized in a forward genetic screen to isolate mutants with constitutive expression of the genes encoding pectinases. One particular gene was found to be mutated in several mutants. This gene was called the repressor of GA utilization, *gaaX*. *gaaX* and *gaaR* are located side-by-side on the chromosome of *A. niger*, similar to the activator-repressor modules *qutA-qutR* and *qa-1F-qa-1S* involved in the utilization of quinic acid in *Aspergillus nidulans and Neurospora crassa*, respectively. In all systems, deletion of the repressor results in constitutive expression of target genes which requires the presence of the corresponding Zn_2Cys_6 type transcriptional activator (Lamb et al., 1996; Grant et al., 1988; Giles et al., 1985, Chapter 4). Moreover, *A. nidulans* strains expressing multiple copies of *qutA* showed constitutive expression of the genes involved in quinic acid utilization (Lamb et al., 1996). Similarly, overexpression of *gaaR* in *A. niger* resulted in constitutive expression of the genes involved in the breakdown of pectin and utilization of GA (Chapter 6). These results indicate that the activity of GaaR is inhibited by GaaX under non-inducing conditions via protein-protein interaction, and that GA-responsive gene expression only requires the presence of active GaaR relieved/escaped from GaaX inhibition. This suggests that the

inducer binds to GaaX resulting in a GaaX-unbound and active form of GaaR under inducing conditions. Because overexpression of the activators involved in GA or quinic acid utilization have similar effects, it is likely that activation mechanism of both QutA/QutR and GaaR/GaaX are conserved and comparable. Whether overexpression of *gaaX* results in a superrespressed phenotype like overexpression of *qutR* does (Lamb et al., 1996) requires further investigation.

A. niger catabolizes GA into pyruvate and glycerol via a four step catabolic pathway involving four enzymes (GaaA-GaaD) (Martens-Uzunova and Schaap, 2008). Comparative analysis of GA catabolic pathway deletion mutants (*ΔgaaA*, *ΔgaaB*, *ΔgaaC*, and *ΔgaaD*) revealed that 2-keto-3-deoxy-L-galactonate is the physiological inducer of the GA-responsive genes. *ΔgaaB* mutant accumulated the preceding metabolite L-galactonate, could not grow on GA and lacked the induced expression of GA-responsive genes, indicating that it cannot produce 2-keto-3-deoxy-L-galactonate. 2-keto-3-deoxy-L-galactonate accumulated in the *ΔgaaC* mutant that is not able to catabolize it further and could not grow on GA. This accumulation resulted in a significant increase in the expression of the genes encoding pectinases (Chapter 5). In addition, 2-keto-3-deoxy-L-galactonate was found to be more potent than GA in inducing the expression of pectinase gene *pgaX* (Chapter 7). This latter observation also supports the idea that 2-keto-3-deoxy-L-galactonate is the physiological inducer.

GaaX of *A. niger*, QutR of *A. nidu*lans and qa-1S of *N. crassa* are all multidomain proteins with sequence similarity to the three C-terminal domains of AROM, a large pentafunctional protein involved in the shikimate pathway. This indicates that these repressors share a common evolutionary origin. The last three domains of the AROM protein encode the shikimate kinase, 3-dehydroquinate dehydratase and shikimate dehydrogenase enzymes (Figure 1) (Lamb et al., 1996). In *A. nidulans*, dehydroquinic acid can be converted to dehydroshikimic acid via different 3-dehydroquinate dehydratase isoenzymes, AROM 3-dehydroquinate dehydratase or QutE, in the biosynthetic shikimate pathway or the quinic acid catabolic pathway, respectively (Hawkins et al., 1993). In *N. crassa*, dehydroquinic acid was suggested to be the physiological inducer of the quinic acid catabolic pathway genes, as mutants that accumulate or cannot produce dehydroquinic acid showed induced or diminished expression of these genes, respectively (Giles et al., 1967). Huiet and Giles identified two missense mutations in non-inducible *N. crassa* strains in the qa-1S domain showing homology with the AROM shikimate dehydrogenase, indicating that this domain might interact with the inducer (Huiet and Giles, 1986). As the outcome of these experiments are not in full agreement with each other, direct interaction studies are required to determine which part (or parts) of the repressor interacts with the inducer. Hawkins et al. have proposed that the QutR domain homologous to the AROM 3-dehydroquinate dehydratase can recognize and bind to dehydroquinic acid produced from quinic acid, but does not react on it enzymatically in *A. nidulans* (Hawkins et al., 1993). The GA catabolic pathway enzyme GaaC acting on the inducer 2-keto-3-deoxy-L-galactonate in *A. niger* is an aldolase, thus not a 3-dehydroquinate dehydratase. However, both the aldolase domain of GaaC and the 3-dehydroquinate dehydratase domain of GaaX belong to the same Aldolase_Class I_superfamily (CD-search, https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (Marchler-Bauer et al., 2015). These findings and propositions suggest that 2-keto-3-

deoxy-L-galactonate binds to the GaaX domain homologous to the AROM 3-dehydroquinate dehydratase, and frees GaaR under inducing conditions.

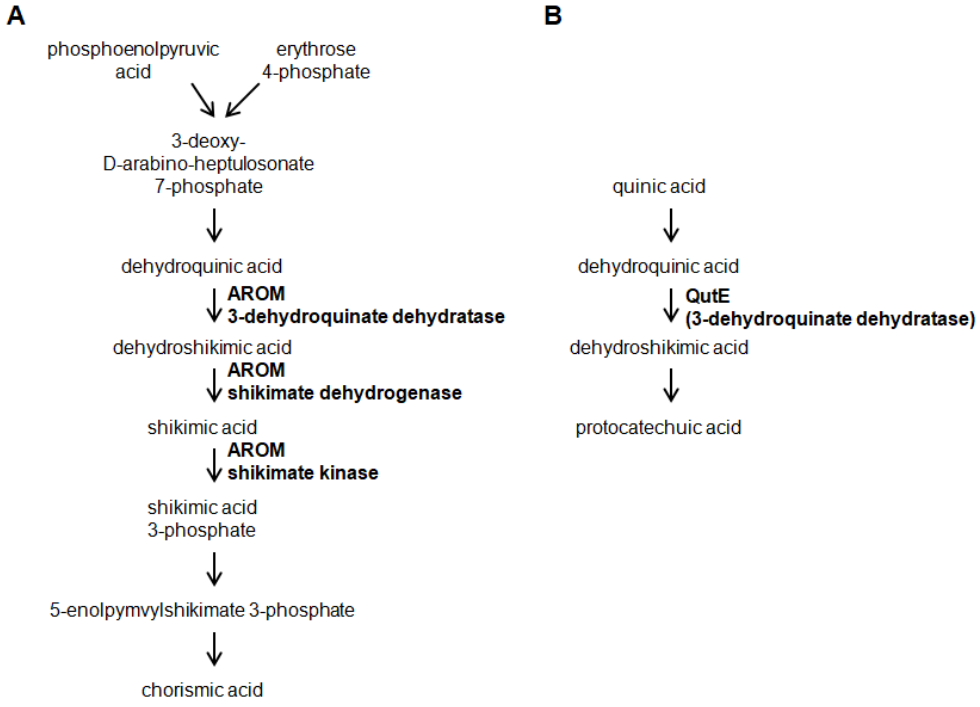

Figure 1. Comparison of the biosynthetic shikimate pathway **(A)** and the quinic acid catabolic pathway **(B)** in *A. nidulans*. Dehydroquinic acid is converted to dehydroshikimic acid in both pathways via two distinct classes of 3-dehydroquinate dehydratases (Hawkins et al., 1993).

Mutants with constitutive expression of the genes encoding pectinases obtained via the forward genetic screen contained loss-of-function mutations, including nonsense, frameshift or missense mutations, in all three domains of *gaaX* (Chapter 7). As it is not possible to predict whether the missense mutations result in a general loss of function or in a loss of the suggested interaction between the activator and repressor, it is difficult to predict the domain of GaaX interacting with GaaR. In *A. nidulans*, the 88 N-terminal amino acids of QutR were found to be sufficient to interact with QutA (Watts et al., 2002), and deletion of the QutR domains homologous to the AROM shikimate kinase, 3-dehydroquinate dehydratase or shikimate dehydrogenase still resulted in a functional repressor (Levett et al., 2000). This indicates that the mechanism of interaction of GaaR and GaaX might be different than that of QutA and QutR. The same forward genetic screen in *A. niger* also yielded to one mutant with the W361R mutation in GaaR. This mutation in the fungal-specific transcription factor domain of GaaR was found to be responsible for the constitutive expression of the genes encoding pectinases, and possibly disrupts the interaction between GaaX and GaaR (Chapter 7). Watts et al. reported that in *A. nidulans* QutR binds to QutA amino acids 449-468 (Watts et al., 2002), which are

located in the fungal-specific transcription factor domain of QutA. This comparison suggests that the fungal-specific transcription factor domains of QutA and GaaR are important for the interaction with their corresponding repressors. Future experiments, such as surface plasmon resonance measurements and structure elucidation of GaaR and GaaX proteins via crystallography would reveal the nature of interactions between GaaR, GaaX and 2-keto-3-deoxy-L-galactonate.

The subcellular localization of GaaX and GaaR was also studied in this thesis. GaaX is transcriptionally induced under inducing conditions and is localized in both the cytosol and the nucleus roughly at the same level (Chapter 4; Chapter 6). RNA seq analysis has shown that the expression of *gaaR* is not dramatically induced under inducing conditions (Chapter 3; Chapter 4). The presence of both N- or C-terminally eGFP-tagged GaaR in a *ΔgaaR* strain results in the complementation of growth on GA. However, the strain expressing the N-terminally tagged version (*eGFP-gaaR*) showed a reduced growth on GA compared to the reference strain, indicating that the N-terminal eGFP-tag might affect GaaR function. Both eGFP-GaaR and GaaR-eGFP localize mainly in the nucleus under inducing conditions, where they activate the transcription of GA-responsive genes. However, while eGFP-GaaR is localized mainly in nucleus also under non-inducing conditions, GaaR-eGFP is not. This indicates that C-terminal eGFP-tagging of GaaR restrains its nuclear accumulation under non-inducing conditions. Supporting evidence that GaaR-eGFP is restrained to accumulate in the nucleus under non-inducing conditions came from the observation that C-terminal eGFP-tagging of the constitutively active $GaaR^{W361}$ abolishes inducer-independent expression of the genes encoding pectinases, and that $GaaR^{W361}$-eGFP also does not accumulate in the nucleus under non-inducing conditions (Chapter 6; Chapter 7). The nuclear localization of the eGFP-GaaR under non-inducing conditions indicates that GaaX exhibits its repressing activity also in the nucleus. Use of more advanced techniques, such as bimolecular fluorescence complementation, would help to investigate the subcellular localization of the interaction of GaaX and GaaR.

To sum up, an important milestone in our current understanding of the transcriptional regulation of pectinases in *A. niger* was the identification and functional characterization of the transcriptional activator GaaR. After the identification of GaaR further discovery of the elements constituting the transcriptional regulatory system of pectinases, and the usage of this information to construct strains with improved pectinase production proceeded hand in hand (Figure 2). During the investigation of the physiological inducer of GA-responsive genes, the *ΔgaaC* mutant was constructed, which shows hyper-induced expression of the genes encoding pectinases on GA. The use of the *ΔgaaC* mutant to increase pectinase production has not been fully exploited yet. It is reasonable to suggest that cultivation of the *ΔgaaC* strain on GA and an additional C-source, could lead to further increased expression of the genes encoding pectinases. It is important to keep in mind that the GA-responsive genes are sensitive to CCR. Therefore, the *ΔgaaC* mutation should be combined with deletion of *creA.* The repressor of GA-utilization GaaX was identified as a result of a screen conducted to isolate mutants with constitutive expression of the genes encoding pectinases. Achievement of constitutive pectinase production via overexpression of *gaaR* provided further evidence that the presence of the inducer is not obligatory for GaaR activity, and that GaaX-unbound GaaR can activate the expression of GA-responsive genes. Finally, identification of the constitutively active *gaaR*

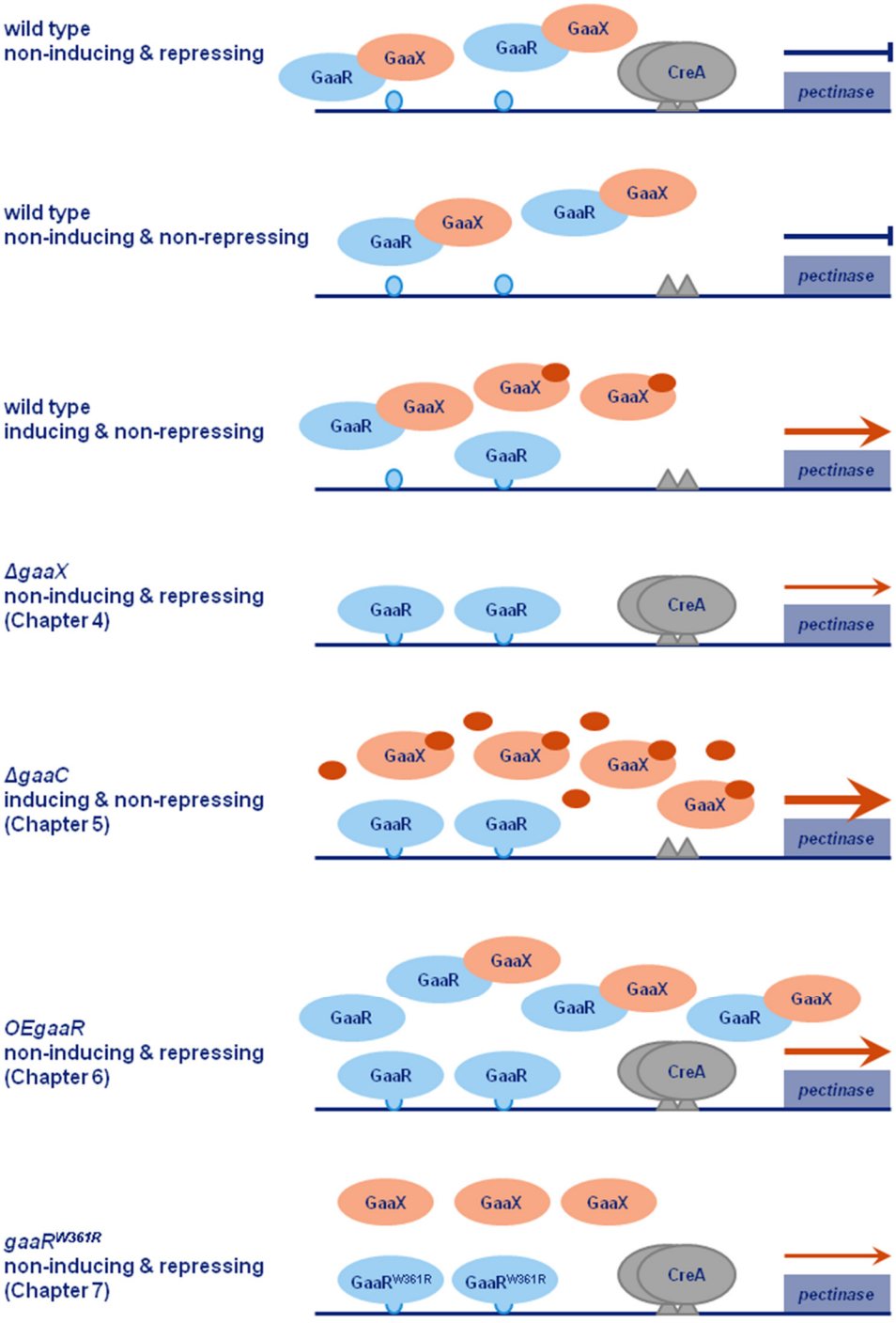
wild type
non-inducing & repressing
GaaX
GaaR
CreA
pectinase
wild type
non-inducing & non-repressing
wild type
inducing & non-repressing
ΔgaaX
non-inducing & repressing
(Chapter 4)
ΔgaaC
inducing & non-repressing
(Chapter 5)
OEgaaR
non-inducing & repressing
(Chapter 6)
gaaR^W361R
non-inducing & repressing
(Chapter 7)
GaaR^W361R

Figure 2. Current model describing the transcriptional regulation and approaches leading to increased or inducer-independent expression of most of the genes encoding pectinases in *A. niger*. Note that the amounts of the GaaR, GaaX and CreA proteins and the inducer (the small orange circle) in this scheme may not represent the real (relative) levels. Transcription factor binding sites are indicated by the blue circle and the gray triangle, and represent the binding sites of GaaR and CreA, respectively. Relative expression level of the genes encoding pectinases under each condition is roughly indicated by the thickness of the orange arrow. Deletion of CreA further increases the level of pectinase gene expression in the *ΔgaaX* and *OEgaaR* strains under repressing conditions (Chapter 4; Chapter 6).

allele which contains the W361R mutation made it possible to suggest that W361 is important for the interaction of GaaR and GaaX while it does not affect the transcriptional activity of GaaR. In this thesis, we also show the use of the state of the art CRISPR-Cas9 technology to introduce point mutated or eGFP-tagged versions of GaaR at the endogenous *gaaR* locus (Chapter 7). The main repressor CreA involved in carbon catabolite repression was found to repress the expression of different genes encoding pectinases at different degrees. Deletion of *creA* was not enough for constitutive expression of the genes encoding pectinases but further enhanced the pectinase production capacity of strains lacking *gaaX* or overexpressing *gaaR* (Chapter 4; Chapter 6). Combining the above mentioned techniques, an industrial *A. niger* strain overexpressing the constitutively active $gaaR^{W361R}$ and lacking *gaaX* and *creA* can be constructed for boosted and constitutive expression of the genes encoding pectinases.

REFERENCES

de Vries RP, Riley R, Wiebenga A, Aguilar-Osorio G, Amillis S, Uchima CA, Anderluh G, Asadollahi M, Askin M, Barry K, Battaglia E, Bayram Ö, Benocci T, Braus-Stromeyer SA, Caldana C, Cánovas D, Cerqueira GC, Chen F, Chen W, Choi C, Clum A, Dos Santos RA, Damásio AR, Diallinas G, Emri T, Fekete E, Flipphi M, Freyberg S, Gallo A, Gournas C, Habgood R, Hainaut M, Harispe ML, Henrissat B, Hildén KS, Hope R, Hossain A, Karabika E, Karaffa L, Karányi Z, Kraševec N, Kuo A, Kusch H, LaButti K, Lagendijk EL, Lapidus A, Levasseur A, Lindquist E, Lipzen A, Logrieco AF, MacCabe A, Mäkelä MR, Malavazi I, Melin P, Meyer V, Mielnichuk N, Miskei M, Molnár ÁP, Mulé G, Ngan CY, Orejas M, Orosz E, Ouedraogo JP, Overkamp KM, Park HS, Perrone G, Piumi F, Punt PJ, Ram AF, Ramón A, Rauscher S, Record E, Riaño-Pachón DM, Robert V, Röhrig J, Ruller R, Salamov A, Salih NS, Samson RA, Sándor E, Sanguinetti M, Schütze T, Sepčić K, Shelest E, Sherlock G, Sophianopoulou V, Squina FM, Sun H, Susca A, Todd RB, Tsang A, Unkles SE, van de Wiele N, van Rossen-Uffink D, Oliveira JV, Vesth TC, Visser J, Yu JH, Zhou M, Andersen MR, Archer DB, Baker SE, Benoit I, Brakhage AA, Braus GH, Fischer R, Frisvad JC, Goldman GH, Appl Microbiol Biotechnol Houbraken J, Oakley B, Pócsi I, Scazzocchio C, Seiboth B, vanKuyk PA, Wortman J, Dyer PS, Grigoriev IV (2017) Comparative genomics reveals high biological diversity and specific adaptations in the industrially and medically important fungal genus *Aspergillus*. *Genome Biology* 18(28)

Edwards MC, Doran-Peterson J (2012) Pectin-rich biomass as feedstock for fuel ethanol production. *Applied Microbiology and Biotechnology* 95:565–575

Giles NH, Partridge CW, Ahmed SI, Case ME (1967) The occurrence of two dehydroquinases in *Neurospora crassa*, one constitutive and one inducible. *Proceedings of the National Academy of Sciences of the United States of America* 58(5):1930-1937

Grant S, Roberts CF, Lamb H, Stout M, Hawkins AR (1988) Genetic regulation of the quinic acid utilization (QUT) gene cluster in *Aspergillus nidulans*. *Journal of General Microbiology* 134:347–358

Hawkins AR, Lamb HK, Moore JD, Charles IG, Roberts CF (1993) The pre-chorismate (shikimate) and quinate pathways in filamentous fungi: theoretical and practical aspects. *Journal of General Microbiology* 139:2891-2899

Huiet L, Giles NH (1986) The *qa* repressor gene of *Neurospora crassa*: Wild-type and mutant nucleotide sequences. *Proceedings of the National Academy of Sciences of the United States of America* 83(10):3381-3385

Kashyap D, Vohra P, Chopra S, Tewari R (2001) Applications of pectinases in the commercial sector: a review. *Bioresource Technology* 77:215–227

Khan M, Nakkeeran E, Umesh-Kumar S (2013) Potential application of pectinase in developing functional foods. *Annual Review of Food Science and Technology* 4:21–34

Lamb HK, Newton GH, Levett LJ, Cairns E, Roberts CF, Hawkins AR (1996). The QUTA activator and QUTR repressor proteins of *Aspergillus nidulans* interact to regulate transcription of the quinate utilization pathway genes. *Microbiology* 142:1477–1490

Levett LJ, Si-Hoe SM, Liddle S, Wheeler K, Smith D, Lamb HK, Newton GH, Coggins JR, Hawkins AR (2000) Identification of domains responsible for signal recognition and transduction within the QUTR transcription repressor protein. *Biochemical Journal* 350:189-197

Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Research* 43:D222-D226

Martens-Uzunova ES, Schaap PJ (2008) An evolutionary conserved D-galacturonic acid metabolic pathway operates across filamentous fungi capable of pectin degradation. *Fungal Genetics and Biology* 45:1449–1457

Niu J, Homan TG, Arentshorst M, de Vries RP, Visser J, Ram AF (2015) The interaction of induction and repression mechanisms in the regulation of galacturonic acid-induced genes in *Aspergillus niger*. *Fungal Genetics and Biology* 82:32–42

Toushik SH, Lee K, Lee J, Kim K (2017) Functional applications of lignocellulolytic enzymes in the fruit and vegetable processing industries. *Journal of Food Science* 88(3):583–595

Watts C, Si-Hoe SM, Lamb HK, Levett LJ, Coggins JR, Hawkins AR (2002) Kinetic analysis of the interaction between the QutA and QutR transcription-regulating proteins. *Proteins* 48:161-168

Zhang L, Lubbers RJM, Simon A, Stassen JHM, Ribera PRV, Viaud M, van Kan JAL (2016) A novel Zn_2Cys_6 transcription factor BcGaaR regulates D-galacturonic acid utilization in *Botrytis cinerea*. *Molecular Microbiology* 100:247–262

Summary

Pectin is a plant cell wall polysaccharide made of mainly D-galacturonic acid (GA) subunits. The potency of *Aspergillus niger* to naturally secrete high amounts of pectinases to degrade pectin has been utilized for the industrial production of pectinases. Industrially produced pectinases by *A. niger* are subsequently used mainly in the food industry, and for the hydrolysis of plant biomass to produce renewable energy. In this thesis the transcriptional activator GaaR, required for the expression of the genes encoding several pectinases, the (putative) GA transporters and the GA catabolic pathway enzymes was identified (Chapter 3). The *ΔgaaR* strain could not grow on GA or PGA, but showed residual growth on pectin, indicating that deletion of *gaaR* does not affect the utilization of monomeric carbon sources other than GA which are present in pectin. Transcriptome analysis of the *ΔgaaR* strain on pectin confirmed that the *ΔgaaR* strain still expresses several enzymes which mainly act on the side chains of rhamnogalacturonan I. Metabolic analysis of the GA catabolic pathway deletion mutants resulted in the identification of 2-keto-3-deoxy-L-galactonate accumulating in the *ΔgaaC* mutant. Gene expression analysis showed that accumulation of 2-keto-3-deoxy-L-galactonate in the *ΔgaaC* mutant resulted in hyperinduction of GA-responsive genes, indicating that 2-keto-3-deoxy-L-galactonate is the physiological inducer of the GA-responsive genes (Chapter 5).

Genome sequencing of several mutants that showed constitutive expression of the genes encoding pectinases resulted in the identification of the repressor protein GaaX that likely inhibits the transcriptional activity of GaaR via protein-protein interaction under non-inducing conditions (Chapter 4). One mutant with constitutive expression of the genes encoding pectinases contained the W361R mutation in GaaR, indicating that W361 is important for the proper interaction of GaaX and GaaR under non-inducing conditions (Chapter 7).

Overexpression of *gaaR* via the strong constitutive *A. nidulans gpdA* promoter brought about constitutive expression of the genes encoding pectinases, providing further evidence that GaaR is active as long as it is not inhibited by GaaX. This suggests that in the presence of the inducer, the inducer binds to GaaX resulting in a GaaX-unbound and active form of GaaR. Deletion of the main repressor involved in carbon catabolite repression, *creA*, enhanced the constitutive expression pectinases in the *OEgaaR* strain (Chapter 6).

Constitutive expression of the genes encoding plant biomass degrading enzymes like pectinases would benefit especially the use of cheap and renewable plant biomass as a feedstock for fungal fermentation. This thesis also contains a chapter reviewing the approaches to modulate the transcriptional regulation in order to increae or constitutively express plant biomass degrading enzymes in filamentous fungi (Chapter 2). These approaches include accumulation of the inducer, deletion of the carbon source specific repressor, constitutive activation and overexpression of the transcriptional activator and elimination of carbon catabolite repression. All of these approaches have been implemented successfully in this thesis to increase and/or constitutively express pectinases in *A. niger*.

Samenvatting

Pectine is een complexe suikerpolymeer dat voorkomt in plantencelwanden en bestaat voornamelijk uit D-galacturonzuur (GA) subeenheden. De potentie van *Aspergillus niger* om van nature grote hoeveelheden pectinasen uit te scheiden om pectine af te breken, is gebruikt voor de industriële productie van pectinasen. Industrieel geproduceerde pectinasen van *A. niger* worden vervolgens voornamelijk gebruikt in de voedingsmiddelenindustrie en voor de hydrolyse van plantaardige biomassa om hernieuwbare energie te produceren. In dit proefschrift is de transcriptionele activator GaaR geïdentificeerd, die vereist is voor de expressie van de genen die coderen voor verschillende pectinasen, (vermoedelijke) GA transporters en de GA katabole pathway enzymen (Hoofdstuk 3). De *ΔgaaR* stam kan niet groeien op GA of PGA, maar vertoont restgroei op pectine, hetgeen aangeeft dat deletie van *gaaR* geen invloed heeft op het gebruik van andere monomere koolstofbronnen dan GA die aanwezig zijn in pectine. Transcriptome analyse van de *ΔgaaR* stam op pectine bevestigde dat de *ΔgaaR* stam nog steeds verschillende enzymen tot expressie brengt die voornamelijk werken op de zijketens van rhamnogalacturonan I. Metaboliet analyse van de GA katabole pathway deletie-mutanten resulteerde in de identificatie van 2-keto-3-deoxy-L-galactonaat als metaboliet dat accumuleert in de *ΔgaaC* mutant. Expressie analyse heeft laten zien dat de accumulatie van 2-keto-3-deoxy-L-galactonaat in de *ΔgaaC* mutant leidt tot hyperinductie van GA-reagerende genen wat impliceert dat 2-keto-3-deoxy-L-galactonaat de fysiologische inducer is van de GA-reagerende genen (Hoofdstuk 5).
Genoom sequencing van verschillende mutanten die constitutieve expressie van pectinasen vertoonden resulteerde in de identificatie van het repressor eiwit (GaaX) dat waarschijnlijk de transcriptionele activiteit van GaaR remt via eiwit-eiwit interactie onder niet-inducerende omstandigheden (Hoofdstuk 4). Eén mutant met constitutieve expressie van pectinasen bevatte de W361R mutatie in GaaR, wat aangeeft dat W361 belangrijk is voor de juiste interactie van GaaX en GaaR onder niet-inducerende omstandigheden (Hoofdstuk 7).
Overexpressie van *gaaR* via de sterke constitutieve *A. nidulans gpdA* promotor bracht de constitutieve expressie van pectinasen teweeg, en leverde verder bewijs dat GaaR actief is, zolang het niet door GaaX wordt geïnactiveerd. Dit suggereert dat in de aanwezigheid van de inducer, deze zich bindt aan GaaX, wat resulteert in een GaaX-ongebonden en actieve vorm van GaaR. Deletie van de belangrijkste repressor betrokken bij koolstof katabolietrepressie, *creA*, verhoogde de constitutieve expressie pectinasen in de *OegaaR* stam (Hoofdstuk 6).
Constitutieve expressie van enzymen zoals pectinasen komt vooral het gebruik van goedkope en hernieuwbare plantenbiomassa als grondstof voor schimmelfermentatie ten goede. Dit proefschrift bevat ook een hoofdstuk waarin de benaderingen voor het moduleren van de transcriptionele regulatie worden beschreven om de productie plantenbiomassa afbrekende enzymen in filamenteuze schimmels te verbeteren of constitutief tot expressie te laten komen (Hoofdstuk 2). Deze benaderingen omvatten accumulatie van de inducer, het uitschakelen van de koolstofbron-specifieke repressor eiwitten, constitutieve activatie en overexpressie van de transcriptionele activator en het elimineren van koolstof katabolietrepressie. Al deze benaderingen zijn met succes

geïmplementeerd in dit proefschrift om de productie pectinasen in *A. niger* te verhogen en/of deze enzymen constitutief tot expressie te brengen.

Curriculum vitae

Ebru Alazi was born on the 6th of September 1985 in Üsküdar (Istanbul), Turkey. She graduated from Cagaloglu Anatolian High School in 2003, and started her study in Chemical Engineering in Bogazici University, specializing in Biochemical Engineering. After obtaining her Bachelor of Science degree in 2008, she continued with her Master of Science study in Chemical Engineering in the same university. She conducted a research project in the Biosystems Engineering Research Group and graduated in 2010. Following her enthusiasm for biotechnology, she started her Master of Science study in Biology in Leiden University in the Netherlands, specializing in Molecular and Cellular Biosciences. She performed her research project in the Molecular Microbiology and Biotechnology Research Group and graduated *cum laude* in 2013. In 2014, she started her PhD in the same research group under the supervision of Dr. A.F.J. Ram and Prof. dr. P.J. Punt. The results of her PhD study are described in this thesis.

Publications

Alazi E, Ram AFJ (2018) Modulating transcriptional regulation of plant biomass degrading enzyme networks for rational design of industrial fungal strains. Submitted to *Frontiers in Bioengineering and Biotechnology*

Alazi E, Niu J, Otto SB, Pham TTM, Tsang A, Ram AFJ (2018) W361R mutation in GaaR, the regulator of D-galacturonic acid responsive genes, leads to constitutive expression of pectinases in *Aspergillus niger*. Submitted to *MicrobiologyOpen*

Alazi E, Knetsch T, Di Falco M, Reid ID, Arentshorst M, Visser J, Tsang A, Ram AFJ (2018) Inducer-independent production of pectinases in *Aspergillus niger* by overexpression of the D-galacturonic acid responsive transcription factor *gaaR*. *Applied Microbiology and Biotechnology* 102:2723-2736

Alazi E, Khosravi C, Homan TG, du Pré S, Arentshorst M, Di Falco M, Pham TTM, Peng M, Aguilar-Pontes MV, Visser J, Tsang A, de Vries RP, Ram AFJ (2017) The pathway intermediate 2-keto-3-deoxy-L-galactonate mediates the induction of genes involved in D-galacturonic acid utilization in *Aspergillus niger*. *FEBS Letters* 591:1408-1418

Niu J, Alazi E, Reid ID, Arentshorst M, Punt PJ, Visser J, Tsang A, Ram AFJ (2017) An evolutionarily conserved transcriptional activator-repressor module controls expression of genes for D-galacturonic acid utilization in *Aspergillus niger*. *Genetics* 205:169-183

Alazi E, Niu J, Kowalczyk JE, Peng M, Aguilar Pontes MV, van Kan JAL, Visser J, de Vries RP, Ram AFJ (2016) The transcriptional activator GaaR of *Aspergillus niger* is required for release and utilization of D-galacturonic acid from pectin. *FEBS Letters* 590:1804-1815

Park J, Hulsman M, Arentshorst M, Breeman M, Alazi E, Lagendijk EL, Rocha MC, Malavazi I, Nitsche BM, van den Hondel CAMJJ, Meyer V, Ram AFJ (2016) Transcriptomic and molecular genetic analysis of the cell wall salvage response of *Aspergillus niger* to the absence of galactofuranose synthesis. *Cellular Microbiology* 18(9):1268-1284