Identification and characterization of the 3,4-dihydroxybenzoic acid and quinic acid catabolic pathway genes in *Aspergillus niger*

Michael Sgro

A Thesis
in
The Department
of
Biology

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science (Biology) at
Concordia University
Montreal, Quebec, Canada

June 2020

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Michael Sgro

Entitled: Identification and characterization of the 3,4-dihydroxybenzoic acid and quinic acid catabolic pathway genes in *Aspergillus niger*

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Biology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____ Chair

Dr. Brandon Findlay

_____ Examiner

Dr. Justin Powlowski

_____ Examiner

Dr. Aashiq Kachroo

_____ Supervisor

Dr. Adrian Tsang

Approved by _____

Dr. Robert Weladji, Graduate program director

September 1, 2020 _____

Dr. André Roy, Dean of Arts and Science

ABSTRACT

Identification and characterization of the 3,4-dihydroxybenzoic acid and quinic acid catabolic pathway genes in *Aspergillus niger*

Michael Sgro

Plants produce an abundant supply of many aromatic compounds including those derived from the aromatic polymer lignin, which comprises approximately 25% of biomass on land, and quinic acid, which can comprise up to 10% of the weight of dry leaf litter. Many aromatic compounds have industrial applications, e.g. as building blocks for plastics, biofuels, drugs, preservatives, flavoring agents, or antioxidants. Many fungal species are able to degrade aromatic compounds and utilize them as a carbon source. While aromatic metabolism is well understood in bacteria, the central pathways remain poorly understood in fungi and the genes involved largely unknown and lacking proper annotation. In this study, we have used data from comparative transcriptomics and the characterization of knockout mutants by growth phenotype and metabolic analysis to determine the genes and enzymes involved in each step in the catabolism of quinic acid and its derivative 3,4-dihydroxybenzoic acid in the filamentous fungus *Aspergillus niger.* The genes encoding quinate dehydrogenase, protocatechuate-3,4-dioxygenase, 3-carboxy-cis,cis-muconate cyclase, 3-carboxymuconolactone hydrolase/decarboxylase, β-ketoadipate CoA transferase, and β-ketoadipyl-CoA transferase in *Aspergillus niger* were determined to be NRRL3_08520, NRRL3_01405, NRRL3_02586, NRRL3_01409, NRRL3_01886, and NRRL3_01526, respectively. Two quinic acid transporters, NRRL3_11036 and NRRL3_05631, were also identified. NRRL3_05631 is a previously uncharacterized gene which functions at a lower pH than NRRL3_11036. Continued research will determine the entire metabolic network formed by the catabolic pathways of the other central aromatic intermediates, providing a comprehensive understanding of fungal aromatic metabolism for the first time. Knowledge of these pathways and the genes involved can lead to improved production of various valuable compounds, utilization of lignin, bioremediation or tolerance of toxic aromatic compounds.

Acknowledgments

I would like to express my deepest gratitude to my supervisor Dr. Adrian Tsang for developing this project and guiding me through it every step of the way. He gave me a chance as a graduate diploma student and I wouldn't be where I am now, continuing on to get my PhD, without him.

My sincere thanks to Dr. Justin Powlowski, who is collaborating with us on this project and helped in more ways than I could list.

There are many colleagues I would like to thank. Dr. My Pham, who somehow finds the time to help me and the other students in our lab every day despite having ten different projects of her own. Sylvester Palys, who spent months teaching me all the techniques we use in this lab and helping me get started. Dr. Marcos DiFalco, our expert on mass spectrometry who has spent many hours teaching me about it or running samples for me. Marie-Claude Moisan, for help with the RNA sequencing and working with DHBs. Dr. Jean-Paul Ouedraogo and Dr. Letian Song who developed some of the techniques that were used. Additional thanks to Idowu Bello-Osagie, who has begun working on the expanded project, Logan Robeck, Gregory Evdokias, Marie Beigas, and the rest of the students and employees in our lab who have all helped me in many ways.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

2,3-DHB	2,3-dihydroxybenzoic acid
2,4-DHB	2,4-dihydroxybenzoic acid
2,5-DHB	2,5-dihydroxybenzoic acid
3,4-DHB	3,4-dihydroxybenzoic acid
CM	Complete Media
CoA	Coenzyme A
CRISPR	Clustered Regularly Interspaced Palindromic Repeats
LC-MS	Liquid Chromatography-Mass Spectrometry
LIC	Ligation Independent Cloning
MM	Minimal Media
MS	Mass Spectrometry
PEG	Poly-Ethylene Glycol
SRM	Selective Regeneration Media
TCA	Tricarboxylic acid

1. Introduction

1.1 Aromatic compounds in plants

Plants contain a wide variety of aromatic compounds. These compounds are present in many forms: in polymers such as lignin, linked to polysaccharides, as large polyphenolic compounds or as monocyclic compounds found freely in plant cells (Mäkelä et al., 2015; Lubbers et al., 2019). Although some are primary metabolites, aromatic compounds in plants are generally secondary metabolites, compounds that are not required for the normal growth or survival of the species. These aromatic secondary metabolites have many different beneficial effects including defense against herbivores, insects, and pathogenic fungi; UV protection; pigmentation; and production of lignin (Pagare et al., 2016)

1.1.1 Lignin

Lignin is a complex polymer of aromatic compounds consisting of *p*-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol which are incorporated into lignin as *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) derivates, respectively (Figure 1A) (Schoenherr et al., 2018). The relative abundance of these monomers varies between and within species, as does the structure of the polymer (Whetten and Sederoff, 1995). One example of a portion of the structure of lignin is shown in Figure 1B.

Lignin is one of the three main components of lignocellulose, along with cellulose and hemicellulose. After cellulose, it is the second most abundant polymer on Earth, accounting for an estimated 30% of all organic carbon (Boerjan, Ralph, & Baucher, 2003). Lignin binds both covalently and non-covalently to hemicellulose in the plant cell wall to enhance rigidity and make the cells impermeable to water, allowing the necessary water transport functions of terrestrial plants to occur (Nishimura et al., 2018). The relative abundances of lignin, cellulose, and hemicellulose are variable between and within species, and vary between types of tissue within individual plants (Whetten and Sederoff, 1995).

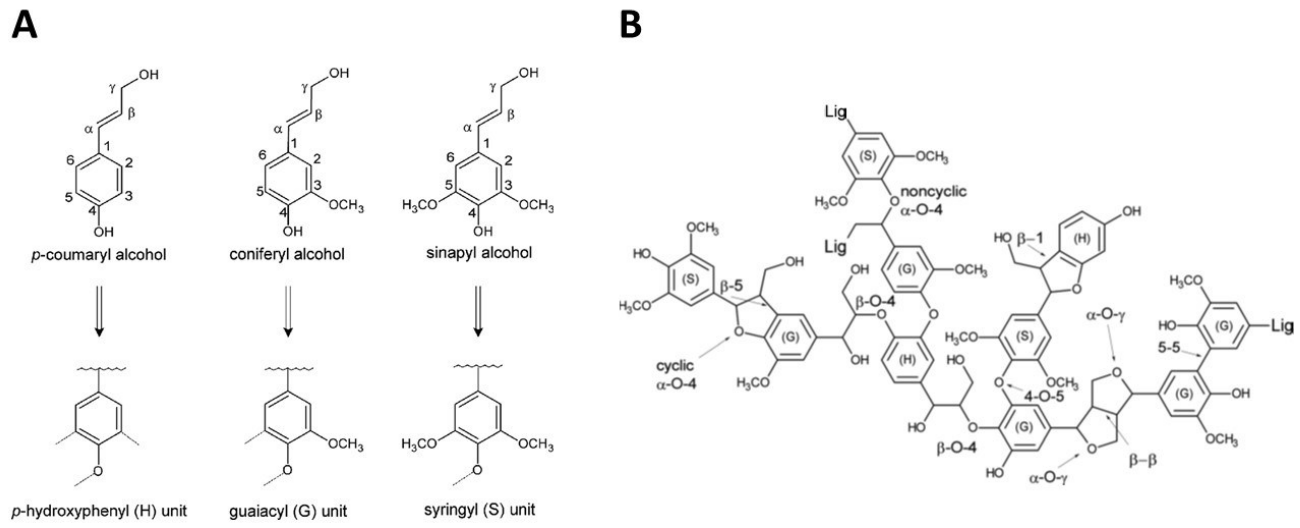

Figure 1. The building blocks (A) and an example of the structure of lignin (B). Modified from Schoenherr et al. (2018).

Due to increasing global population, climate change, and the limited supply of petroleum, global demand for alternative energy, fuels and chemicals is increasing. The abundance, cheap cost, and carbon neutrality of lignocellulosic biomass make it a very attractive option for these purposes (Wenger & Stern, 2019; Sun et al., 2018). Although it is being used in a limited capacity to produce bioethanol, biorefineries currently focus on the cellulose and hemicellulose fractions, while using the lignin for combustion to meet internal energy demands. However, this use of lignin has low value and only requires about 40% of the total amount of lignin produced (Wang et al., 2019; Sun et al., 2018).

To make lignocellulosic biorefineries more economically feasible, lignin valorization, i.e. conversion of lignin into high value fuels or chemicals, will be essential (Beckham et al., 2016). There are many possible industrial uses of lignin-derived aromatic compounds including production of a variety of plastics and polymers, adhesives, drugs, cosmetics, and food additives including antioxidants, flavoring agents, and preservatives (Lubbers et al., 2019; Sun et al., 2018). For this reason, lignin depolymerization and valorization have become increasingly

important areas of research (Ragauskas et al., 2014; Xu et al., 2014). This is seen in the trends in scientific literature in the past decade. The cumulative number of publications regarding lignin valorization increased by more than 7-fold between 2010 and 2016, to over 350 (Abejón et al., 2018). Yearly publications regarding lignin depolymerization approximately doubled between 2010 and 2013 to over 60 per year (Xu et al., 2014). Similarly, publications containing both the terms 'biorefinery' and 'feedstock' approximately double between 2010 and 2013, and nearly doubled again between 2013 and 2017, to approximately 150 per year (Wenger & Stern, 2019).

1.1.2 Quinic acid and shikimic acid

The plant metabolites shikimic acid and quinic acid are also highly abundant in many plants, comprising up to 10% of the dry weight of leaf litter (Hawkins et al., 1993). These compounds are closely related and can be interconverted in plants using the enzyme quinate dehydratase (Herrmann and Weaver 1999). Shikimic acid is a particularly important aromatic compound as the shikimate pathway is the major aromatic biosynthesis pathway in plants. The shikimate pathway begins with phosphoenol pyruvate and ends with chorismate, the precursor of the aromatic amino acids phenylalanine, tyrosine, and tryptophan (Herrmann et al., 1995). These amino acids are then used as precursors for many other plant secondary metabolites including lignin and flavonoids, as well as some plant hormones including salicylic acid (Tzin and Galili, 2010; Guo et al., 2014). Esters of quinate and shikimate have also been shown to have a role in lignin biosynthesis. Silencing of a gene in tobacco encoding an enzyme that catalyzes the synthesis of shikimate and quinate esters such as chlorogenic acid caused a 15% reduction in lignin content (Boudet et al., 2003).

1.1.3 Other plant aromatics

A wide variety of other aromatic compounds are also found linked to polysaccharides or freely in the cells. These include monocyclic aromatic compounds such as ferulic acid, coumaric acid, vanillic acid, hydroxybenzoic acids, and caffeic acid (Mäkelä et al., 2015). Polyphenols in plants include classes of compounds such as phenylpropanoids, flavonoids, coumarins, xanthones, stilbenoids, and phenolic acids and alcohols (Abbas et al., 2017; Mäkelä et al., 2015).

Thousands of these compounds are present in plants. It has been estimated there may be up to 100,000 plant secondary metabolites, many of which are aromatics (Metcalf and Kogan, 1987). One of the largest and most diverse groups is the flavonoids, with at least 4000-6500 different flavonoids having already been identified in plants (Ververidis et al., 2007; Abbas et al., 2017). Polyphenols are further diversified by prenylation, the addition of a 15- or 20-carbon isoprenoid group. Over 1000 prenylated polyphenols have been discovered in plants (Yazaki et al., 2009). Polyphenols have been found to have many beneficial biological activities in humans including antioxidant, anticarcinogenic, antibacterial, antiviral, antityrosinase, estrogenic, inhibition of sulfotransferase, antinitric oxide, and inhibition of phospholipase (Sugiyama et al., 2011).

Tannins, the second most abundant group of aromatic compounds in plants, are found in most plant species and tissues and can comprise up to 20% of the dry weight of plants (Adamczyk et al., 2018; Mäkelä et al., 2015). Two classes of tannins exist in plants: the flavonoid-based condensed tannins, and hydrolyzable tannins, consisting of gallic acid esters with a central polyol (Gross, 2008). Tannins bind to proteins and carbohydrates to decrease the digestibility of those nutrients, and also give plants an astringent taste and odor to make them less appealing to animals and insects (Dykes and Rooney, 2007). Tannins, particularly the condensed tannins, also have high antioxidant activity and numerous other health benefits in humans including anticarcinogenic, gastroprotective, antiulcerogenic, and cholesterol-lowering properties (Prior and Gu, 2005).

1.2 Degradation of lignin and tannins

Soil fungi and bacteria play an important role in the carbon cycle by recycling the large quantity of aromatic compounds produced by plants. Lignocellulosic biomass is degraded mainly by fungi, particularly the white-rot and brown-rot basidiomycetes (Hatakka and Hammel, 2011). The white-rot fungi are capable of degrading lignin by producing extracellular enzymes such as peroxidases and laccase (Waldner et al., 1988; Niku-Paavola et al., 1988). These enzymes decay plant matter turning it soft and white, or light in color. White-rot fungi comprise more than 90% of all wood-rotting basidiomycetes (Gilbertson, 1980).

The brown-rot fungi are also capable of degrading cellulose and hemicellulose, but leave behind demethoxylated lignin (Dey et al., 1994). Similarly, many ascomycete fungi are capable of soft-rot decay of plant matter, which degrades cellulose and hemicellulose but not lignin. Brown-rot and soft-rot fungi potentially have higher cellulolytic and hemicellulolytic activity than white-rot fungi (Sista Kameshwar and Qin, 2018). Degradation of lignin by white-rot fungi results in the formation of low molecular weight bi- or monocyclic aromatic compounds such as vanillin, vanillic acid, ferulic acid, syringic acid, syringaldehyde, and benzoic acids like 3,4-dihydroxybenzoic (Kamimura et al., 2017). Figure 2 shows some of the major compounds produced by the white rot fungi and some of the pathways bacteria use to convert them into common intermediates and then into TCA cycle intermediates.

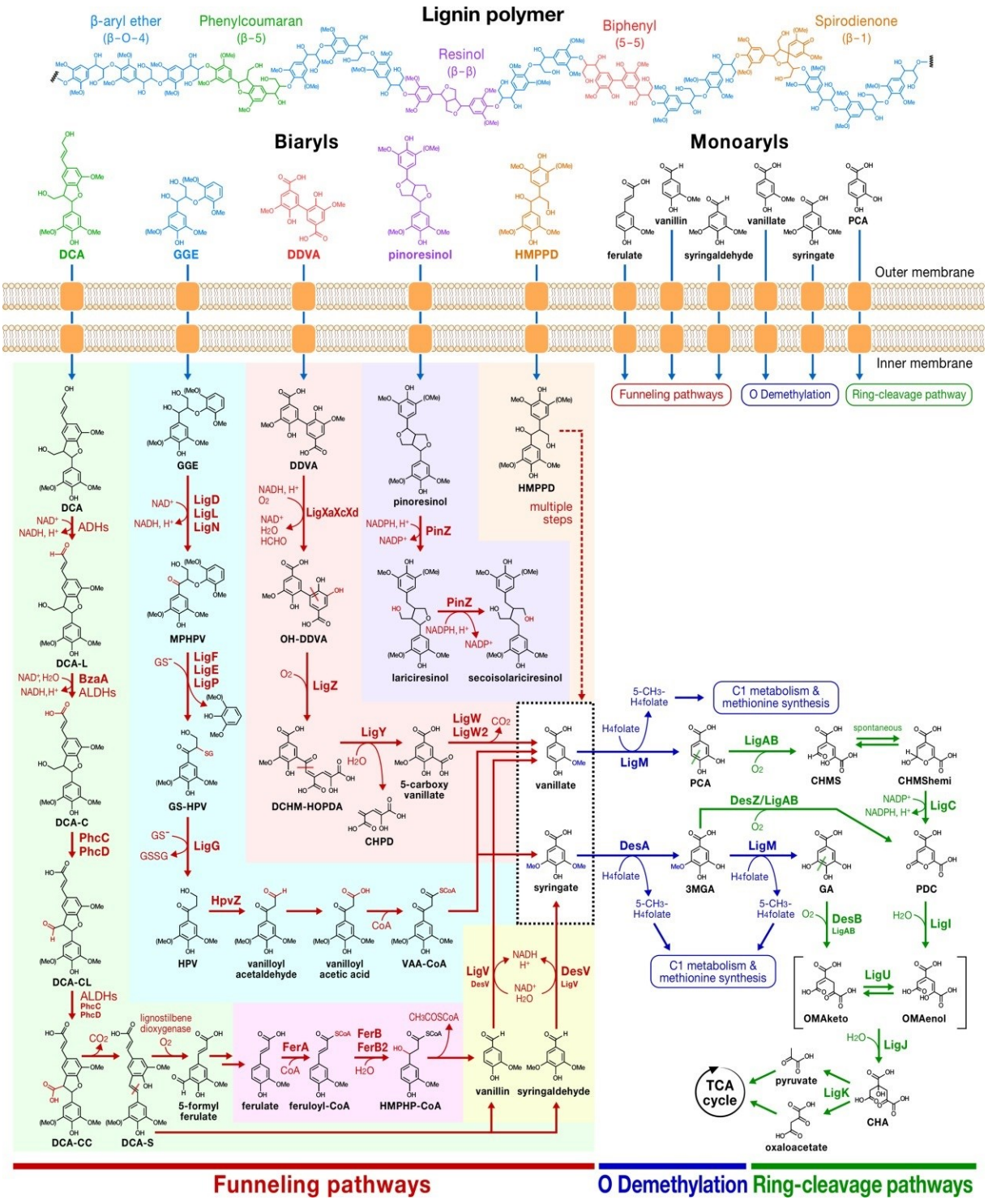

Figure 2. Lignin degradation products and metabolic pathways for their utilization in bacteria. DCA = dehydrodiconiferyl alcohol. GGE = guaiacylglycerol-β-guaiacyl ether. DDVA = 5,5′-dehydrodivanillate. HMPPD = 1,2-bis(4-hydroxy-3-methoxyphehyl)-1,3-propanediol. PCA =

protocatechuic acid (3,4-dihydroxybenzoic acid). 3MGA = 3-*O*-methylgallate. Modified from Kamimura et al. (2017).

Similar to degradation by white-rot fungi, pre-treatment and hydrolysis of lignocellulosic biomass to prepare for fermentation produces many lower molecular weight aromatic compounds including vanillin, syringic acid, ferulic acid, coniferyl aldehyde, 4-hydroxybenzoic acid, acetoguaiacone, catechol, and coumaric acid (Adeboye et al., 2014; Jung and Kim, 2015). Hydrolyzable tannins can also be degraded by tannase enzymes in fungi to produce the aromatic gallic acid and glucose (Aguilar et al., 2007), and condensed tannins can also be degraded through flavonoid intermediates into phloroglucinol carboxylic acid and protocatechuic acid (Bhat et al., 1998).

1.3 Catabolism of aromatic compounds

Due to the high levels released into the environment from the degradation of plant matter, most soil fungi and bacteria are able to catabolize a wide variety of aromatic compounds and utilize them as a carbon source. This ability is also important because many of these plant-derived aromatic acids and toxic to microorganisms, even at low concentrations (Mäkelä et al., 2015; Jung and Kim, 2015).

In aerobic fungi and bacteria, aromatic compounds that can be utilized as a carbon source are funneled into one of seven known dihydroxylated aromatic intermediates (Diaz et al., 2013; Lubbers et al., 2019). The most common of the seven common aromatic intermediates are 3,4-dihydroxybenzoic acid (3,4-DHB), also known as protocatechuic acid (PCA, as seen in Figure 2), and catechol (Harwood and Parales, 1996; Marín et al., 2010). Hydroxyquinol, gentisic acid, gallic acid, hydroquinone, and pyrogallol are also central intermediates used by fungi and bacteria (Lubbers et al., 2019). An example of this funneling is shown in Figure 2.

Following conversion to one of the central intermediates, the compounds then undergo ring fission, which reduces or eliminates the toxicity, followed by conversion into intermediates

which enter the TCA cycle (Figure 2). Ring fission is catalyzed by specific metal ion-dependent dioxygenases that insert both atoms of an O_2 molecule into these central intermediates, all of which are hydroxylated at two positions either *ortho* or *para* to each other (Vaillancourt et al. 2006). The central pathways for catabolism of plant-derived aromatic compounds appear to be widely conserved among fungi and bacteria, although the fungal pathways are not quite identical to the bacterial pathways (Harwood & Parales, 1996; Lubbers, 2019). Figure 3 shows the current state of knowledge of these central pathways in bacteria, yeast, and filamentous fungi. The seven common intermediates are indicated by the letters A to G in Figure 3. As shown by the lack of blue boxes in this figure, the enzymes or genes involved in the catabolism of five of the seven compounds are unknown in filamentous fungi.

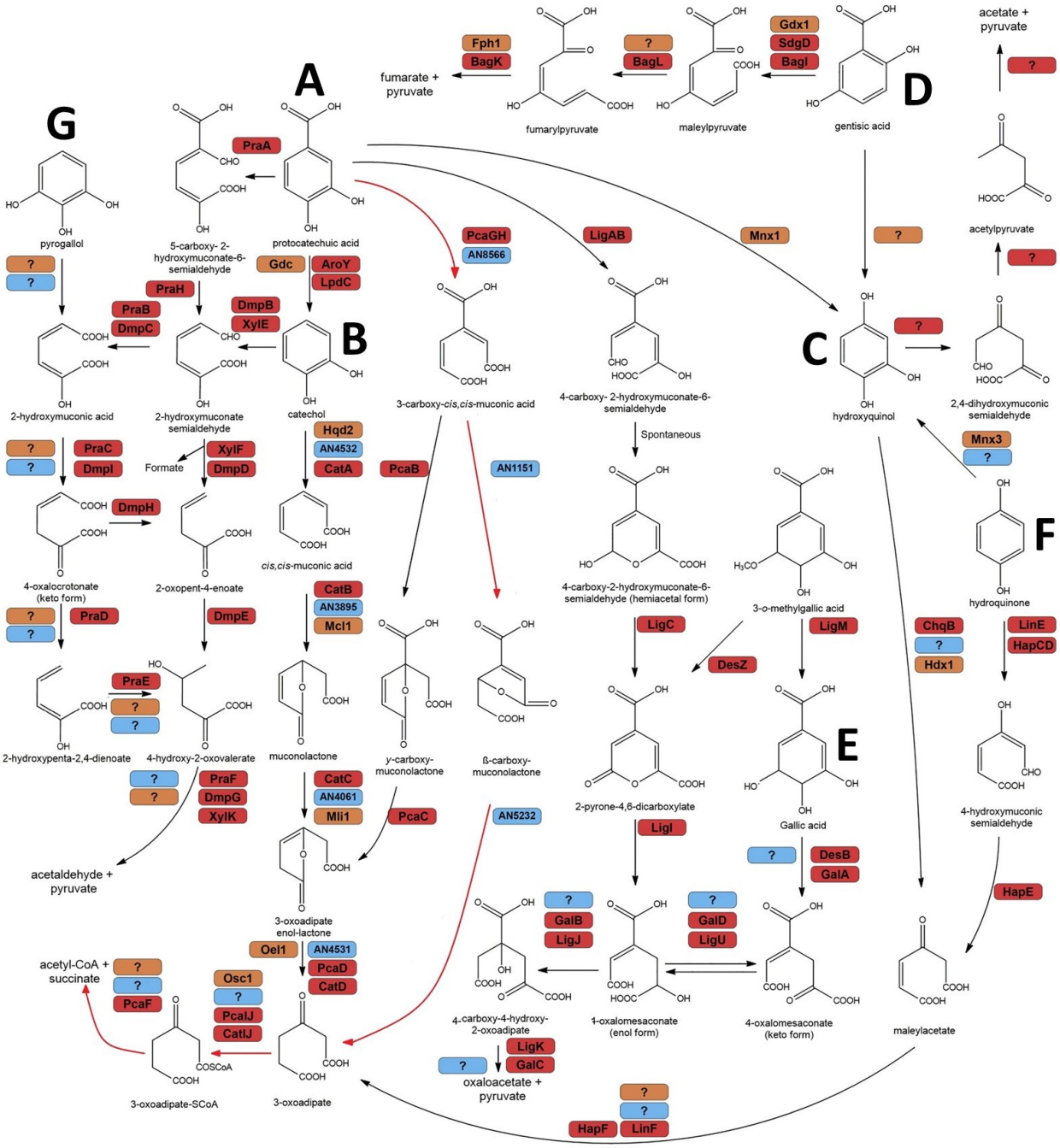

Figure 3. Central pathways of aromatic catabolism in filamentous fungi, bacteria, and yeast. The 3,4-DHB pathway in filamentous fungi is indicated by the red arrows. Enzymes or genes involved in each step are indicated in the boxes when known, in red for bacteria, orange for yeast, and blue for filamentous fungi. The seven common intermediates are indicated by the letters A to G. Modified from Lubbers et al. (2019).

1.3.1 Catabolism of quinic acid

As a common plant aromatic compound, quinic acid is utilized as a carbon source by a wide variety of fungi and bacteria (Grund and Kutzner, 1998; Niu et al, 2016). The initial steps of quinic acid catabolism convert quinic acid into the common intermediate 3,4-DHB. The genes involved in these steps of the pathway have been studied for decades in the filamentous fungus *Neurospora crassa*. Like many of the aromatic catabolic pathways leading to the central intermediates, the genes were found to be clustered. This gene cluster was among the first to be discovered in eukaryotes, and the first to include a regulatory gene (Valone et al., 1971). This cluster, called the *qa* gene cluster in *Neurospora crassa*, includes seven genes. Three of these genes encode the enzymes quinate dehydrogenase, which converts quinic acid to dehydroquinic acid; dehydroquinate dehydratase, which converts dehydroquinic acid to dehydroshikimic acid; and dehydroshikimate dehydratase, which converts dehydroshikimic acid to 3,4-dihydroxybenzoic acid. Another two genes encode an activator and repressor controlling expression of the cluster, and two more encode a quinate permease and a protein of unknown function (Giles et al., 1991; Huiet, 1984). The *qa* gene cluster is conserved in some fungi and has also been studied in the fungus *Aspergillus nidulans*, where it is known as the *QUT* cluster. The *QUT* cluster contains orthologs of the seven *qa* genes, although in a slightly different order, and contains an eighth gene with unknown function (Giles et al., 1991; Lamb et al., 1992). Orthologs of the *QUT* cluster have also been identified in *Aspergillus niger* although these genes have not been characterized (Niu et al, 2016).

Catabolism of the related compound shikimic acid occurs via the shikimate pathway. There are seven steps in the prechorismate section of this pathway, five of which are catalyzed by the pentafunctional AROM polypeptide (Lamb et al., 1991). One of these five reactions is identical to a reaction of the quinic acid pathway: conversion of dehydroquinic acid to dehydroshikimic acid.

The quinate permease gene encodes a quinate:H+ symporter in the major facilitator superfamily of transporters (Pao et al., 1998). Whittington et al. (1987) showed that this permease in *Aspergillus nidulans*, QutD, is required for the transport of quinic acid into the mycelium at pH

6.5. Knockout mutants lacking QutD were able to grow well in media with quinic acid as the sole carbon source at pH of 3.5 but failed to grow at pH 6.5. This was hypothesized to be due to the protonated quinic acid molecule, with an apparent pKa of 3.6, being able to enter the mycelium by diffusion at low pH, while the deprotonated quinate molecule present at the higher pH requires the permease for transport into the mycelium.

1.3.2 Central catabolic pathways of aromatic compounds

Three of the central intermediates, protocatechuic acid (3,4-DHB), catechol, and hydroxyquinol, are degraded by pathways which converge at β-ketoadipate (3-oxoadipate) before further degradation to TCA cycle intermediates (Figure 3). The 3,4-DHB and catechol branches of this pathway are known as the β-ketoadipate pathway, which uses analogous sets of enzymes and intermediates before converging on β-ketoadipate (Wright, 1993; Harwood and Parales, 1996). In bacteria these central pathways, and the β-ketoadipate pathway in particular, have been well characterized, as suggested by the identification of an enzyme at nearly every step in Figure 3. Studies of bacteria have characterized the biochemistry, enzymology, genetics, and regulation of these pathways. However, in fungi, particularly the filamentous fungi, these pathways remain poorly understood (reviewed in Harwood and Parales, 1996; Wright, 1993; Lubbers et al., 2019). This is also suggested by Figure 3, where enzymes have not been assigned to most steps.

The 3,4-DHB branch of this pathway is one of the most well understood in fungi. It comprises three enzymes converting 3,4-DHB to β-ketoadipate: protocatechuate 3,4-dioxygenase, which cleaves 3,4-DHB to form 3-carboxy-*cis,cis*-muconic acid; carboxy-*cis,cis*-muconate cyclase, which converts 3-carboxy-*cis,cis*-muconic acid into 3-carboxymuconolactone; and 3-carboxymuconolactone hydrolase, which converts 3-carboxymuconolactone into β-ketoadipate (Cain et al 1968). Two further reactions convert β-ketoadipate into β-ketoadipyl-CoA then succinyl-CoA and acetyl-CoA, catalyzed by β-ketoadipate-CoA transferase and β-ketoadipyl-CoA thiolase, respectively (Katagiri and Hayaishi, 1957; Canovas and Stanier, 1967). These enzymatic reactions are indicated by the red arrows in Figure 3. The 3,4-DHB catabolic pathway in bacteria is slightly different than in bacteria. In bacteria, 3-carboxy-*cis,cis*-muconic acid is converted to 4-carboxymuconolactone rather than 3-carboxymuconolactone. Then 4-

carboxymuconolactone is converted to 3-oxoadipate-enol-lactone, where the convergence with the catechol branch occurs (Figure 3). The 3-oxoadipate-enol-lactone is converted into 3-oxoadipate, where the fungal and bacterial pathways once again converge.

One area of knowledge that is especially lacking in fungi is the genes involved in these pathways. Previous studies, although often spread across different fungal species, have identified many of the intermediates and enzymatic activities involved in these pathways. As mentioned, Cain et al., in a number of papers (1968; Thatcher and Cain, 1970; 1974; 1975), identified the intermediates and enzymatic activities involved in the 3,4-DHB branch of the β-ketoadipate pathway in *Aspergillus niger* and several other fungal species decades ago. However, although the genes involved in some of the funneling pathways like the quinic acid catabolic pathway have been identified in fungi, the genes involved in these central pathways remain unknown in most cases, and the enzymes uncharacterized.

Recent studies have begun to examine the genes of these pathways. Martins et al. (2015), determined some of the genes involved in the β-ketoadipate pathway in *Aspergillus nidulans.* Another study in the yeast *Candida parapsilosis* showed several of the genes involved in gentisic acid and hydroxyquinol metabolism (Holesova et al., 2011). However, there remains much work to be done in this area.

1.4 *Aspergillus niger* as a model organism

Aspergillus niger is an effective model organism to study the genetics of aromatic compound catabolism for a number of reasons. *Aspergillus niger* is a filamentous ascomycete fungus that is ubiquitous in the environment (Baker, 2006). As a soil fungus, it requires the ability to degrade a wide range of aromatic compounds produced by decaying plant matter, and is known to be able to utilize many different monocyclic aromatic compounds as a sole carbon source (Shailubhai et al., 1983a; Semana and Powlowski, 2019; Milstein et al., 1988; Lubbers et al., 2019; Tsang lab unpublished data). This includes utilization of 3,4-DHB, catechol, and hydroxyquinol.

Aspergillus niger is also an important species both industrially, for the production of enzymes and organic acids, and in studies of fungal physiology, biochemistry, biotechnology, metabolism, and genetics. Industrially, most of the global supply of citric acid is produced by fermentation using *A. niger* (Tong et al., 2019). It is also an important source of enzymes involved in degradation of plant polysaccharides (Hu et al., 2011). Along with *Aspergillus nidulans, A. niger* is one of the most common organisms for filamentous fungi research. It has the advantage of being genetically tractable, and having tools for genetic manipulation, including by CRISPR-Cas9, already developed (Song and Ouedraogo et al., 2018). *Aspergillus niger* also has several high-quality genome sequences of available (Pel et al., 2007; Andersen et al., 2011). The *A. niger* strain CBS 138852 is a descendant of the NRRL3 strain. The genome of NRRL3 has been manually curated and annotated (http://gbrowse.fungalgenomics.ca/cgi-bin/gb2/gbrowse/Aspni_nrrl3_public/).

Aspergillus niger strain CBS 138852 contains several mutations which are useful for studies involving genetic manipulation. It is derived from a mutant, N402 (*cspA1)*, itself derived from the NRRL3 strain, with dense sporulation and much shorter conidiophores than the wild type, which makes it easier to work with on solid media (Bos et al., 1988). The N593 strain was derived from N402 and also carries a knockout mutation in the *pyrG* gene, which encodes orotidine-5'-phosphate decarboxylase, an enzyme involved in uridine biosynthesis. Mutants lacking expression of this enzyme are unable to produce uridine and become auxotrophic for uridine and uracil (Goosen et al., 1987). This gives N593 a simple and very effective selectable marker by the inclusion of the *pyrG* gene on plasmids used in transformation of this strain. When plated on media lacking uracil and uridine, only successful transformants will be able to grow. Finally, CBS 138852 was derived from N593 and carries a knockout mutation in the *kusA* gene. This gene is involved in the non-homologous end joining pathway, which repairs double stranded breaks in DNA without the need of a rescue template. With this mutation, DNA can only be repaired by the homology directed repair pathway, greatly increasing the efficiency of homologous recombination (Meyer et al., 2007).

1.5 CRISPR-Cas9

CRISPR-Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats / CRISPR associated protein 9) was discovered as an adaptive immune system used by some bacteria for acquired resistance against bacteriophages by uptake of invasive DNA (Barrangou et al., 2007). In recent years, the CRISPR-Cas9 system has been adapted for use as a method of simple, precise and efficient genome editing in a wide variety of species, including filamentous fungi like *Aspergillus niger* (Jiang et al., 2013; Nødvig et al., 2015; Song and Ouedraogo et al., 2018) .

The CRISPR-Cas9 system involves two main components: the Cas9 protein, an endonuclease that cleaves DNA causing double stranded breaks, and a guide RNA (gRNA). The gRNA itself has two components, the protospacer, or CRISPR RNA (crRNA), which is used for targeting specific sequences and is at the 5' end of an RNA sequence which associates with the Cas9 protein (Nødvig et al., 2015; Song and Ouedraogo et al., 2018). The 20 base pair crRNA can be used to specifically direct Cas9 to create a double stranded break at any location in the genome adjacent to a Protospacer Adjacent Motif (PAM) site, which has the sequence NGG (Jiang et al., 2013). As this is a common sequence, it is generally easy to find target sites in throughout the genome. However, not all crRNA sequences are equally efficient (see section 2.3).

1.6 Rationale for the study of aromatic compound catabolism in *A niger*

As previously mentioned, there is a large gap in the fundamental understanding of not only the 3,4-DHB pathway, but all central catabolic pathways for plant-derived aromatic compounds in fungi. Very few genes have been assigned in these pathways in any fungi, and even less so in filamentous fungi compared to yeast. The vast majority of the enzymes involved have also not been characterized. Additionally, the intermediates and enzymes involved in some of these pathways remain unknown in filamentous fungi. For example, although 3,4-DHB and 2,3-DHB are known to be degraded through the β-ketoadipate pathway in fungi (the catechol branch is used to degrade 2,3-DHB), the closely related compound 2,4-DHB also appears to be degraded through an unknown pathway that converges at β-ketoadipate (unpublished data).

As 3,4-DHB is one of the most common intermediates, this study begins with an examination of this pathway in *A. niger*. However, using the same approach for the other central pathways can reveal the entire network of pathways and the genes and enzymes involved for the first time. *Aspergillus niger* is not only a useful model organism for the reasons mentioned above, but the high-quality genome assembly and annotation will allow the acquired information about these genes to be easily applied to other fungi.

There are several potentially valuable applications of this research. As previously mentioned, many aromatic compounds are toxic to microorganisms even at low concentrations. This toxicity inhibits fermentation when lignocellulosic biomass is used as a feedstock (Larsson et al., 2000, Adeboye et al., 2014). Understanding the pathways which degrade these compounds can lead to improved tolerance, or heterologous expression of these pathways in organisms used for fermentation to improve yields. Toxic aromatic compounds, both mono- and polycyclic, are also present in soil and water and many have negative effects on human health including carcinogenic, mutagenic, and teratogenic effects (Cao et al., 2009; Singh and Jain, 2003). Fungi have shown the ability to degrade many of these toxic compounds. Many toxic chlorinated aromatic compounds, introduced into the environment from pesticide use or industrial waste, can be degraded by *A. niger* through the β-ketoadipate pathway (Sahasrabudhe et al., 1985; Shailubhai et al., 1983b). *Aspergillus niger* has also been shown to degrade benzene, toluene, ethylbenzene and xylene, toxic aromatic pollutants that form part of petroleum (Perdana et al., 2019). To efficiently use fungi for bioremediation of toxic aromatic compounds, it will be necessary to fully understand the catabolic pathways involved in degrading them.

As mentioned above, there are also many industrial uses for aromatic compounds derived from lignin that are potential targets for lignin valorization. The funneling of lignin-derived aromatic compounds into central pathways provides a natural solution to the problem of the heterogeneity of lignin as a feedstock (Beckham et al., 2016). Many intermediates of these pathways have industrial value. For example, *cis,cis*-muconic acid in the catechol branch of the pathway is valuable as an intermediate for several chemicals, primarily adipic acid which is used in the production of nylon and other polymers (Vardon et al., 2015). Adipic acid can also be produced from another intermediate of the β-ketoadipate pathway, β-ketoadipyl-CoA (Kallscheuer et al.,

2017). Other potentially valuable intermediates of the central pathways include 2-pyrone-4,6-dicarboxylic acid (PDC), pyridine 2,4-dicarboxylic acid, pyridine 2,5-dicarboxylic acid, muconolactone, 2-hydroxymuconic semialdehyde, and 3-carboxymuconate (Beckham et al., 2016). Quinic and shikimic acid may also be used as starting molecules for synthesis of antiviral drugs (Nie and Shi 2009).

1.7 Specific goals of this study

The main goals of this study are to identify and characterize the genes involved in each step of the catabolism of quinic acid and 3,4-DHB in *A. niger*. These goals require several stages of research. First, identifying potential genes involved in this pathway using transcriptomics, gene annotation, and where possible, homology to other genes including those in the bacterial 3,4-DHB catabolic pathway. Second, using CRISPR-Cas9 to create knockout mutants for the genes identified as the most likely to encode each enzyme. Thirdly, growth phenotype of these mutants using 3,4-DHB or quinic acid as a sole carbon source can be used to determine involvement in the pathway. Metabolic analysis by liquid chromatography-mass spectrometry using 3,4-DHB or quinic acid as a substrate can provide further evidence and confirm the correct identification of the genes encoding each enzyme by accumulation of intermediates.

2. Materials and Methods

2.1 *Aspergillus niger* strains

The parental strain used in this study was CBS 138852 (Kun et al., 2020). This strain is a derivative of *A. niger* N593, which is auxotrophic for uridine (Goosen et al., 1987). CBS 138852 carries an additional deletion in the *kusA* gene, which increases the efficiency of homologous recombination by preventing nonhomologous end joining DNA repair (Meyer et al., 2007). Various *A. niger* mutant strains were generated in this study and are listed in Table 1. *Escherichia coli* strain DH5α was used for the maintenance and propagation of cloned plasmids.

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

Strain	Genotype	Source
CBS 138852	*cspA1*, Δ*pyrG*, Δ*kusA*	Kun et al. (2020)
MS1	CBS 138852 ΔNRRL3_08520	This study
MS2	CBS 138852 ΔNRRL3_11037	This study
MS3	CBS 138852 ΔNRRL3_11035	This study
MS4	CBS 138852 ΔNRRL3_11036	This study
MS5	CBS 138852 ΔNRRL3_05631	This study
MS6	CBS 138852 ΔNRRL3_11036, ΔNRRL3_05631	This study
MS7	CBS 138852 ΔNRRL3_01405	This study
MS8	CBS 138852 ΔNRRL3_02586	This study
MS9	CBS 138852 ΔNRRL3_00837	This study
MS10	CBS 138852 ΔNRRL3_01409	This study
MS11	CBS 138852 ΔNRRL3_00837, ΔNRRL3_01409	This study
MS12	CBS 138852 ΔNRRL3_08340	This study
MS13	CBS 138852 ΔNRRL3_01886	This study
MS14	CBS 138852 ΔNRRL3_01526	This study

2.2 Culture conditions

Table 2 shows the media and solutions used in this study. Following transformation of CRISPR plasmids and gene-editing oligonucleotides into *A. niger*, protoplasts were plated on selective

regeneration media (SRM) plates lacking uridine or uracil (Hill and Kafer, 2001) and grown for up to 5 days at 30°C.

Screening for successful deletions by PCR was performed using spores from single colonies of the transformant plate inoculated into 250 µL of complete media (CM: minimal media supplemented with 5 g/L yeast extract, 1 g/L casamino acids, and 10mM uridine) and grown at 30°C for 17 hours prior to DNA extraction.

For transcriptomics, cultures of *A. niger* CBS 138852 were grown using spores at a concentration of $2x10^6$ spores/mL in CM for 24 hours at 37°C, shaking at 220 rpm. Mycelia were then collected by gravity filtration using Miracloth (MilliporeSigma, Burlington, MA), washed with MM lacking any carbon source and transferred into 50 mL of either MM + 0.5% fructose or MM + 0.5% 3,4-DHB. These cultures were incubated at 32°C, 220 rpm for 2 hours prior to collection of mycelia as above for RNA extraction.

For miniprep of plasmids, single *E. coli* colonies were grown in 5 mL of LB medium with ampicillin overnight at 37°C, shaking at 220 rpm.

Fungal colonies for phenotypic testing were grown by spotting 1000 fresh spores in a total volume of 2 µL of saline/Tween onto minimal media plates with 0.5% of either fructose or a DHB as the sole carbon source.

For mass spectrometry, 200 mL of complete media was inoculated with fresh spores to a concentration of $2x10^6$ spores/mL. Mycelia collected by gravity filtration using Miracloth, washed with minimal media with no carbon source, dried with paper towel around the Miracloth, and 250 mg were transferred to 25 mL of minimal media with quinic acid or 3,4-DHB as the sole carbon source and grown at 30°C for up to 120 hours, shaking at 220 rpm.

Table 2 media and solutions used in this study

Minimal Media (MM)	As described by Käfer, E., (1977). 5 g/L carbon source (0.5%) used unless

	otherwise noted, supplemented with 10 mL of 100 mM uridine solution and 15 g of agar for plates
Complete Media (CM)	MM supplemented with 10 g/L glucose, 5 g/L yeast extract, 2 g/L peptone, 1 g/L casamino acids, and 15 g of agar if plates
DNA Extraction Lysis Buffer	20 mM Tris-HCL, pH 8.0; 400 mM NaCl, 0.3% SDS
Tris-EDTA (TE) Buffer	10 mM Tris-Cl, pH 8.0; 1 mM EDTA, pH 8.0
Selective Regeneration Media (SRM)	MM lacking uridine or uracil and supplemented with 30 g/L sucrose and 15 g/L agar
LB media (Bertani, G.,1951)	10 g bacto-tryptone, 5 g YE, 10 g NaCl, pH 7.0
Saline Tween	0.5% NaCl, 0.002% Tween 80
Sorbitol MES buffer Calcium Chloride (SMC) buffer	1.33 M sorbitol, 50 mM $CaCl_2$, and 20 mM MES buffer, pH 5.8
Sorbitol Tris Calcium Chloride (STC) buffer	1.33 M sorbitol, 50 mM $CaCl_2$, and 10 mM Tris-HCl, pH 7.5
Tris Calcium Chloride (TC) Buffer	50 mM $CaCl_2$, and 10 mM Tris-HCl, pH 7.5
Polyethylene glycol (PEG)-TC buffer	25% PEG-6000 in TC Buffer

2.3 Plasmid construction

Mutants were generated using CRISPR-Cas9 by creating deletions in target genes as previously described (Song and Ouedraogo et al., 2018). CRISPR-Cas9 guide RNA sequences were identified using Geneious R9.1 (http://www.geneious.com, Kearse et al., 2012). Possible guide RNA sequences were evaluated based on 3 criteria:

1. Doench activity score (Doench et al., 2014), which predicts the efficiency of gRNA binding and Cas9 cleavage. Higher scores were prioritized.
2. Number of off-target sites. Sequences with zero off-target sites were chosen.
3. Location of the cut site, prioritizing sequences near the 5' end of the gene to cause frameshift mutations after creating deletions.

Sequences chosen as the CRISPR target sequences are shown in Table 3.

The plasmid used for generating the knockout mutants used in this project was the previously designed ANEp8-Cas9-LIC (Song and Ouedraogo et al., 2018), shown in Figure 4A. This plasmid is 15,569 base pairs in size and expresses Cas9, *bla* (beta-lactamase, conferring resistance to ampicillin for bacterial selection), and *pyrG* (orotidine-5'-phosphate decarboxylase, a gene deleted from CBS 138852 to cause uridine auxotrophy). The plasmid also contains AMA1, a sequence that allows extrachromosomal plasmid replication (Aleksenko and Clutterbuck, 1996) and a 38 base pair ligation independent cloning (LIC) site containing a SwaI restriction site in the center. The LIC site is used for inserting the guide RNA cassette (Figure 4B) into the plasmid following the LIC method (Aslanidis and De Jong, 1990).

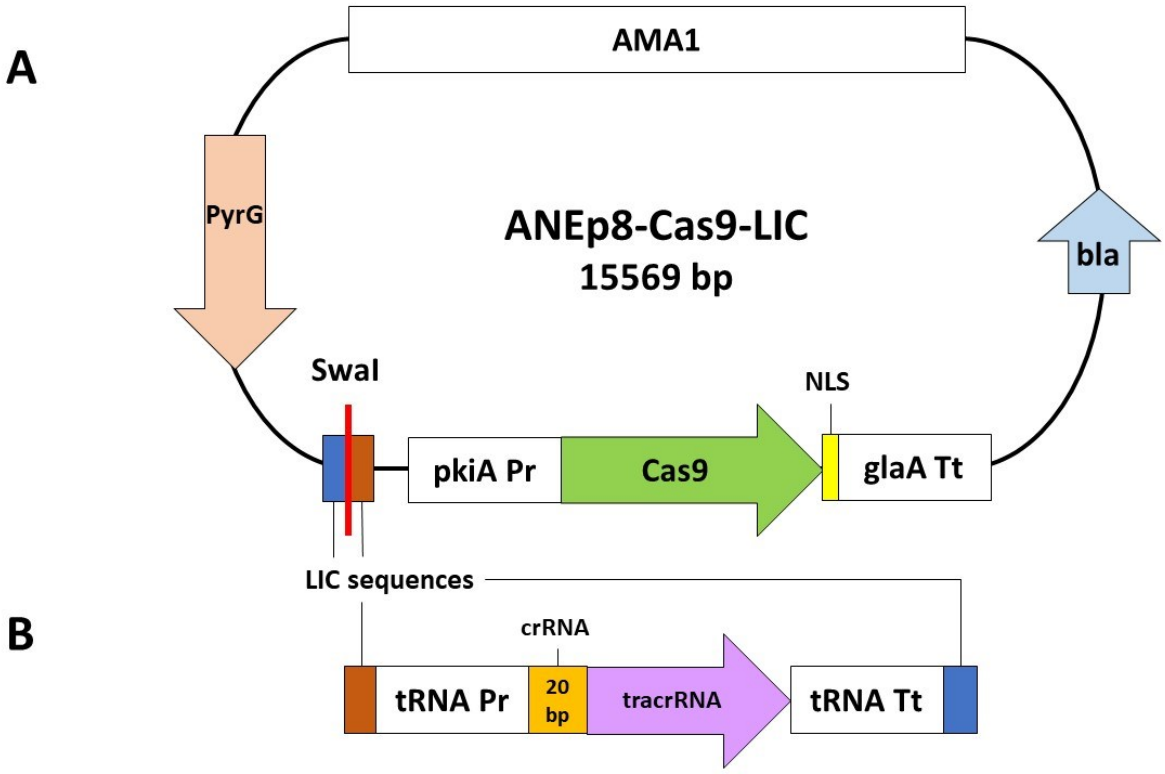

Figure 4 Map of CRISPR-Cas9 plasmid ANEp8-Cas9-LIC (A) and gRNA cassette (B) to be inserted into the plasmid by LIC. Abbreviations: bla, beta-lactamase; *pkiA* Pr, pyruvate kinase gene promoter; *glaA* Tt, glucoamylase gene terminator; NLS, nuclear localization signal; tRNA Pr, $tRNA^{pro1}$ gene promoter; tRNA Tt, $tRNA^{pro1}$ gene terminator; crRNA, CRISPR RNA; tracrRNA, trans-activating crRNA.

Guide RNA expression cassettes were created using primers listed in Table 3. These cassettes express the two components of the gRNA: the 20 base pair crRNA used to target the CRISPR cut site, and the 80 base pair trans-activating CRISPR RNA (tracrRNA) sequence required for cas9 binding (Sander and Joung, 2014). For all genes to be deleted, four primers were used to create two fragments of the gRNA expression cassette by PCR using a previously constructed ANEp8-Cas9-LIC plasmid as the template. These fragments were then combined using fusion PCR as shown in Figure 5. The first fragment was created using the $tRNA^{pro1}$ LIC Fw, which is the forward primer binding to the 5' end of promoter and containing an LIC adapter sequence, and the reverse primer binding to the 3' end of the promoter and containing a tail with the 20 bp sequence of the CRISPR target site. The second fragment was created using a forward primer binding to the 5' end of the tracrRNA (or gRNA scaffold) and containing a complementary 20 bp

CRISPR target site tail for binding the two fragments through fusion PCR. The reverse primer for this fragment was tRNA^{pro1} LIC Rv, which binds to the 3' end of the terminator and contains an LIC adapter sequence.

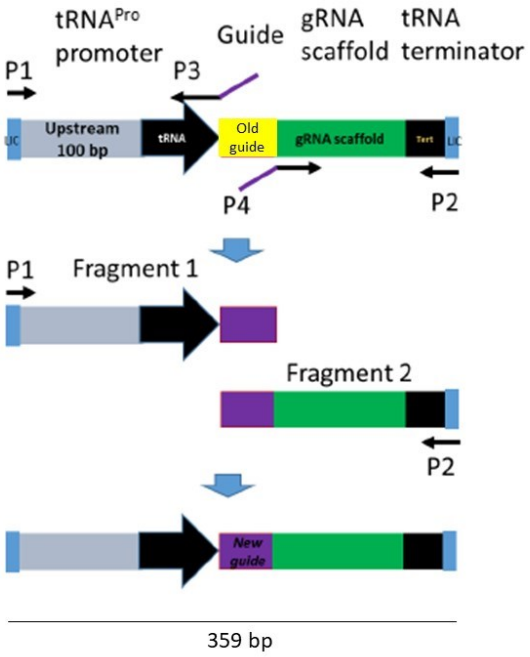

Figure 5. Diagram of PCR used to create the gRNA cassette. Modified from Song and Tsang (2017).

Competent cells of the DH5α strain of *Escherichia coli* were transformed with cloned plasmids as described (New England Biolabs, n.d.) and plated on LB plates containing ampicillin for selection. Resulting colonies were screened for successful LIC, i.e. presence for the gRNA cassette insert, by colony PCR. Colonies were picked, placed in 20 µL ddH_2O, and vortexed. From the colony suspension, 1 µL was used as the template for PCR. If successful, the remaining 19 µL was used for inoculation of LB media in preparation for miniprep. Primers (LIC screen Fw and Rv; Table 3) flanking the insertion site were used for the PCR. If the 359 bp insertion was present, the fragment resulting from the PCR would be 490 bp. In the absence of the insertion, the fragment would be 131 bp.

Table 3. Primers and oligonucleotides used in this study.

Name	Sequence	Description
$tRNA^{pro1}$ LIC Fw	**CAACCTCCAATCCAATTTG**ACTCCG CCGAACGTACTG	Forward primer for tRNA promoter, LIC tail in bold
$tRNA^{pro1}$ LIC Rv	**ACTACTCTACCACTATTTG**AAAAGC AAAAAAGGAAGGTACAAAAAAGC	Reverse primer for tRNA terminator, LIC tail in bold
NRRL3_08520 Fw	GCACCTGGACGACCTTACCGAGG GTTTTAGAGCTAGAAATAGCAAG	Forward primer for gRNA scaffold, gRNA sequence in bold
NRRL3_08520 Rv	CCTCGGTAAGGTCGTCCAGGTGC GACGAGCTTACTCGTTTCG	Reverse primer for tRNA promoter, gRNA sequence in bold
NRRL3_11037 Fw	TCGACCGCATCCACGCCGCG GTTTTAGAGCTAGAAATAGCAAG	Forward primer for gRNA scaffold, gRNA sequence in bold
NRRL3_11037 Rv	CGCGGCGTGGATGCGGTCGA GACGAGCTTACTCGTTTCG	Reverse primer for tRNA promoter, gRNA sequence in bold
NRRL3_11035 Fw	GCAACCCTTCTCCTTCTACG GTTTTAGAGCTAGAAATAGCAAG	Forward primer for gRNA scaffold, gRNA sequence in bold
NRRL3_11035 Rv	CGTAGAAGGAGAAGGGTTGC GACGAGCTTACTCGTTTCG	Reverse primer for tRNA promoter, gRNA sequence in bold
NRRL3_11036 Fw	**ACAAGAGGTGAAGGAGGCAA**GTTTT AGAGCTAGAAATAGCAAG	Forward primer for gRNA scaffold, gRNA sequence in bold
NRRL3_11036 Rv	**TTGCCTCCTTCACCTCTTGT**GACGA GCTTACTCGTTTCG	Reverse primer for tRNA promoter, gRNA sequence in bold
NRRL3_05631 Fw	**GAGTCGGACGATCCTCAACG**GTTTT AGAGCTAGAAATAGCAAG	Forward primer for gRNA scaffold, gRNA sequence in bold
NRRL3_05631 Rv	**CGTTGAGGATCGTCCGACTC**GACGA GCTTACTCGTTTCG	Reverse primer for tRNA promoter, gRNA sequence in bold
NRRL3_01405 Fw	**AGATACGGTGGCACTCGTTA**GTTTT AGAGCTAGAAATAGCAAG	Forward primer for gRNA scaffold, gRNA sequence in bold
NRRL3_01405 Rv	**TAACGAGTGCCACCGTATCT**GACGA GCTTACTCGTTTCG	Reverse primer for tRNA promoter, gRNA sequence in bold

NRRL3_02586 Fw	**CAAGAAGACCGACATCCCCG**GTTTTAGAGCTAGAAATAGCAAG	Forward primer for gRNA scaffold, gRNA sequence in bold
NRRL3_02586 Rv	**CGGGGATGTCGGTCTTCTTG**GACGAGCTTACTCGTTTCG	Reverse primer for tRNA promoter, gRNA sequence in bold
NRRL3_00837 Fw	**ACGCAAACAATCATCGACGG**GTTTTAGAGCTAGAAATAGCAAG	Forward primer for gRNA scaffold, gRNA sequence in bold
NRRL3_00837 Rv	**CCGTCGATGATTGTTTGCGT**GACGAGCTTACTCGTTTCG	Reverse primer for tRNA promoter, gRNA sequence in bold
NRRL3_01409 Fw	**GTACCAGTCGTGGAATTGAG**GTTTTAGAGCTAGAAATAGCAAG	Forward primer for gRNA scaffold, gRNA sequence in bold
NRRL3_01409 Rv	**CTCAATTCCACGACTGGTAC**GACGAGCTTACTCGTTTCG	Reverse primer for tRNA promoter, gRNA sequence in bold
NRRL3_01886 Fw	**TCAGATCCTTGATGTCCGGG**GTTTTAGAGCTAGAAATAGCAAG	Forward primer for gRNA scaffold, gRNA sequence in bold
NRRL3_01886 Rv	**CCCGGACATCAAGGATCTGA**GACGAGCTTACTCGTTTCG	Reverse primer for tRNA promoter, gRNA sequence in bold
NRRL3_01526 Fw	**GGTCCTCCAGAAATCTCCCT**GTTTTAGAGCTAGAAATAGCAAG	Forward primer for gRNA scaffold, gRNA sequence in bold
NRRL3_01526 Rv	**AGGGAGATTTCTGGAGGACC**GACGAGCTTACTCGTTTCG	Reverse primer for tRNA promoter, gRNA sequence in bold
NRRL3_08340 Fw	**AGGTATCGCCAACGACGGCG**GTTTTAGAGCTAGAAATAGCAAG	Forward primer for gRNA scaffold, gRNA sequence in bold
NRRL3_08340 Rv	**CGCCGTCGTTGGCGATACCT**GACGAGCTTACTCGTTTCG	Reverse primer for tRNA promoter, gRNA sequence in bold
LIC Screen Fw	TTTTCTCTTCCATTTACGC	Forward primer to screen for successful LIC
LIC Screen Rv	GGGGATCATAATAGTACTAGCCA	Reverse primer to screen for successful LIC
GE-Oligo NRRL3_11037	**ACCACGAGGGCGCCATCGTCGACCG CATCC**GGGCTCCGTCGACGCAATCATCATCAACCC	Gene-editing oligo to create deletion in target gene between bold and unbold sequences

GE-Oligo NRRL3_11035	**CCTCACCATCATCTGCCTGCAACCC TTCTC**GGGCTCGTCGACCGCAAGCA AACCGAGTAT	Gene-editing oligo to create deletion in target gene between bold and unbold sequences
GE-Oligo NRRL3_11036	**CTACAACTGGCGTATCTACCTCCTG GCGGC**TTCACCTCTTGTATGATCGG CTACGACAGT	Gene-editing oligo to create deletion in target gene between bold and unbold sequences
GE-Oligo NRRL3_05631	**TTCTCCCATACAATGGGTGGAATTC TTCAC**ATCGTCCGACTCCCAAAAAT GTCTACAATT	Gene-editing oligo to create deletion in target gene between bold and unbold sequences
GE-Oligo NRRL3_01405	**AACTCCATTGGAAAGATCAGCACTC CCATC**GCCACCGTATCTGCGATGTT ATCGGTCTGG	Gene-editing oligo to create deletion in target gene between bold and unbold sequences
GE-Oligo NRRL3_02586	**CCTCACCCTCGACCTGGTCAAGAAG ACCGA**GCTGAGCCCATCTCCTGGAT GACCTTTTCG	Gene-editing oligo to create deletion in target gene between bold and unbold sequences
GE-Oligo NRRL3_00837	**TTCCTGGGCGCCATCCGCACGCAAA CAATC**GATTGCTCGAGCTCGCCGTG TGTCGCGTGG	Gene-editing oligo to create deletion in target gene between bold and unbold sequences
GE-Oligo NRRL3_01409	**TTCTCTGGGTGACGATGAACCCTCG TCCCT**CAACACCGAACACGGCCCTC TGCGCCTGCG	Gene-editing oligo to create deletion in target gene between bold and unbold sequences
GE-Oligo NRRL3_01886	**CCTACTACCTTGATTGATGCCGTGC GTGAC**TCAAGGATCTGACTGTCGTC TCCAACAATG	Gene-editing oligo to create deletion in target gene
GE-Oligo NRRL3_01526	**CCTCTCCAATCCCCCGGGGCCTGCG CCAGG**TCTCTCTTCCCTCCGTACCC CGGTCACCCG	Gene-editing oligo to create deletion in target gene between bold and unbold sequences
GE-Oligo NRRL3_08340	**GCTCCACGAAACCCTCTTCGACGAG GGTAT**TACGTCGATAGAGCCTTGGC TCAAGGATCC	Gene-editing oligo to create deletion in target gene

NRRL3_11037 Check Fw	ACGAGGGCGCCATCGTC	Forward primer to verify deletion by PCR
NRRL3_11037 Check Rv	GGGTTGATGATGATTGCGTC	Reverse primer to verify deletion by PCR
NRRL3_11035 Check Fw	AACTCCATTGGAAAGATCAG	Forward primer to verify deletion by PCR
NRRL3_11035 Check Rv	AGACCGATAACATCGCAG	Reverse primer to verify deletion by PCR
NRRL3_11036 Check Fw	CTACAACTGGCGTATCTACC	Forward primer to verify deletion by PCR
NRRL3_11036 Check Rv	ACTGTCGTAGCCGATCATAC	Reverse primer to verify deletion by PCR
NRRL3_05631 Check Fw	TTCTCCCATACAATGGGT	Forward primer to verify deletion by PCR
NRRL3_05631 Check Rv	AATTGTAGACATTTTTGGGA	Reverse primer to verify deletion by PCR
NRRL3_01405 Check Fw	AACTCCATTGGAAAGATCAG	Forward primer to verify deletion by PCR
NRRL3_01405 Check Rv	AGACCGATAACATCGCAG	Reverse primer to verify deletion by PCR
NRRL3_02586 Check Fw	CTCACCCTCGACCTGGTCA	Forward primer to verify deletion by PCR
NRRL3_02586 Check Rv	CGAAAAGGTCATCCAGGAGA	Reverse primer to verify deletion by PCR
NRRL3_00837 Check Fw	AACAGCTTCCTGGGCGCCAT	Forward primer to verify deletion by PCR
NRRL3_00837 Check Rv	ACGCGACACACGGCGAG	Reverse primer to verify deletion by PCR
NRRL3_01409 Check Fw	CGATGAACCCTCGTCCCT	Forward primer to verify deletion by PCR
NRRL3_01409 Check Rv	AGGGCCGTGTTCGGTGTT	Reverse primer to verify deletion by PCR
NRRL3_01886 Check Fw	CCTACTACCTTGATTGATGC	Forward primer to verify deletion by PCR
NRRL3_01886 Check Rv	CATTGTTGGAGACGACAGTC	Reverse primer to verify deletion by PCR
NRRL3_01526 Check Fw	GGCCTCTCCAATCCCC	Forward primer to verify deletion by PCR
NRRL3_01526 Check Rv	GGTACGGAGGGAAGAGAGAA	Reverse primer to verify deletion by PCR
NRRL3_08340 Check Fw	TTGACCAGCTCCACGAAA	Forward primer to verify deletion by PCR
NRRL3_08340 Check Rv	CTGGATCCTTGAGCCAAG	Reverse primer to verify deletion by PCR

NRRL3_08520 5′ flank Fw	TGAGCTCCTCATGCCTGCTC	Forward primer for 5′ flank of homologous recombination oligo
NRRL3_08520 5′ flank Rv	**AAATGGATTGGAAGTGGATAAC**GTTGGGCGGGTTGGTAGATT	Reverse primer for 5′ flank of homologous recombination oligo. Overlap sequence in bold
NRRL3_08520 3′ flank Fw	**AAATGGATTGGATTGGAAGTAC**GATCAGCAGCGCAACCCGTC	Forward primer for 3′ flank of homologous recombination oligo. Overlap sequence in bold
NRRL3_08520 3′ flank Rv	GCTGATAGAATATCTATAAT	Reverse primer for 3′ flank of homologous recombination oligo
PyrG Fw	**TTATCCACTTCCAATCCATTTG**CCCCTTTTAGTCAATACCG	Forward primer for pyrG fragment of homologous recombination oligo. Overlap sequence in bold
PyrG Rv	**TACTTCCAATCCAATCCATTTG**CGCAACTTCCTCGAGAAC	Reverse primer for pyrG fragment of homologous recombination oligo. Overlap sequence in bold
NRRL3_08520 Check Fw	ATGAATCGAGTAAACGAGTT	Forward primer to verify homologous recombination
NRRL3_08520 Check Rv	GATATACACCGTCACATACT	Reverse primer to verify homologous recombination

2.4 Protoplastation and transformation of *A. niger*

Aspergillus niger protoplasts were created using young hyphae as previously described (Debets and Bos, 1986). Transformations were performed as previously described (Master et al., 2008) using the SMC, STC, and PEG-TC buffers listed in Table 2. Protoplasts were generated using VinoTaste Pro enzymes (Novozymes, Copenhagen, Denmark) in SMC buffer. Following incubation at 37ºC with shaking at 75 rpm for 2 hours, protoplasts were collected by gravity filtration of the mixture through Miracloth and washed with STC three times. Protoplasts were observed using a light microscope and the concentration determined using a hemocytometer. Transformations were performed using 1.5 µg of CRISPR plasmid DNA and 1 nmol of the corresponding gene-editing oligo in a 10 µL volume (GE-oligo; Table 3) added to 200 µL of

protoplasts with a minimum concentration of 1 x 10^7 protoplasts/mL. Samples with no plasmid and with a plasmid lacking gRNA were also included as negative and positive controls, respectively. A volume of 0.67 mL PEG-TC was added and mixed gently before allowing the mixture to rest for exactly 5 minutes. After 5 minutes, 1.33 mL of STC was added. The transformation mixture was then plated on SRM.

2.5 Mutant generation using CRISPR-Cas9

Single-stranded 60-base gene-editing oligonucleotides with 30 bases of homology on each side of the CRISPR cut site along with the CRISPR/Cas9 plasmids were introduced into the protoplasts by co-transformation. With the deletion of the *kusA* gene in *A. niger* CBS 138852 preventing non-homologous end-joining, homology-directed repair using these oligonucleotides was the only way for the cells to repair the double stranded break created in the target genes. These GE-oligos created deletions of 10 or 31 (depending on the gene) base pairs in the repaired DNA, as shown in Figure 6. These deletions also caused frameshift mutation in the sequence of the gene on the 3' end of the deletion. Table 3 lists all the GE-oligos used in this study.

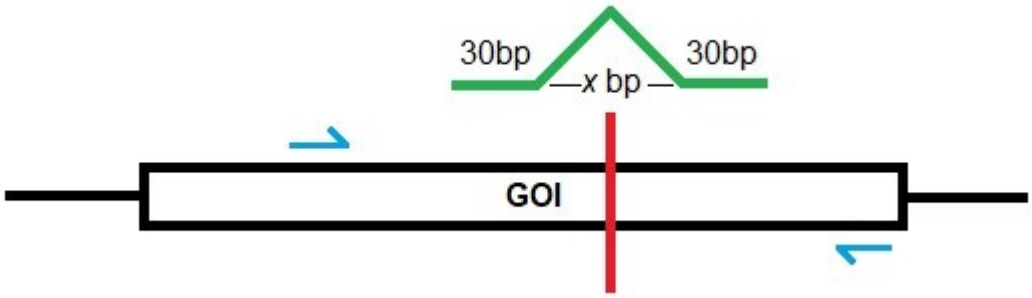

Figure 6. Diagram of mutant generation and PCR verification using CRISPR/Cas9 and rescue oligonucleotide. Red line indicates the location of the double-stranded break created in the gene of interest. The gene-editing oligonucleotide (green) used for homology directed repair creates a deletion of *x* bp (10 bp or 31 bp in this study). Primers (blue) around the deletion are used for PCR verification of the deletion.

The only exception to this method of mutant generation was the NRRL3_08520 mutant which was generated by homologous recombination. Primers were used to create 700 bp fragments

homologous to the sequences upstream and downstream of the gene, and a fragment containing the full *pyrG* sequence. These fragments included overlapping sequences which were used to combine them by fusion PCR. The resulting DNA fragment (10 µg) was introduced into *A. niger* CBS 138852 as above to replace NRRL3_08520 with *pyrG* by homologous recombination.

2.6 Extraction of genomic DNA

Spores from the transformant colonies were touched with a wet cotton tip and swirled into 250 µL of CM and incubated at 30°C overnight. Mycelia were then collected from the surface, placed on the sides of the wells of a 96-well PCR plate (Thermo Scientific, Waltham, MA), and frozen at -80°C for at least one hour. Mycelia were thawed and centrifuged at room temperature and 3000 rpm for 15 minutes to remove any media or other liquid. Lysis buffer (Table 2) was mixed with proteinase K (New England Biolabs, Ipswich, MA) in a 49:1 ratio for a final concentration of 8 units of proteinase K per 50 µL. The mixture was added to the mycelia (50 µL per well) and incubated at 55°C for 75 minutes. TE buffer (Table 2) was added (100 µL per well) and incubated at 95°C for 5 minutes. Plates were centrifuged again at 3000 rpm for 15 minutes to remove cell debris and 20-30 µL of the supernatant containing the genomic DNA were placed in new wells to be used as PCR template (1 µL/20 µL PCR reaction). DNA was stored at -20°C.

2.7 Mutant verification

Mutations in the transformants were verified by PCR amplification of sequences surrounding the deletion site as shown in Figure 6. Primers used for verification are listed in Table 3. PCR was run on genomic DNA extracted from both the parent strain transformed with the empty ANEp8 plasmid as a control (expected to lack the deletion), and several of the deletion transformants from each transformation. Due to the small size of some of the deletions, amplified DNA fragments were visualized following gel electrophoresis on 3% agarose gel to increase resolution, rather than the standard 0.8% used in all other cases in this study. Additionally, agarose gels used for deletion verification were created and electrophoresis was run using TBE buffer rather than TAE buffer to increase resolution of small fragments (Miura et al., 1999).

2.8 Growth profile of permease mutants

Wells in a 96-well plate were filled with 190 µL of MM at various pH levels with 1% quinic acid as the sole carbon source. Wells were inoculated to a final concentration of 2.5 x 10^5 spores/mL using 10 µL of spore solution. Growth occurred at 30°C over 120 hours with optical density at 595 nm being measured every hour using a Tecan Sunrise microplate reader (Tecan, Männedorf, Switzerland). Results are the average of triplicate samples.

2.9 Analysis of mutants by liquid chromatography-mass spectrometry

Spores from parental and mutant strains were used to inoculate 200 mL of CM and cultures were grown overnight at 32°C with shaking at 220 rpm. Mycelia were collected by gravity filtration through Miracloth and washed well with MM lacking a carbon source. Excess liquid was removed from the mycelia by gentle pressing with paper towel and separated into triplicate samples of 250 mg each. These samples were quickly placed in 25 mL of MM containing either quinic acid or 3,4-DHB and incubated at 30°C with shaking at 220 rpm. Samples of the cultures were collected at various time points (0, 6, 24, and 48 hours) in 1.5 mL microcentrifuge tubes. Samples were centrifuged at 13,300 rpm for 30 minutes to remove mycelia. Supernatants were transferred to new tubes and an equal volume of -20°C methanol was added. Samples were incubated on ice for 10 minutes and centrifuged at 13,300 rpm for 30 minutes to precipitate proteins. Supernatants were transferred to wells of a microtiter plate and an equal volume of 0.1% formic acid was added to each and mixed.

Standards were created to quantify the drop in quinic acid in the media over time. Eight standards were created with concentrations of 0.5%, 0.2%, 0.05%, 0.02%, 0.005%, 0.002%, 0.0005%, and 0.0002% quinic acid in ddH_2O. These standards were used to create a graph of the area of the quinic acid peaks vs concentration of quinic acid. This was used to derive an equation to estimate the amount of quinic acid remaining in each sample at each time point.

Electrospray liquid chromatography-mass spectrometry was performed in both negative and positive mode on an Agilent 6560 Ion mobility Q-TOF and 1290 Infinity II LC/MS system

(Agilent Technologies, Santa Clara, CA). Scan range was from 100 to 1400 m/z. Reversed-phase liquid chromatography (RPLC) was performed using a Synergi 4 µm Hydro-RP 80 Å, 150 x 2.00 mm column (Phenomenex, Torrance, CA). Solvents used to generate the gradient for the RPLC separation were 0.1 % formic acid in water for Solvent A and 0.1% formic acid in acetonitrile for Solvent B. Analysis of the data was performed using Agilent MassHunter workstation software. Specifically, MassHunter Qualitative Analysis was used to visualize the total ion and extracted ion chromatograms and mass spectra, and MassHunter Profinder B.08.00 batch recursive feature extraction wizard was used with the default settings to find the area of the peaks for each compound in each sample. Microsoft Excel was used to create graphs of this data.

3. Results

3.1 Identification of potential 3,4-DHB pathway genes

Although the β-ketoadipate pathway has been known for decades to exist in fungi such as *A. niger* and *Neurospora crassa* (Cain et al 1968; Mazur et al., 1994), few enzymes of the pathway have been linked to genes. To identify the genes involved in the 3,4-DHB catabolic pathway, predictions were first made based on functional annotation of the *A. niger* genome and on homology with any genes are known to be involved, as well as the literature about these enzymes in fungi. Comparative transcriptomics was then used to provide experimental evidence in support of the candidate genes.

3.1.1 Predictions based on functional annotation and homology

Although no genes were annotated as protocatechuate-3,4-dioxygenase, the first enzyme of the pathway, the *A. niger* NRRL3 genome contains five genes annotated as non-secreted intradiol ring cleavage dioxygenases (Aguilar-Pontes et al. 2018). In *A. nidulans*, the gene AN8566 was upregulated both at the gene and protein levels by growth on benzoic acid, which is catabolized through the 3,4-DHB pathway (Martins et al. 2015). The *A. niger* orthologue of this gene is NRRL3_01405, which has 94% identity with AN8566 in the protein sequence. This gene is annotated as catechol-1,2-dioxygenase, a ring fission enzyme that cleaves catechol to form

cis,cis-muconic acid as the first step of the catechol branch of the β-ketoadipate pathway. Due to the similarity of protocatechuate-3,4-dioxygenase and catechol-1,2-dioxygenase, NRRL3_01405 appeared to be a strong candidate. Additional evidence is found in a recent study of the substrate specificities of four purified intradiol dioxygenases from *A. niger* where NRRL3_01405 was shown to have significant activity on 3,4-DHB (Semana and Powlowski, 2019).

The second enzyme in the pathway, carboxy-cis,cis-muconate cyclase, has been identified in *A. niger* and characterized quite extensively decades ago, but no gene has been identified (Cain et al., 1968; Thatcher and Cain, 1974; Thatcher and Cain,1975). Only one gene in the NRRL3 genome, NRRL3_02586, is annotated as carboxy-cis,cis-muconate cyclase. The characterized *N. crassa* orthologue, CMLE_NEUCR, shares 67% identity with this protein (Mazur et al., 1994). The corresponding *A. nidulans* orthologue AN1151 (91% identity in protein sequence with NRRL3_02586), was upregulated in both the transcriptome and proteome when grown on benzoic acid, and a knockout of this gene prevented growth on benzoate media and resulted in accumulation of a compound putatively identified as 3-carboxymuconate (Martins et al. 2015).

3-carboxymuconolactone hydrolase/decarboxylase is more difficult to predict. Two genes are annotated as carboxymuconolactone decarboxylase-like proteins: NRRL3_08340 and NRRL3_03759. Cain et al. (1968) identified a hydrolase/decarboxylase enzyme in *A. niger* with a mass of 54(±5) kDa that converted 3-carboxymuconolactone to β -ketoadipate in a single step, however it was not linked to any gene. NRRL3_08340 and NRRL3_03759 appear to encode proteins too small to be this enzyme, with molecular weights of only 15.88 and 21.55 respectively. In bacteria, the pathway diverges from the fungal pathway at the previous step, where 3-carboxy-cis,cis-muconic acid is instead lactonized to 4-carboxymuconolactone, converted by a decarboxylase and then a hydrolase into β-ketoadipate (Harwood and Parales, 1996). In *A. nidulans*, Martins et al. identified this gene as AN5232, encoding a 26 kDa protein only weakly upregulated in the transcriptome and proteome when grown on benzoic acid. This gene has a weak similarity to 4-carboxymuconolactone decarboxylase. However, when knocked out this gene did prevent growth on benzoate and resulted in some accumulation of all intermediates up to 3-carboxymuconolactone. On this basis AN5232 was assigned as 3-carboxymuconolactone hydrolase (decarboxylating).

The *A. niger* orthologue of AN5232 is NRRL3_00837 (72% identity). However, this gene has a predicted protein molecular weight of 25.48 kDa, about half that of the enzyme identified by Cain et al. The NRRL3 annotation of this gene indicates the presence of AhpD domain (InterPro entry IPR029032), based on similarity to a characterized alkylperoxide reductase from *M. tuberculosis* (AHPD_MYCTU). No significant similarity was found between APHD_MYCTU and NRRL3_00837 in a BLASTP comparison under default settings. The function of this gene is unclear based on homology.

During the course of this work, another candidate for this enzyme came our attention from the work of Powlowski and Olyaie (personal communication) who had expressed and characterized NRRL3_01409 which has this activity. This candidate was identified via a BLAST search of the NRRL3 genome using the sequence of ELH2_ACIAD, a β-ketoadipate enol-lactonase from *Acinetobacter baylyii,* an enzyme induced specifically with other enzymes of the 3,4-DHB branch of the β-ketoadipate pathway (Cánovas and Stanier, 1967). NRRL3_01409 is 28% identical and is annotated as an alpha/beta hydrolase fold-1 domain-containing protein, but unlike the other three candidates, has a predicted molecular weight of 60.95 kDa, similar to that predicted by Cain et al., making it a possible candidate.

The next step of the pathway involves conversion of β-ketoadipate to β-ketoadipyl-CoA through the activity of β-ketoadipate-CoA transferase. Three different genes were annotated succinyl-CoA:3-ketoacid coenzyme A transferases (the systematic name of this enzyme): NRRL3_01886, NRRL3_11640 and NRRL3_01593. Finally, the β-ketoadipyl-CoA transferase which converts β-ketoadipyl-CoA into succinyl-CoA and acetyl-CoA, also had three candidates. NRRL3_01526, NRRL3_07786 and NRRL3_11162 were all annotated as 3-ketoacyl-CoA thiolase, an alternative name for this enzyme. Figure 7 shows the possible genes for each of the enzymes in the 3,4-DHB pathway based on functional annotation and homology, and literature.

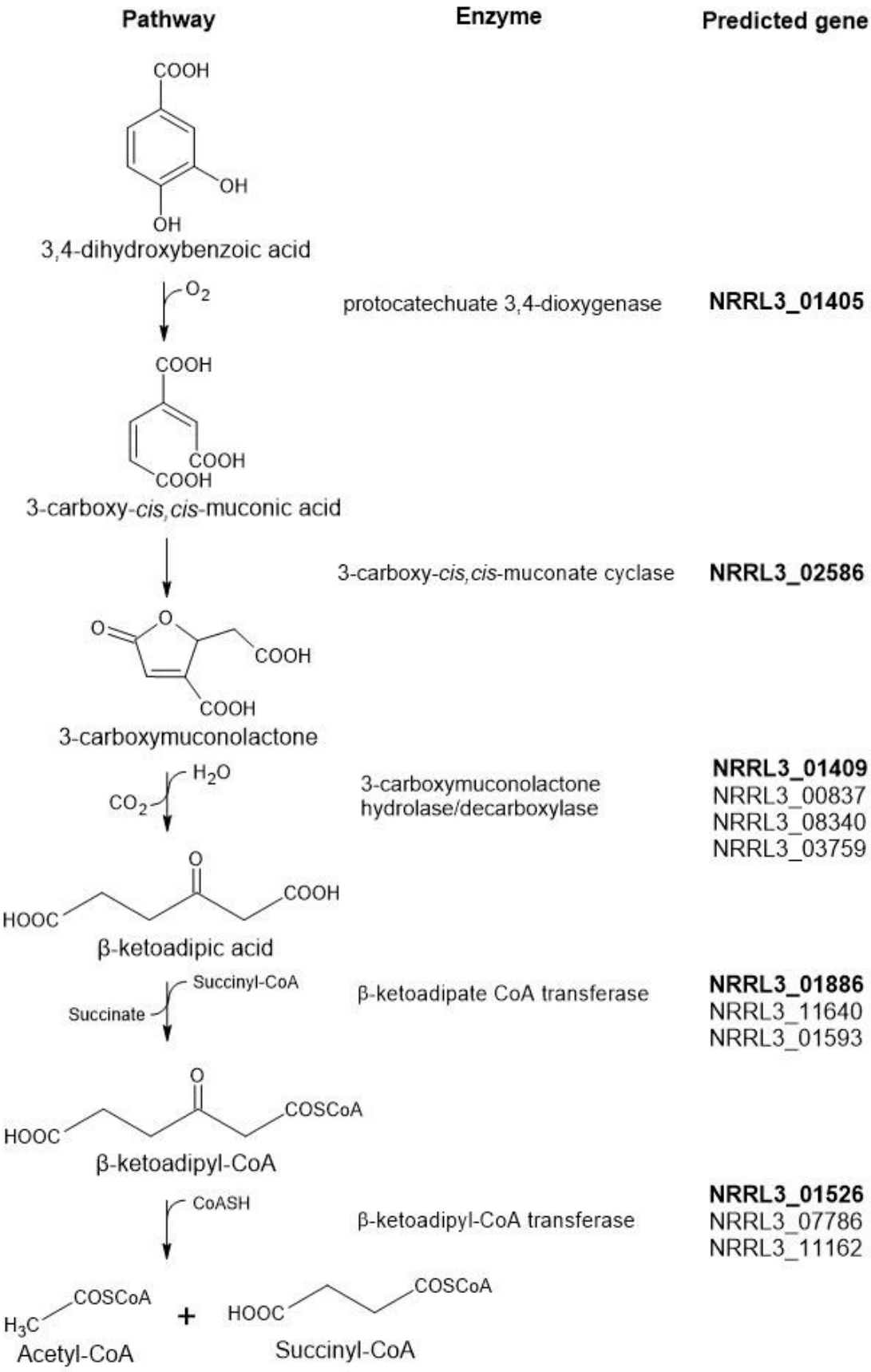
Pathway
Enzyme
Predicted gene
COOH
OH
OH
3,4-dihydroxybenzoic acid
O2
protocatechuate 3,4-dioxygenase
NRRL3_01405
COOH
COOH
COOH
3-carboxy-cis,cis-muconic acid
3-carboxy-cis,cis-muconate cyclase
NRRL3_02586
O
O
COOH
COOH
3-carboxymuconolactone
H2O
CO2
3-carboxymuconolactone hydrolase/decarboxylase
NRRL3_01409
NRRL3_00837
NRRL3_08340
NRRL3_03759
O
COOH
HOOC
β-ketoadipic acid
Succinyl-CoA
Succinate
β-ketoadipate CoA transferase
NRRL3_01886
NRRL3_11640
NRRL3_01593
O
COSCoA
HOOC
β-ketoadipyl-CoA
CoASH
β-ketoadipyl-CoA transferase
NRRL3_01526
NRRL3_07786
NRRL3_11162
COSCoA
H3C
+
HOOC
COSCoA
Acetyl-CoA
Succinyl-CoA

Figure 7. The 3,4-dihydroxybenzoic acid catabolic pathway. Possible genes based on functional annotation and homology are listed for each enzyme, with the predicted gene in bold.

3.1.2 Comparative transcriptomics of *A. niger* grown on 3,4-DHB and fructose as a sole carbon source

RNA sequencing was performed using *A. niger* CBS 138852 grown in media containing either fructose or 3,4-DHB as the sole carbon source. Results from select genes are shown in Table 4 including all of the possible genes, as identified in the previous section, encoding the enzymes of the 3,4-DHB pathway. Clear candidates were observed for four of the five enzymes based on the level of expression in the 3,4-DHB media and upregulation in that media compared to the fructose media.

The only gene identified as a likely candidate for protocatechuate-3,4-dioxygenase, NRRL3_01405, was highly expressed on 3,4-DHB. It was the third most highly expressed gene in the entire genome (full data not shown). Compared to fructose, the expression level also showed a fold-change of over 2000, making this a very strong candidate.

The single gene identified as the possible 3-carboxy-*cis,cis*-muconate cyclase-encoding gene, NRRL3_02586, also showed strong expression levels and upregulation. Though not as high as NRRL3_01405, this gene had a high transcripts per kilobase million (TPM) count (a measure of the amount of the total mRNA molecules belonging to this gene and normalizing for gene length and sequencing depth) of over 1000. It also had a fold-change of 60 in expression levels in the 3,4-DHB samples compared to the fructose samples. Again, this provides strong evidence in favor of the only candidate.

The RNA-seq data failed to distinguish which gene encodes the third enzyme or enzymes in the pathway, 3-carboxymuconolactone hydrolase/decarboxylase. Among the two genes with related annotation, NRRL3_03759 had low expression levels on both media and was only minimally upregulated on the 3,4-DHB. However, NRRL3_08340 was among the most highly expressed genes on 3,4-DHB with an average TPM count of over 3600. It was also very highly

upregulated, being about 241-fold more highly expressed on 3,4-DHB than fructose. While NRRL3_08340 appeared to be a very strong candidate on this basis, NRRL3_01409 and NRRL3_00837 were both also highly expressed on 3,4-DHB (728 and 1038 TPM respectively), and upregulated in this sample compared to fructose (69x and 48x respectively). Although one gene can essentially be ruled out, three possibilities exist for this enzyme.

The final two enzymes, β-ketoadipate-CoA transferase and β-ketoadipyl-CoA transferase, respectively, each had one clear highly expressed candidate among the three possibilities identified for each by annotation. For β-ketoadipate-CoA transferase this was NRRL3_01886, and for β-ketoadipyl-CoA transferase it was NRRL3_01526. These genes had much higher expression levels and upregulation on 3,4-DHB than the other four genes.

Among the rest of the genome that had not been identified as possibly encoding enzymes in this pathway, 1076 genes had a fold change greater than 4 when grown on 3,4-DHB compared to fructose as the sole carbon source. Several noteworthy genes with high expression levels on 3,4-DHB and high upregulation are included in the bottom half of Table 4. NRRL3_01635 and the adjacent NRRL3_01634 encode proteins with alcohol dehydrogenase and aldehyde dehydrogenase domains, respectively. NRRL3_01635 was the most highly upregulated gene in the entire genome on 3,4-DHB compared to fructose, and among the most highly expressed. The relationship of activities of the enzymes encoded by these genes to the 3,4-DHB pathway, if any, is unclear. An alcohol dehydrogenase could potentially convert 3,4-dihydroxybenzyl alcohol into an aldehyde, which the NRRL3_01634 enzyme could then convert to 3,4-DHB, however there is no evidence of this. Another possibility is that the high expression level of this gene is a sign of oxidative stress, as studies have shown this can induce expression of alcohol dehydrogenase (Kelly et al., 1990). Several of the other genes included in Table 4 could potentially act on aromatic compounds or cellulose, and were perhaps upregulated due to a transcription factor active in the presence of 3,4-DHB or other aromatic compounds, though no clear candidate for a 3,4-DHB transcription factor stands out in the data. NRRL3_02833 encodes a transporter which could potentially be involved in transport of 3,4-DHB or other aromatic compounds into or out of the cell.

Table 4. RNA-sequencing data for *A. niger* CBS 138852 grown on fructose or 3,4-DHB as sole carbon source. Expression results are shown as the average transcripts per kilobase million (TPM) of duplicate samples for each carbon source. Genes above the break were identified as described in the text as candidates for 3,4-DHB pathway enzymes.

# Transcript	Description	3,4-DHB TPM	Fructose TPM	Fold change (3,4-DHB vs. Fructose)
NRRL3_01405	catechol 1,2-dioxygenase	18330.48	8.24	2223.4
NRRL3_02586	carboxy-cis,cis-muconate cyclase	1016.83	16.88	60.23
NRRL3_00837	AhpD-like domain-containing protein	1038.39	21.48	48.34
NRRL3_01409	alpha/beta hydrolase fold-1 domain-containing protein	727.89	10.51	69.23
NRRL3_08340	carboxymuconolactone decarboxylase-like protein	3661.57	15.2	240.94
NRRL3_03759	carboxymuconolactone decarboxylase-like protein	6.76	1.51	4.49
NRRL3_01886	succinyl-CoA:3-ketoacid coenzyme A transferase	1499.22	20.66	72.58
NRRL3_11640	succinyl-CoA:3-ketoacid coenzyme A transferase 1, mitochondrial	360.56	32.14	11.22
NRRL3_01593	succinyl-CoA:3-ketoacid coenzyme A transferase	55.19	32.52	1.7
NRRL3_01526	3-ketoacyl-CoA thiolase	1410.97	19.21	73.47
NRRL3_07786	3-ketoacyl-CoA thiolase	254.02	274.13	0.93
NRRL3_11162	3-ketoacyl-CoA thiolase	143.4	37.38	3.84
NRRL3_01635	alcohol dehydrogenase domain-containing protein	7756.66	0.42	18301.02
NRRL3_01634	aldehyde dehydrogenase domain-containing protein	5928.22	29.52	200.80
NRRL3_05639	methionine synthase domain-containing protein	1748.09	3.95	442.02
NRRL3_08551	phenol hydroxylase - C-terminal dimerization domain-containing protein	1547.63	5.90	262.30
NRRL3_02833	major facilitator superfamily protein	800.37	0.92	872.93
NRRL3_07975	beta-glucosidase	617.44	0.17	3555.25
NRRL3_10132	flavin monooxygenase-like protein	360.49	0.74	484.83
NRRL3_08804	flavin-dependent halogenase family protein	199.17	0.03	7715.90

3.2 Identification of potential quinic acid pathway genes

The genetics of the quinic acid catabolic pathway leading to 3,4-DHB (Figure 8a) have been studied previously in fungi (Giles et al., 1991; Hawkins et al., 1993). This makes the identification of candidate genes simpler for this pathway compared to the 3,4-DHB pathway, as orthologs of the *Neurospora crassa* and *Aspergillus nidulans* quinic acid pathway genes are present in *Aspergillus niger*. Functional annotation, homology, and comparative transcriptomics were used for this pathway as well to determine whether any other genes are potentially involved in this pathway and to provide experimental evidence in support of the candidate genes.

3.2.1 Predictions based on sequence similarity and functional annotation

As a first step, orthologs of the *N. crassa* and *A. nidulans* quinic acid cluster genes were identified in *A. niger.* The *A. niger* gene NRRL3_11038 was found to be the ortholog of the activator, with 60% identity to QutA (AN1134) of *A. nidulans* and similar length. Many of the genes involved in quinic acid catabolism in *A. niger* were similarly clustered on Chromosome_8, with the repressor ortholog in *A. niger* being NRRL3_11039 (66% protein sequence identity with QutR), the dehydroquinate dehydrase ortholog being NRRL3_11037 (82% identity with QutG), the dehydroshikimate dehydrase ortholog being NRRL3_11035 (56% identity with QutC), and the quinate permease ortholog being NRRL3_11036 (77% identity with QutD). Unlike in *N. crassa* and *A. nidulans*, some of the genes involved in this pathway have translocated outside of the cluster. The quinate dehydrogenase ortholog was found to be NRRL3_08520 (72% identity with QutB) on Chromosome_6 and one of the genes of unknown function relocated along with it, NRRL3_08521 (79% identity with QutG). An additional gene with unknown function, QutH, also had an ortholog outside of the cluster, NRRL3_10284, with 53% identity. This gene cluster in *A. nidulans* and the corresponding cluster and genes in *A. niger* are shown in Figure 8b. All eight genes were of similar length to the *A. nidulans* orthologs and based on sequence similarity and protein motif similarity all were given the functional annotation of the corresponding *A. nidulans* genes, except for the genes of unknown function. NRRL3_8521 was annotated as oxidoreductase domain-containing protein. NRRL3_10284 was annotated as inositol monophosphatase-like protein based on sequence similarity to a bovine monophosphatase gene,

though it is largely based on a conserved metal ion binding motif and core structure and the substrate is likely different (York et al., 1995). It has been proposed that this gene may have a function in downregulating quinic acid cluster expression by dephosphorylating the activator protein (Levesley 2013).

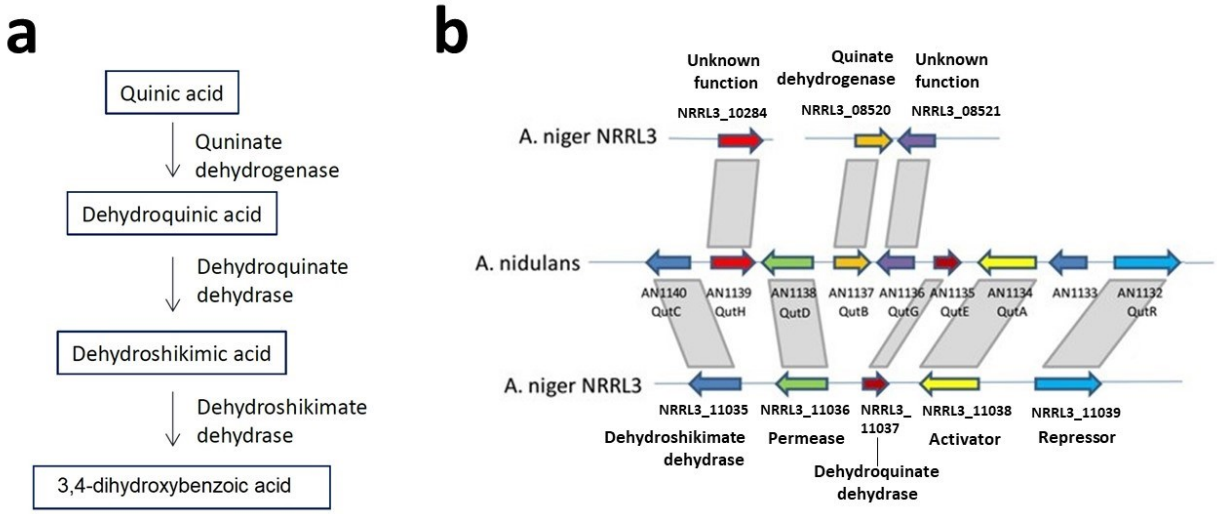

Figure 8. Quinic acid catabolic pathway (a) and gene cluster (b). Gene cluster in *Aspergillus nidulans* is shown with the corresponding orthologs in *Aspergillus niger* NRRL3 and gene function. In *A. niger* three genes are located outside of the cluster on either the same (NRRL3_10284) or a different chromosome (NRRL3_08520 and NRRL3_08521).

Another gene cluster in the *A. niger* NRRL3 genome contained the only other genes with protein sequence identity above 40% compared to the quinic acid pathway enzymes. This cluster includes NRRL3_07602, with 42% identity shared with the *A. nidulans* quinate dehydrogenase, NRRL3_07603, with 50% identity shared with the *A. nidulans* dehydroshikimate dehydrase, and NRRL3_07604, an transcription factor with 40% identity compared to the *A. nidulans* activator protein. The genes NRRL3_07602 and NRRL3_07603 are annotated as shikimate dehydrogenase substrate binding domain-containing protein and dehydroshikimate dehydrase, respectively. Two other genes in this cluster also have annotations related to shikimate, NRRL3_7600, shikimate dehydrogenase substrate binding domain-containing protein, and

NRRL3_7605, shikimate kinase/gluconokinase family protein. Another gene, NRRL3_05631, was also found to have high similarity to the *A. nidulans* quinate permease, with 60% identity in the protein sequence. This gene shares 66% identity with NRRL3_11036 and has only three fewer amino acids with 536.

3.2.2 Comparative transcriptomics of *A. niger* grown on quinic acid and fructose as a sole carbon source

RNA sequencing was performed using *A. niger* N593 grown in media containing quinic acid as the sole carbon source. Select genes are shown in Table 5 including the orthologs of the *A. nidulans* QUT cluster, the second *A. niger* cluster identified as having sequence similarity to these genes, the pentafunctional AROM polypeptide, and the second permease.

The genes in the cluster from NRRL3_11035 to NRRL3_11039 were all upregulated to varying degrees. The dehydroquinate dehydrase encoded by NRRL3_11037 was both expressed more than the other genes and had a much higher fold change compared to fructose, over 2000-fold. The other gene in the cluster encoding an enzyme, the dehydroshikimate dehydrase NRRL3_11035, while expressed and upregulated less than NRRL3_11037, had more than a 33-fold upregulation compared to fructose. The activator and repressor genes, NRRL3_11038 and NRRL3_11039 respectively, were expressed lower than the rest of the cluster, but still had fold changes of over 5 and 16 respectively.

The quinate permease NRRL3_11036 was upregulated almost 77-fold and had a TPM of approximately 142. The second quinate permease which was identified outside of the cluster, as discussed in the previous section, NRRL3_05631 was both more highly expressed and more highly upregulated on quinic acid. This gene had a TPM of approximately 266 and a fold change of approximately 127, both of which are almost twice that of in-cluster permease NRRL3_11036.

Among the other QUT cluster homologs, the quinate dehydrogenase NRRL3_08520 and the QutH homolog with unknown function, NRRL3_10284, were both highly expressed and

upregulated. These genes had a higher TPM (617 and 344 respectively) and fold change on quinic acid (approximately 244 and 249 respectively) than any of the genes in the cluster except for NRRL3_11037. However, the QutG ortholog, NRRL3_08521 was expressed nearly equally on both fructose and quinic acid.

The NRRL3_07600 cluster genes were mostly not upregulated on quinic acid compared to fructose. NRRL3_07601, encoding a nuclear transport factor domain-containing protein was the most upregulated, from near zero on fructose to over 100 TPM on quinic acid. The possible quinate dehydrogenase and dehydroshikimate dehydrase orthologs, NRRL3_07602 and NRRL3_07603 respectively, were also upregulated from near zero TPM on fructose to 15 and 6.5 TPM on quinic acid.. The remaining genes in this cluster appear to not be upregulated.

The pentafunctional AROM polypeptide NRRL3_11273 was expressed on both fructose and quinic acid, though approximately 6-fold higher on fructose.

Table 5. RNA-sequencing data for *A. niger* CBS 138852 grown on fructose or quinic acid as sole carbon source. Expression results are shown as the average transcripts per kilobase million (TPM) of duplicate samples for each carbon source. Genes above the break are orthologs of *A. nidulans* QUT cluster genes.

# Transcript	Description	Fructose TPM	Quinic acid TPM	Fold change (Quinic acid vs. Fructose)
NRRL3_11035	dehydroshikimate dehydratase	8.15	273.10	33.53
NRRL3_11036	major facilitator superfamily, sugar/inositol transporter-like protein	1.85	142.20	76.67
NRRL3_11037	catabolic 3-dehydroquinase	1.19	2471.19	2082.09
NRRL3_11038	fungal-specific transcription factor, quinic acid utilization	5.58	30.20	5.41
NRRL3_11039	quinate repressor protein	4.96	79.45	16.01
NRRL3_10284	oxidoreductase domain-containing protein	1.38	344.34	248.89
NRRL3_08520	quinate dehydrogenase	2.52	617.26	244.84
NRRL3_08521	inositol monophosphatase-like protein	22.83	27.76	1.22

NRRL3_07600	shikimate dehydrogenase substrate binding domain-containing protein	0.18	0.72	4.10
NRRL3_07601	nuclear transport factor 2 domain-containing protein	0.08	103.58	1257.31
NRRL3_07602	shikimate dehydrogenase substrate binding domain-containing protein	0.02	14.92	720.02
NRRL3_07603	dehydroshikimate dehydratase	0.23	6.46	28.00
NRRL3_07604	fungal-specific transcription factor	0.13	0.49	3.82
NRRL3_07605	shikimate kinase/gluconokinase family protein	0.92	1.70	1.85
NRRL3_11273	Pentafunctional AROM polypeptide	33.31	5.28	0.16
NRRL3_05631	major facilitator superfamily, sugar/inositol transporter-like protein	2.08	266.12	127.98

3.3 Generation of knockout mutants for candidate 3,4-pathway and quinic acid pathway genes

Based on the information in section 3.1, seven 3,4-DHB pathway genes were chosen to be knocked out. These included the most likely gene encoding each enzyme except for 3-carboxymuconolactone hydrolase/decarboxylase, for which the three possible genes were all chosen. The genes to be knocked out were: NRRL3_01405, NRRL3_02586, NRRL3_01409, NRRL3_00837, NRRL3_08340, NRRL3_01886, and NRRL3_01526. A double mutant was also created with deletions in both NRRL3_01409 and NRRL3_00837 in case both were necessary for that step of the pathway.

The quinic acid pathway genes predicted to encode the enzymes (NRRL3_11035, NRRL3_11037, and NRRL3_08520) and the two quinate permeases (NRRL3_11036 and NRRL3_05631) were also knocked out. The activator NRRL3_11038 was also knocked out by our lab in collaboration with Dr. Arthur Ram of Leiden University. This mutant showed very weak growth on quinic acid compared to the parent strain CBS 138852 (unpublished results), supporting the prediction that this cluster is involved in quinic acid catabolism in *A. niger* as it is in *A. nidulans.*

The method of mutant generation is described in the *Materials and Methods* section. CRISPR plasmids were constructed for each gene to be knocked out which expressed Cas9, a guide RNA specific to each gene, and the *pyrG* selectable marker (Figure 4). *A. niger* CBS 138852 was transformed with each of these plasmids along with the corresponding gene-editing oligo used to create a deletion at the target site during repair (Table 3; Figure 6). Transformant colonies which grew on media lacking uracil or uridine and therefore expected to be expressing the plasmid genes (pyrG) were screened by PCR to verify the presence of the expected deletion (Figure 6). The results of the PCR screening are shown below (Figures 9 and 10). Using primers on either side of the deletion, successful deletions were observed in all seven genes of the 3,4-DHB pathway and five genes of the quinic acid pathway. PCR products using the parent strain were larger than those using the mutant strains genomic DNA as a template, by the number of base pairs which were equal to the expected size of the deletion.

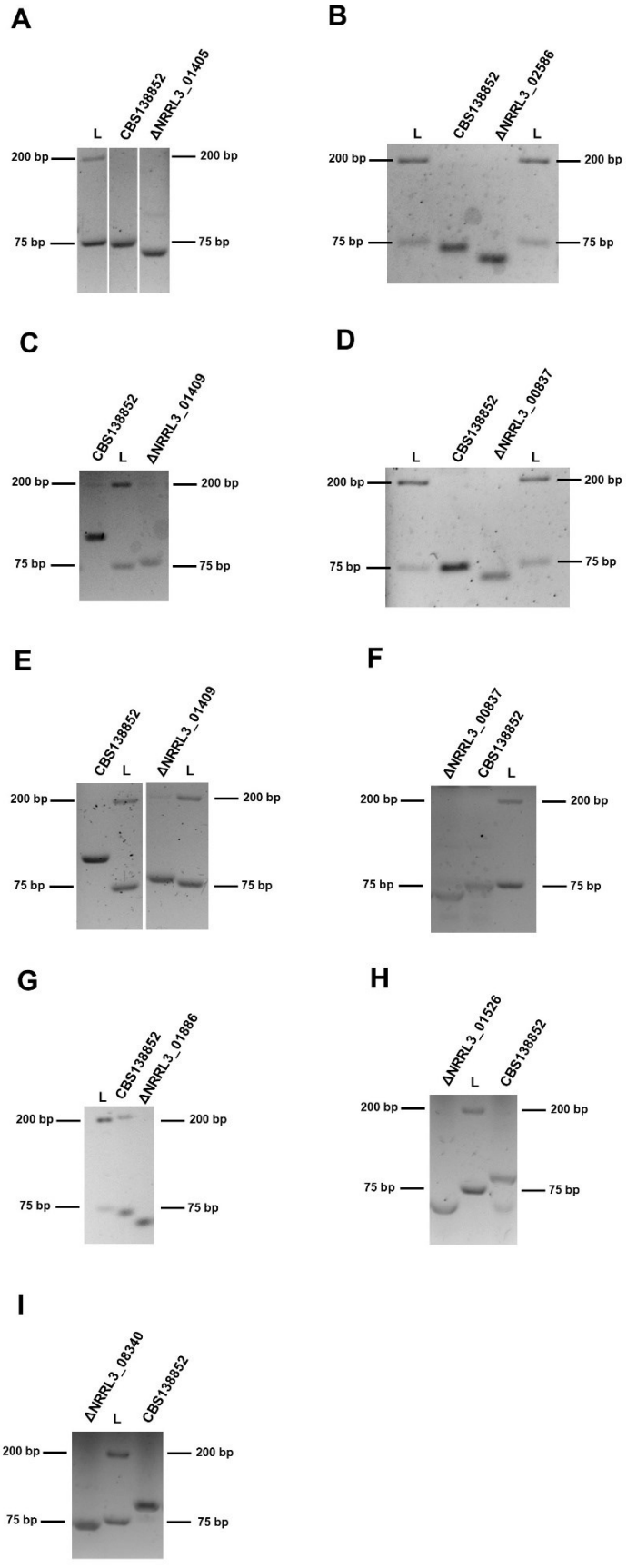
A
L
CBS138852
ΔNRRL3_01405
200 bp
75 bp
B
L
CBS138852
ΔNRRL3_02586
C
CBS138852
L
ΔNRRL3_01409
D
ΔNRRL3_00837
E
F
G
ΔNRRL3_01886
H
ΔNRRL3_01526
I
ΔNRRL3_08340

Figure 9. PCR verification of deletions in the 3,4-DHB pathway genes. Forward primers upstream and reverse primers downstream of deletions were used, as seen in Figure 6. Resulting fragment lengths are (A) 58 bp for ΔNRRL3_01405 and 68 bp for parent strain (CBS 138852), (B) 59 bp for ΔNRRL3_02586 and 69 bp for CBS 138852, (C) 79 bp for ΔNRRL3_01409 and 110 bp for CBS 138852, (D) 64 bp for ΔNRRL3_00837 and 74 bp for CBS 138852, (E) 79 bp for ΔNRRL3_01409 and 110 bp for CBS 138852 in double knockout of NRRL3_01409 and NRRL3_00837, (F) 64 bp for ΔNRRL3_00837 and 74 bp for CBS 138852 in double knockout of NRRL3_01409 and NRRL3_00837, (G) 60 bp for ΔNRRL3_01886 and 70 bp for CBS 138852, (H) 69 bp for ΔNRRL3_01526 and 90 bp for CBS 138852, (I) 69 bp for ΔNRRL3_08340 and 100 bp for CBS 138852. L = ladder (GeneRuler 1 kb Plus DNA Ladder; Thermo Scientific). Image from unpublished manuscript.

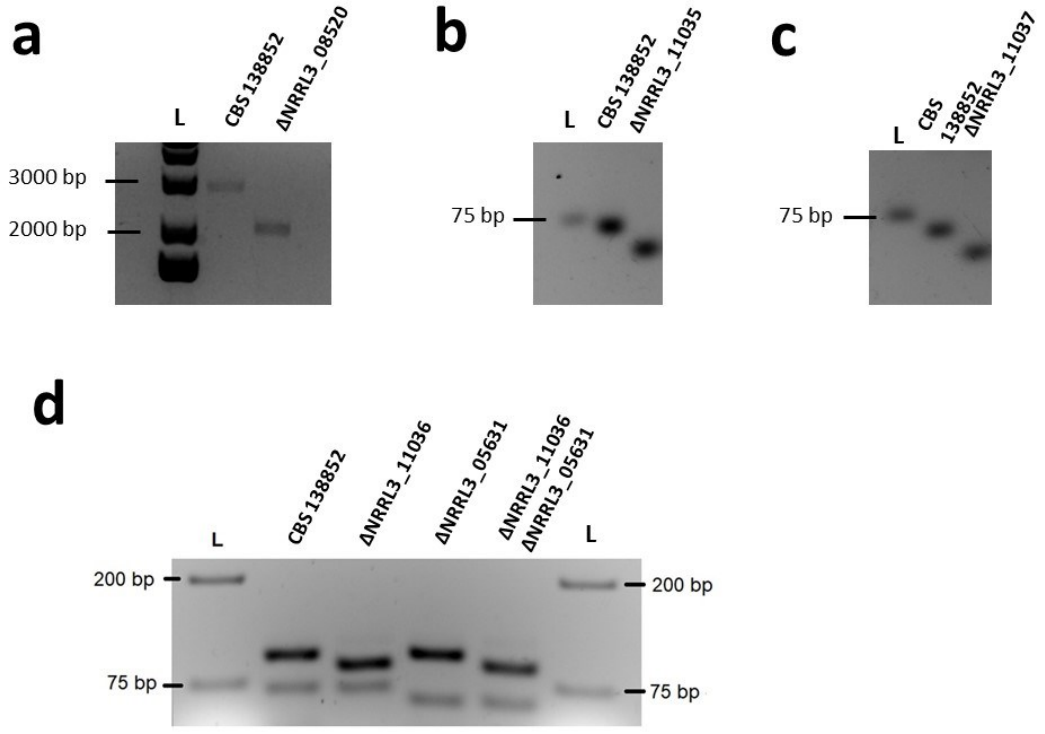

Figure 10. PCR verification of knockouts in the quinic acid pathway genes. Forward primers upstream and reverse primers downstream of deletions or genes were used. Resulting fragment lengths are (a) 2884 bp for ΔNRRL3_08520 and 2049 bp for parent strain (CBS 138852), (b) 60 bp for ΔNRRL3_11035 and 70 bp for CBS 138852, (c) 57 bp for ΔNRRL3_11037 and 67 bp for CBS 138852. For the permease genes, PCR included the primers for both genes (d). Fragments were 90 bp with deletion in NRRL3_11036 and 100 bp without deletion, and 60 bp with deletion

in NRRL3_05631 and 70 bp without deletion. L = ladder (GeneRuler 1 kb Plus DNA Ladder; Thermo Scientific).

3.4 Characterization of the protocatechuate-3,4-dioxygenase mutant

Mutants were grown on minimal media plates with only fructose, 3,4-DHB, or 2,3-dihydroxybenzoic acid (2,3-DHB), which is degraded through the catechol branch of the pathway, as the sole carbon source. All mutants were grown on fructose media as a control for general effects on growth, as these genes have no expected effect on fructose metabolism.

The ΔNRRL3_01405 mutant grew very poorly on 3,4-DHB and failed to sporulate, further supporting the assignment of this gene as the protocatechuate-3,4-dioxygenase (Figure 11a). As expected, growth on both fructose and 2,3-DHB was identical to the parental strain.

Mutants were also grown on complete media and mycelia were transferred to media with quinic acid as the sole carbon source. Quinic acid is known to be degraded through the 3,4-DHB branch of the 3-oxoadipate pathway (Hawkins et al., 1993). Liquid chromatography-mass spectrometry performed using the culture media at several time points showed that degradation of quinic acid in the ΔNRRL3_01405 mutant (Figure 11b) resulted in extracellular accumulation of 3,4-DHB at much greater levels than were observed in the parent strain (Figure 11c).

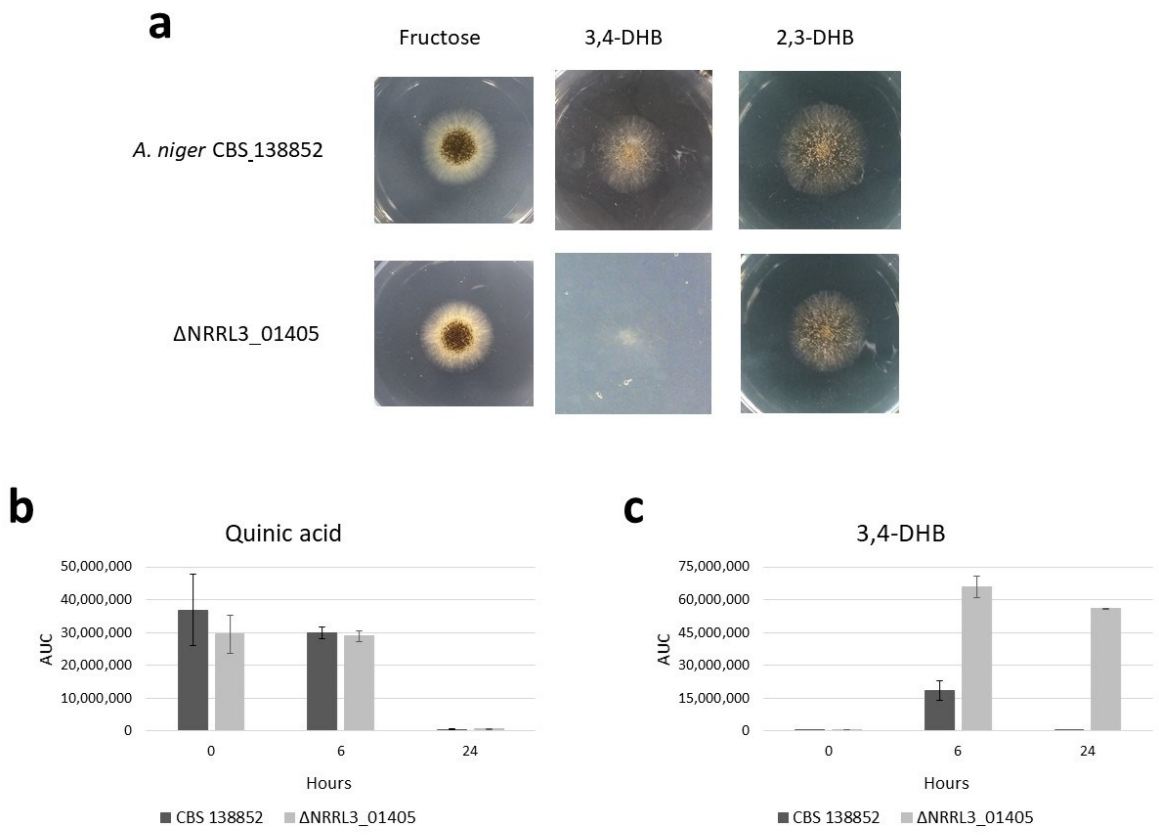

Figure 11. Growth phenotype and MS analysis of *A. niger* CBS 138852 ΔNRRL3_01405 mutant. (a) *Aspergillus niger* strains138852 and the ΔNRRL3_01405 were grown at 30°C for 72 hours on media containing fructose; 3,4-DHB; or 2,3-DHB as sole carbon source. (b) Decrease in quinic acid and (c) accumulation of 3,4-DHB in the ΔNRRL3_01405 mutant in minimal media with quinic acid as sole carbon source. Data from LC-MS in negative mode. AUC = area under curve, calculated by Agilent Profinder software.

3.5 Characterization of the 3-carboxy-cis,cis-muconate cyclase mutant

The 3-carboxy-cis,cis-muconate cyclase mutant, ΔNRRL3_02586, failed to grow on 3,4-DHB media, while growth on 2,3-DHB media appeared to be at most only slightly impaired compared to the parental strain (Figure 12a). The growth phenotype on fructose remained unaffected. Mass spectrometric analysis showed that the mutant consumed quinic acid at a much slower rate than the parent strain (Figure 12b). By 24 hours, the parent had consumed nearly all of the quinic

acid, while the mutant culture only showed a slight reduction at each time point and a large amount remained at 48 hours. However, even with the slower rate of catabolism, accumulation of 3-carboxy-cis,cis-muconic acid over time was observed in the mutant using LC-MS (Figure 12c). The amount of 3-carboxy-cis,cis-muconic acid increased at each time point as the quinic acid was degraded. No 3-carboxy-cis,cis-muconic acid was detected in the parent strain at any time point.

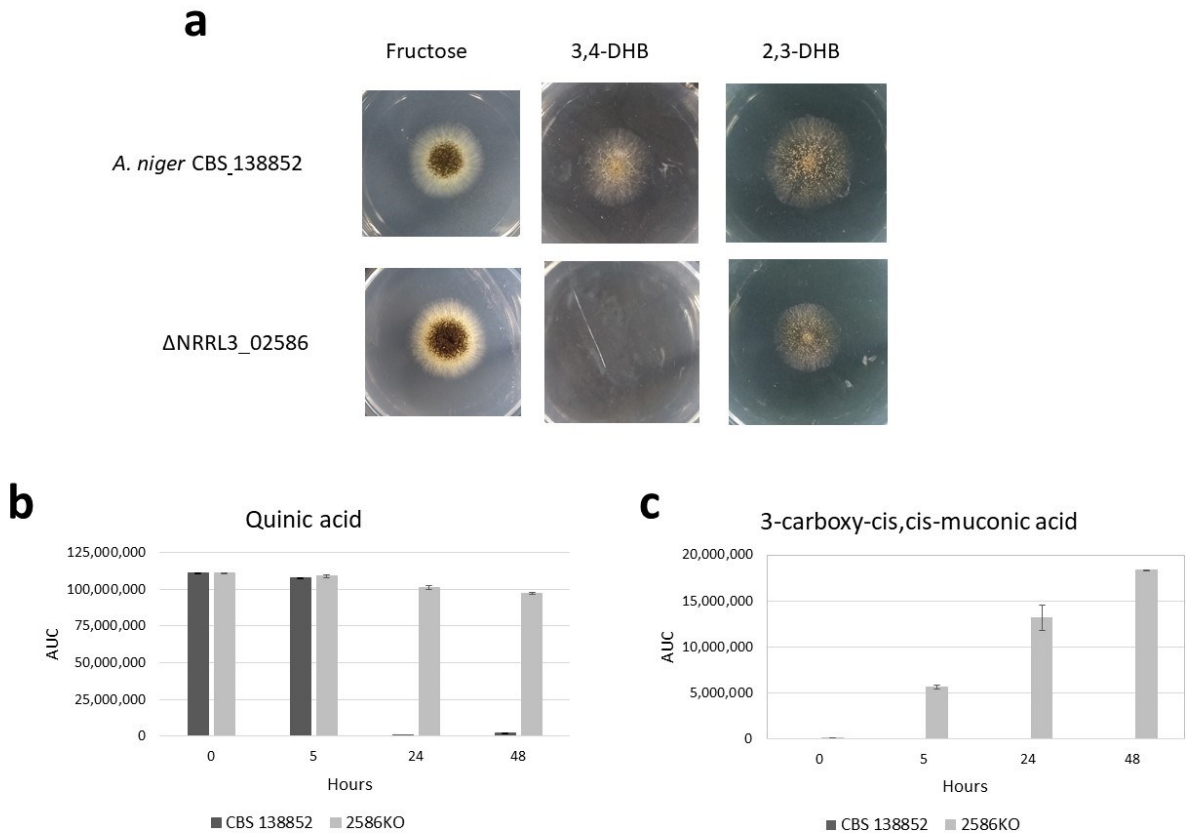

Figure 12. Growth phenotype and MS analysis of *A. niger* CBS 138852 ΔNRRL3_02586 mutant. (a) *Aspergillus niger* strains CBS 138852 and the ΔNRRL3_02586 were grown at 30°C for 72 hours on media containing fructose; 3,4-DHB; or 2,3-DHB as sole carbon source. (b) Decrease in quinic acid and (c) accumulation of 3,4-DHB in the ΔNRRL3_02586 mutant in minimal media with quinic acid as sole carbon source. Data from LC-MS in negative mode. AUC = area under curve, calculated by Agilent Profinder software.

3.6 Characterization of the 3-carboxymuconolactone hydrolase/decarboxylase mutant

The ΔNRRL3_08340 mutant had the same growth phenotype on fructose, 3,4-DHB and 2,3-DHB as the parent strain (Figure 13). However, the other two mutants both had a strong phenotype on 3,4-DHB media, with ΔNRRL3_00837 preventing any growth and ΔNRRL3_01409 growing very poorly. The ΔNRRL3_00837 mutant also slightly impaired growth on 2,3-DHB.

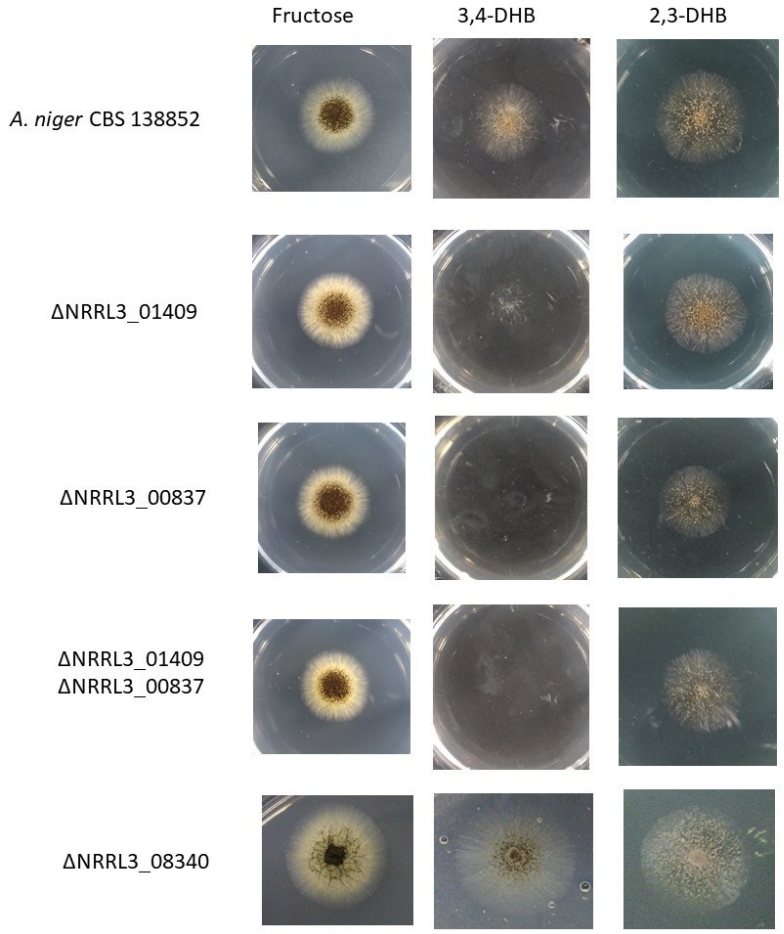

Figure 13. Growth phenotype of *A. niger* CBS 138852 and the possible 3-carboxymuconolactone hydrolase/decarboxylase mutants ΔNRRL3_01409, ΔNRRL3_00837, ΔNRRL3_08340, and double mutant ΔNRRL3_01409ΔNRRL3_00837. *Aspergillus niger* strains were grown at 30°C for 72 hours on media containing fructose; 3,4-DHB; or 2,3-DHB as sole carbon source.

LC-MS was performed on these mutants and the parent strains on 3,4-DHB, however although the level of 3,4-DHB dropped, no 3-carboxymuconolactone was detected in any samples. An additional mutant was generated with a double knockout of both ΔNRRL3_00837 and ΔNRRL3_01409 to see if both are necessary together for accumulation of 3-carboxymuconolactone. However, again none was detected. This double mutant had the same growth phenotype on all media as the ΔNRRL3_00837 mutant, with growth on 2,3-DHB slightly impaired and a lack of growth on 3,4-DHB (Figure 13)

3.7 Characterization of β-ketoadipate-CoA transferase and β-ketoadipyl-CoA transferase mutants

Mutants generated for NRRL3_01886 and NRRL3_01526 prevented any growth on both 3,4-DHB and 2,3-DHB as sole carbon sources, while growing equally well as the parental strain on fructose media (Figure 14a). This result is expected because the enzymes encoded by these genes are shared by multiple branches of the β-ketoadipate pathway including the catechol branch used for 2,3-DHB catabolism.

LC-MS detected accumulation of a compound with a mass of 160.01 g/mol, which corresponds to 3-oxoadipic acid (monoisotopic mass 160.037173 g/mol), in the ΔNRRL3_01886 samples which increased over time as the 3,4-DHB substrate was degraded (Figure 14b,c). This compound was also detected in the ΔNRRL3_01526 samples at slightly lower levels, however no β-ketoadipyl-CoA was detected.

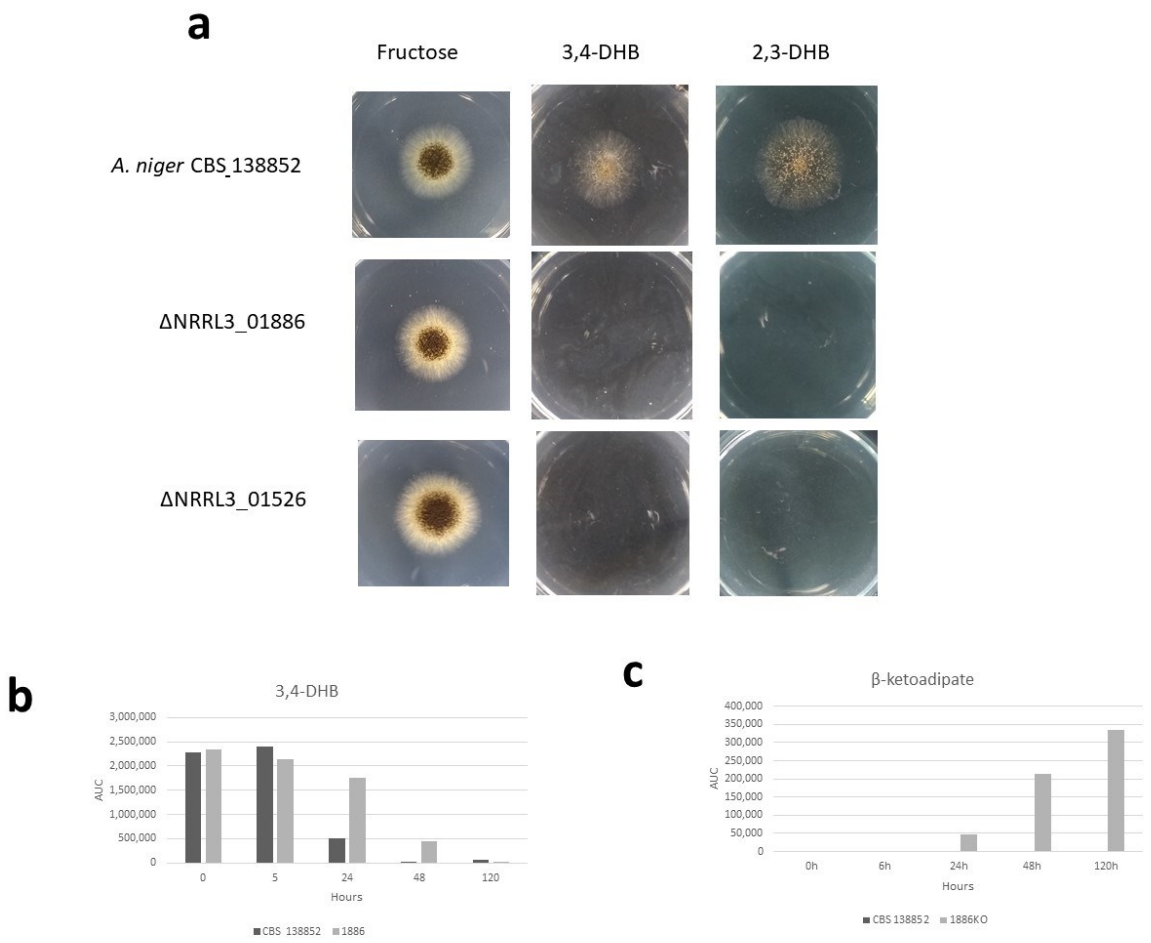

Figure 14. Growth phenotype and MS analysis of *A. niger* CBS 138852 ΔNRRL3_01886 and ΔNRRL3_01526 mutant. (a) *Aspergillus niger* strains CBS 138852, ΔNRRL3_01886, and ΔNRRL3_01526 were grown at 30°C for 72 hours on media containing fructose; 3,4-DHB; or 2,3-DHB as sole carbon source. (b) Decrease in 3,4-DHB and (c) accumulation of β-ketoadipate in the ΔNRRL3_01886 mutant in minimal media with 3,4-DHB as sole carbon source. Data from LC-MS in positive mode. AUC = area under curve, calculated by Agilent Profinder software.

Additional growth phenotype tests were performed on 2,4-DHB and gentisic acid (2,5-DHB) as sole carbon sources. Gentisic acid is another of the seven common intermediates beginning the central pathways of aromatic catabolism in bacteria and fungi (Figure 3; Lubbers et al., 2019). This pathway in filamentous fungi is poorly understood, but *A. niger* is able to utilize gentisate as a sole carbon source, and none of the mutations in this study were found to have any effect on this pathway (Figure 15).

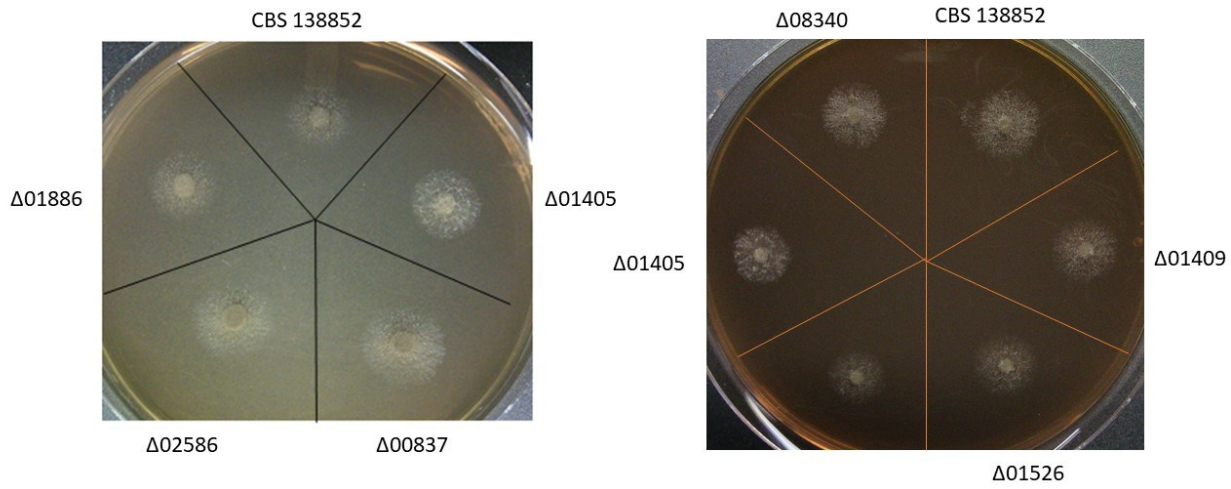

Figure 15. Growth phenotype of knockout mutants generated for each enzyme in the 3,4-DHB pathway on MM with gentisate as sole carbon source. A similar phenotype was observed for all mutants and the parent strain. Strains were grown at 30°C for 72 hours.

Little is known about the 2,4-DHB pathway or the genes involved, especially in filamentous fungi. On 2,4-DHB, *A. niger* CBS 138852 and all mutants grew relatively poorly, but growth was equal to the parent strain in all mutants except for the ΔNRRL3_01886 and ΔNRRL3_01526 mutants, which showed no growth (Figure 16).

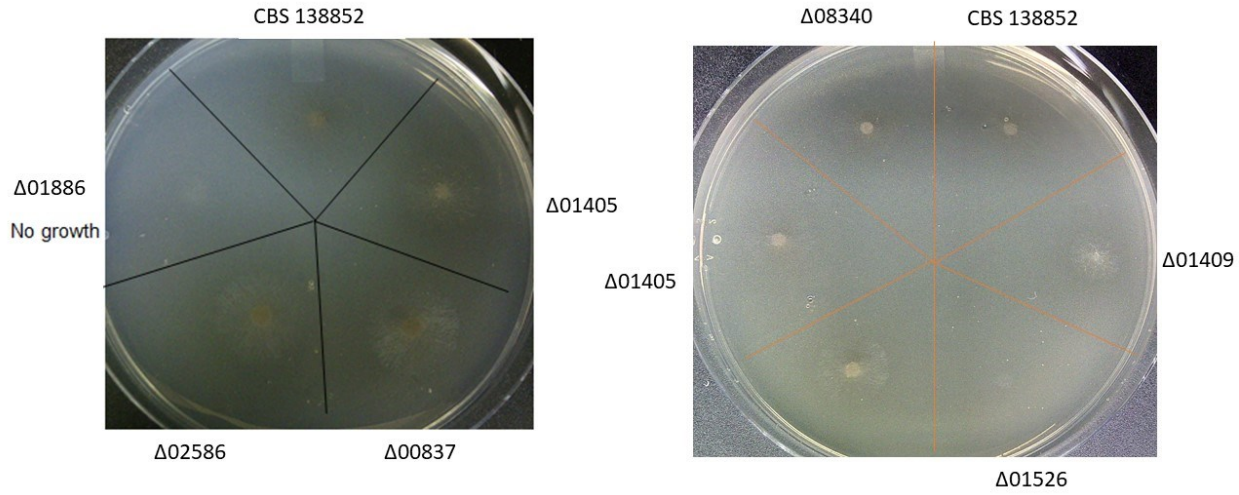

Figure 16. Growth phenotype of knockout mutants generated for each enzyme in the 3,4-DHB pathway on MM with 2,4-DHB as sole carbon source. Strains were grown at 30°C for 72 hours,

3.8 Growth phenotype of quinic acid pathway enzyme mutants

Mutants generated for the three genes predicted to encode the enzymes of the quinic acid pathway, NRRL3_08520, NRRL3_11037 and NRRL3_11035, were grown on minimal media plates with either quinic acid or fructose as the sole carbon source (Figure 17). Growth was equal in all mutants and the parent strain on fructose media. The ΔNRRL3_08520 mutant grew very poorly on quinic acid media compared to the parent, supporting the prediction that this gene encodes the quinate dehydrogenase mutant in *Aspergillus niger*. However, both the predicted dehydroquinate dehydrase mutant and the predicted dehydroshikimate dehydrase mutant, ΔNRRL3_11037 and ΔNRRL3_11035 respectively, showed a phenotype of equal growth to the parent strain on quinic acid media.

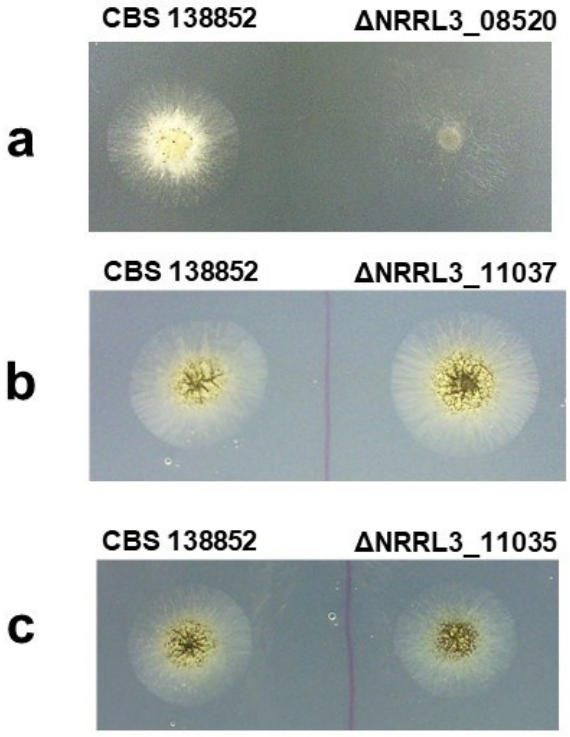

Figure 17. Growth of quinic acid pathway enzyme mutants on MM with 0.5% quinic acid as sole carbon source. The parent strain *A. niger* CBS 138852 is shown on the left in each case next to (a) the predicted quinate dehydrogenase mutant ΔNRRL3_08520, (b) the predicted dehydroquinate dehydrase mutant ΔNRRL3_11037, and (c) the predicted dehydroshikimate dehydrase mutant ΔNRRL3_11035. Strains were grown at 30°C for 72 hours,

3.9 Growth of quinate permease mutants

As previously discussed, the in-cluster *qutD* gene encoding quinate permease in *Aspergillus nidulans* was shown to be required for growth on quinic acid at pH 6.5 but knockout mutants lacking this gene were able to grow on quinic acid at pH 3.5 (Whittington et al., 1987). Transcriptome analysis of *A. niger* showed that the in-cluster ortholog of *qutD*, NRRL3_11036, was upregulated when quinic acid was used as sole carbon source. As well, an off-cluster paralog of NRRL3_11036, NRRL3_05631, was also upregulated when quinic acid was used as sole carbon source. Growth of the three mutants generated for the *A. niger* permease mutants

(ΔNRRL3_11036, ΔNRRL3_05631, and the double mutant lacking both permeases) was tested on quinic acid media and fructose media at pH 6.5 and pH 3.5 (Figure 18). Results for the ΔNRRL3_11036 mutants were consistent with this gene being the ortholog of the *qutD* gene of *A. nidulans*, as predicted by sequence similarity and this gene being located in the cluster. Growth at pH 6.5 was significantly reduced while growth at pH 3.5 remained equal to that of the parent strain. Conversely, the ΔNRRL3_05631 mutant exhibited poor growth at pH 3.5 but grew as well as the parent strain at pH 6.5. The double knockout strain lacking expression of both permeases showed very weak growth at both pH levels and failed to sporulate after 4 days at 30℃. As with the other mutants, growth on media with fructose as the sole carbon source (at both pH levels for the permeases) was used as a control for unexpected general effects on growth. All mutants grew as well as the parent strain on fructose.

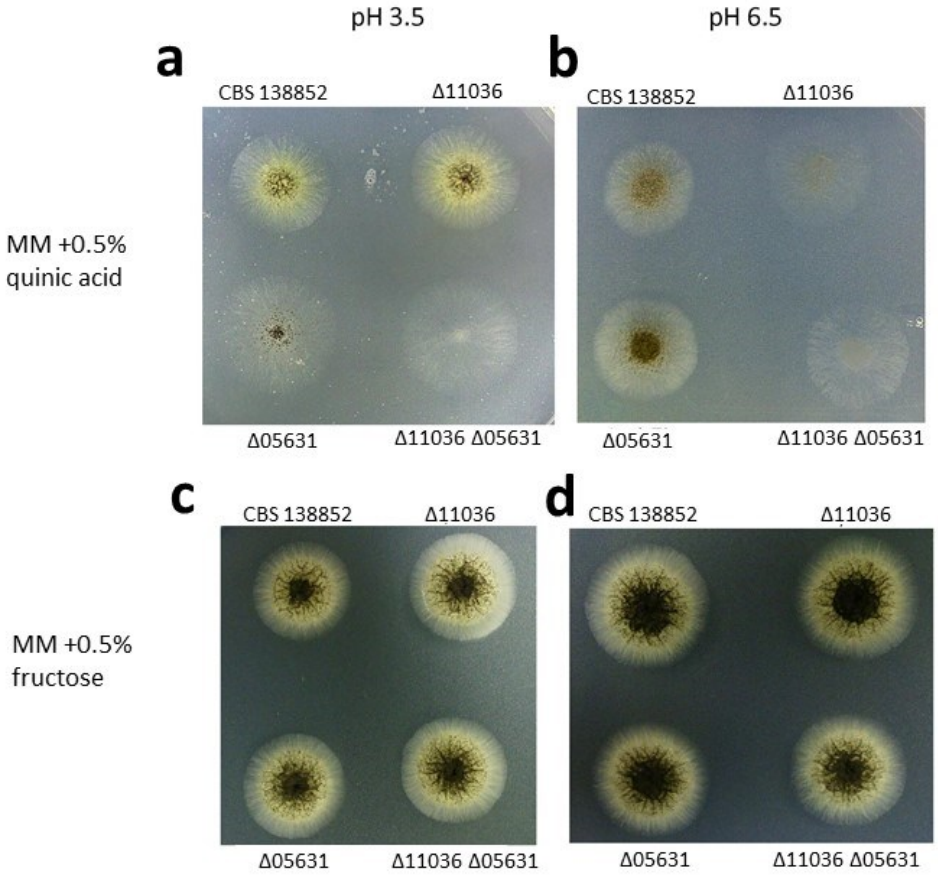

Figure 18. Growth phenotype of CBS 138852 and permease mutants. Colonies were grown for 72 hours at 30℃ on minimal media with 0.5% quinic acid as sole carbon source at pH 3.5 (a) and pH 6.5 (b), or with 0.5% fructose as sole carbon source at pH 3.5 (c) and pH 6.5 (d).

To determine the activity of the two permeases over a more complete range of pH values, a growth assay was performed with these mutants grown on quinic acid media using a microplate reader which measured the optical density at 595 nm over time. The parent strain grew approximately equally well at pH 3.5, 4.5, 5.5, and 6.5, with slightly weaker growth at pH 2.5 (Figure 19A). The mutant strain lacking expression of NRRL3_11036 had optimal growth at pH 3.5, with weaker growth at pH 2.5 and 4.5, but no growth at pH 5.5 or 6.5 (Figure 19B). The NRRL3_05631 mutant also had growth at pH 4.5, showing a slight overlap of function in the two permeases (Figure 19C). This mutant grew optimally at pH 6.5, with growth at pH 5.5 and 4.5 as well, and little to no growth at pH 3.5 or 2.5. The double mutant (Figure 19D) failed to grow sufficiently at any pH level.

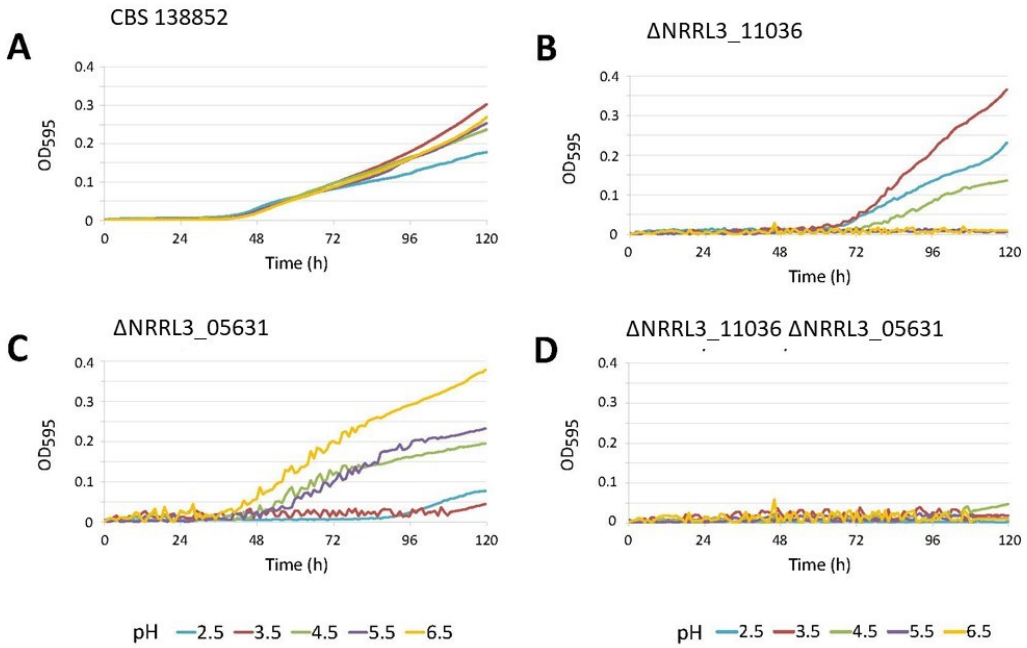

Figure 19. Growth profile of permease mutants in microplate on quinic acid media at various pH levels. Each sample represents the average of triplicates. Measurements taken every hour for 120 hours.

4. Discussion

4.1 Quinic acid pathway genes

The cluster of genes in *A. niger* between NRRL3_11035 and NRRL3_11039, and the genes NRRL3_08520, NRRL3_08521, and NRRL3_10284, are the orthologs of the *A. nidulans* quinic acid utilization gene cluster which converts quinic acid to 3,4-DHB. There is strong evidence that these genes are involved conversion of quinic acid to 3,4-DHB in *A. niger* as well. In support of this, the RNA sequencing data from *A. niger* CBS 138852 grown on quinic acid media showed high levels of expression of these genes and a high degree of upregulation compared to the same strain grown on fructose media. A knockout mutant lacking expression of the activator in this cluster, NRRL3_11038 failed to grow on quinic acid media (unpublished data of Dr. Arthur Ram of Leiden University). However, our results were not as clear as what has been observed in *A. nidulans.* Although knockout mutation in the NRRL3_08520 gene predicted to encode the first enzyme in the pathway, quinate dehydrogenase, inhibited growth on quinic acid as expected, knockout mutations in genes predicted to encode the next two enzymes in the pathway, dehydroquinate dehydrase and dehydroshikimate dehydrase (NRRL3_11037 and NRRL3_11035, respectively) have no effect on growth. One possibility is that the pentafunctional AROM polypeptide of the shikimate pathway is compensating for the loss of the dehydroquinate dehydrase. As previously mentioned, the shikimate pathway also includes the conversion of dehydroquinic acid to dehydroshikimic acid. The pentafunctional AROM polypeptide, NRRL3_11273, was found to be expressed on both quinic acid and fructose, as expected due to the importance of the shikimate pathway. Although expression was higher on fructose, it may be expressed at sufficient levels on quinic acid to compensate for the loss of NRRL3_11037. In *Aspergillus nidulans,* a mutant lacking the quinic acid cluster dehydroquinate dehydrase gene was shown to be able to grow on quinic acid as the sole carbon source, however this required 5-fold overproduction of the AROM polypeptide (Lamb et al., 1991). Another

possibility is that genes from the cluster between NRRL3_07600 and NRRL3_07605, discussed in section 3.2, or other genes have a redundant function to the dehydroshikimate dehydrase, or both of these enzymes. Further experiments including generation of mutants with knockouts in multiple genes will be required.

A knockout mutation in the quinate permease NRRL3_11036, which transports quinic acid into the cells, was found to have an effect on growth consistent with that observed in *A. nidulans.* This mutation inhibited growth at pH levels above 4.5, with optimal growth at pH 3.5 The newly identified second permease NRRL3_05631 was found to be required for growth on quinic acid at lower pH, below 4.5, with optimal growth at pH 6.5. As previously mentioned, Whittington et al. (1987), who identified QutD, the NRRL3_11036 ortholog, in *A. nidulans* proposed that the protonated quinic acid molecule may be entering the cell by diffusion through the membrane while the quinate ion required the transporter. However, NRRL3_05631 has an ortholog in *A. nidulans* (AN6805.2 in strain FGSC A4). This gene encodes a 549 amino acid protein with 78% identity over the length of the NRRL3_05631 protein sequence. The *A. nidulans* gene is also located outside of the quinic acid gene cluster. This second permease is likely responsible for the growth observed at pH 3.5 in *A. nidulans*. The double mutant in *A. niger* only grew very poorly on quinic acid at any pH, showing diffusion of quinic acid through the membrane is not sufficient for growth, at least in this species.

4.2 3,4-dihydroxybenzoic acid catabolic pathway genes

The results from the knockout mutations showed that the predictions made for the genes of the 3,4-DHB pathway appear to be correct in all cases where a single gene was predicted. The mutation in the protocatechuate-3,4-dioxygenase, NRRL3_01405 prevented growth on 3,4-DHB media and accumulated 3,4-DHB when grown on quinic acid media, as expected of a mutant unable to catabolize quinic acid beyond that step of the pathway. When combined, the results confirm the prediction that NRRL3_01405 encodes the protocatechuate-3,4-dioxygenase was correct. Similarly, the results appear to confirm that NRRL3_02586 encodes 3-carboxy-*cis,cis*-muconate cyclase. This mutant was also unable to grow on 3,4-DHB media and accumulated 3-carboxy-*cis,cis*-muconic acid when grown on quinic acid media.

The enzyme with the least clear prediction based on sequence similarity and functional annotation was the 3-carboxymuconolactone hydrolase/decarboxylase. Of the three genes predicted as possible candidates, NRRL3_01409, NRRL3_00837, and NRRL3_08340, two are required for growth on 3,4-DHB media. The ΔNRRL3_08340 mutant had no apparent effect on growth and may be rejected as a candidate for this enzyme. It is unclear why this gene was so highly expressed on 3,4-DHB.

From these results, it remains unclear which gene, NRRL3_01409 or NRRL3_00837, encodes the 3-carboxymuconolactone hydrolase/decarboxylase enzyme. Neither gene accumulated 3-carboxymuconolactone when grown on quinic acid. However, enzymatic data from a collaboration with the Powlowski Lab at Concordia University has shown the enzyme activity of many of these genes. These biochemical data support the results obtained from mutant analyses for the protocatechuate-3,4-dioxygenase and the 3-carboxy-cis,cis-muconate cyclase, i.e. that the genes NRRL3_01405 and NRRL3_02586, respectively, encode these enzymes (Semana and Powlowski, 2019; Powlowski and Olyaie (personal communication)). The enzyme activity data also shows that NRRL3_01409 is the 3-carboxymuconolactone hydrolase/decarboxylase (Powlowski and Olyaie (personal communication)). This is consistent with the mutant results observed in this study. NRRL3_00837 appears to play an essential role in this pathway as deletion of this gene resulted in near absence of growth on 3,4-DHB, but the sequence of the short protein encoded is not similar enough to a characterized protein to be able to predict a function with any certainty.

The predicted β-ketoadipate-CoA transferase and β-ketoadipyl-CoA transferase encoding genes, NRRL3_01886 and NRRL3_01526 respectively, also appear to be correctly predicted based on the growth phenotype and LC-MS results. These two genes form part of the pathway which is shared by the catabolic pathways of several of the other aromatic pathway central intermediates, including catechol and hydroxyquinol (Figure 3). Evidence of this was also seen by the lack of growth of these mutants on 2,3-dihydroxybenzoic acid and 2,4-dihydroxybenzoic acid. The pathway of 2,4-DHB degradation remains unclear in fungi. In bacteria under anaerobic conditions, 2,4-DHB can be converted to resorcinol (Tschech 1985). If 2,4-DHB is converted to

resorcinol in *A. niger,* it could then be degraded through the hydroxyquinol branch of the β-ketoadipate pathway (Shailubhai et al., 1983a). This would involve ring cleavage of the hydroxyquinol to produce maleylacetate, which can then be converted to β-ketoadipate. These results support the idea that 2,4-DHB is being catabolized through hydroxyquinol, the only other known major branch of the β-ketoadipate pathway (Figure 3). However, recent results suggest that a knockout mutant of the predicted hydroxyquinol ring fission enzyme, NRRL3_05330, is still able to grow on 2,4-DHB (not shown). More research will be needed to determine if 2,4-DHB is catabolized through a separate pathway from hydroxyquinol, 3,4-DHB, and catechol, or if NRRL3_05330 is the incorrect gene prediction for this enzyme.

Mutations in NRRL3_01886 and NRRL3_01526, and all other genes in this study, were shown to have no effect on growth of *A. niger* CBS 138852 on 2,5-dihydroxybenzoic acid. In bacteria, 2,5-DHB is cleaved by gentisate-1,2-dioxygenase to form maleylpyruvate, which is converted to fumarylpyruvate by maleylpyruvate isomerase. Fumarylpyruvate is hydrolyzed by fumarylpyruvate hydrolase to form fumarate and pyruvate which enter the TCA cycle (Liu and Zhou, 2012). The same pathway has been shown to be present in the yeast *Candida parapsilosis* (Holesova et al., 2011). If this pathway is used by *A. niger* it would be expected that NRRL3_01886 and NRRL3_01526 are not involved.

5. Conclusions and future work

The genes involved in the quinic acid catabolic pathway in *A. niger* are less easy to confirm than initially predicted. Although *A. niger* contains the quinic acid utilization gene cluster found in other fungi, only one of the genes encoding enzymes, the quinate dehydrogenase gene NRRL3_08520, was required for growth on quinic acid. Knockout mutants of the dehydroquinate dehydrase and dehydroshikimate dehydrase genes, NRRL3_11037 and NRRL3_11035 respectively, still grew on quinic acid. Although it is likely these genes are involved, based on sequence similarity and RNA sequencing data from quinic acid media, there may be redundant genes in *A. niger*, such as the pentafunctional AROM polypeptide. Further testing will be required, starting with the NRRL3_07600 to NRRL3_07605 cluster.

The quinate permease found in the quinate utilization gene cluster, NRRL3_11036, was found to be required for growth on quinic acid at pH 6.5 but not at pH 3.5, consistent with what has been observed in *A. nidulans*. However, a new quinate permease was identified and characterized for the first time in this study. NRRL3_05631, which has an uncharacterized ortholog in *A. nidulans*, was found to be a second quinate permease required for transport of quinic acid into the cell at low pH. Knockout mutants lacking both of these genes showed minimal growth on quinic acid.

The central pathways that make up the final stages of aromatic compound catabolism remain very poorly understood in filamentous fungi. This study is the beginning of an extensive examination of the genes of these pathways. Although a few recent studies have begun to assign genes to some of these enzymes in fungi, they are spread among different fungi, including yeasts which may use different pathways or enzymes than filamentous fungi in some cases. The results of this study have clearly identified the genes in the 3,4-DHB branch of the β-ketoadipate pathway in *A. niger* for the first time. Using ortholog identification and comparative transcriptomics to predict the genes encoding these enzymes have shown to be a successful strategy, with four of the five enzymes being correctly predicted, as shown by the growth phenotype and LC-MS analysis. NRRL3_01405 was shown to be the gene encoding protocatechuate-3,4-dioxygenase, with the knockout mutant exhibiting poor growth on 3,4-DHB media and accumulation of this intermediate in the extracellular media when grown on its precursor quinic acid. Similarly, NRRL3_02586 can be assigned as the enzyme 3-carboxy-cis,cis-muconate cyclase. Knocking out this gene prevented any growth on 3,4-DHB and accumulation of this intermediate was also seen when grown on quinic acid. The β-ketoadipate-CoA transferase enzyme is shared by multiple branches of the β-ketoadipate pathway, as the 3,4-DHB, catechol and hydroxyquinol branches converge. This enzyme was correctly predicted to be NRRL3_01886, and the mutant failed to grow on both 3,4-DHB and 2,3-DHB as expected for the shared steps of the pathway. The mutant also appears to accumulate the intermediate when grown on 3,4-DHB. The final enzyme in the pathway, β-ketoadipyl-CoA transferase, was also correctly predicted to be NRRL3_01526. Mutants lacking expression of this gene were also unable to grow on 3,4-DHB or 2,3-DHB.

Although the results were less clear, the enzyme 3-carboxy-*cis,cis*-muconate hydrolase/decarboxylase could be the NRRL3_01409 gene based on the molecular weight of the gene product and growth phenotype of the knockout. Enzyme activity assays have confirmed this assignment (Powlowski and Olyaie, unpublished). Most of the remaining work on understanding this specific pathway will involve attempting to discover the true functions of NRRL3_00837 and NRRL3_08340. Although the knockout of NRRL3_08340 had no effect on growth on 3,4-DHB, the gene was very highly upregulated and expressed when *A. niger* CBS 138852, and likely has a role in this pathway or a similar pathway. NRRL3_00837 had been predicted or even assigned as the 3-carboxy-*cis,cis*-muconate hydrolase/decarboxylase by previous studies. It was shown to be a necessary gene as the knockout mutant failed to grow on 3,4-DHB, however the role of this protein is also currently unknown.

The methods used in this study can be applied to the catabolic pathways of all of the seven common ring fission intermediates in *A. niger*. This will be the first time the genes and enzymes of very important network of pathways will be identified and characterized. This knowledge can then be used for many future studies. Using sequence similarity to the *A. niger* genes, the orthologs in other species of fungi can be easily identified. This will not only be useful for further studies involving specific pathways in other species but will also provide a quick method of discovering the distribution of each of the pathways among fungi, as all seven are not present in all species. Understanding these pathways will lead to many future areas of research in the *Aspergilli* and other fungi, including improving tolerance of aromatic compounds during fermentation, bioremediation of toxic aromatic compounds, and production of valuable products from plant sources.

References

Abbas, M., Saeed, F., Anjum, F. M., Afzaal, M., Tufail, T., Bashir, M. S., Ishtiaq, A., Hussain, S., & Suleria, H. A. R. (2017). Natural polyphenols: An overview. *International Journal of Food Properties*, *20*(8), 1689-1699.

Abejón, R., Pérez-Acebo, H., & Clavijo, L. (2018). Alternatives for chemical and biochemical lignin valorization: hot topics from a bibliometric analysis of the research published during the 2000–2016 period. *Processes*, *6*(8), 98.

Adamczyk, B., Simon, J., Kitunen, V., Adamczyk, S., & Smolander, A. (2017). Tannins and their complex interaction with different organic nitrogen compounds and enzymes: old paradigms versus recent advances. *ChemistryOpen*, *6*(5), 610-614.

Adeboye, P. T., Bettiga, M., & Olsson, L. (2014). The chemical nature of phenolic compounds determines their toxicity and induces distinct physiological responses in Saccharomyces cerevisiae in lignocellulose hydrolysates. *Amb Express*, *4*(1), 1-10.

Aguilar, C. N., Rodríguez, R., Gutiérrez-Sánchez, G., Augur, C., Favela-Torres, E., Prado-Barragan, L. A., ... & Contreras-Esquivel, J. C. (2007). Microbial tannases: advances and perspectives. *Applied microbiology and biotechnology*, *76*(1), 47-59.

Aguilar-Pontes M.V., Brandl, J., McDonnell, E., Strasser, K., Nguyen, T.T.M., Riley, R., Mondo, S., Salamov, A., Nybo. J.L., Vesth, T.C., Grigoriev, I.V., Andersen, M.R., Tsang, A., de Vries R.P. (2018). The gold-standard genome of *Aspergillus niger* NRRL 3 enables a detailed view of the diversity of sugar catabolism in fungi. *Stud Mycol 91*:61–78. https://doi.org/10.1016/j.simyco.2018.10.001.

Aleksenko, A., & Clutterbuck, A. J. (1996). The plasmid replicator AMA1 in Aspergillus nidulans is an inverted duplication of a low-copy-number dispersed genomic repeat. *Molecular microbiology*, *19*(3), 565-574. https://doi.org/10.1046/j.1365-2958.1996.400937.x

Andersen, M. R., Salazar, M. P., Schaap, P. J., Van De Vondervoort, P. J., Culley, D., Thykaer, J., ... & Baker, S.E. (2011). Comparative genomics of citric-acid-producing Aspergillus niger ATCC 1015 versus enzyme-producing CBS 513.88. *Genome research*.

Arnett, D. R., Lorimer, H. E., & Asch, D. K. (2009). Catabolite repression directly affects transcription of the qa-y gene of Neurospora crassa. *Fungal genetics and biology*, *46*(5), 377-380.

Aslanidis, C., & De Jong, P. J. (1990). Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic acids research*, *18*(20), 6069-6074.

Baker, S. E. (2006). Aspergillus niger genomics: past, present and into the future. *Medical mycology*, *44*(Supplement_1), S17-S21.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., ... & Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, *315*(5819), 1709-1712.

Beckham, G. T., Johnson, C. W., Karp, E. M., Salvachúa, D., & Vardon, D. R. (2016). Opportunities and challenges in biological lignin valorization. *Current Opinion in Biotechnology*, *42*, 40–53. https://doi.org/10.1016/j.copbio.2016.02.030

Bertani, G. (1951). Studies on lysogenesis I.: the mode of phage liberation by lysogenic Escherichia coli1. *Journal of bacteriology*, *62*(3), 293.

Bhat, T. K., Singh, B., & Sharma, O. P. (1998). Microbial degradation of tannins–a current perspective. *Biodegradation*, *9*(5), 343-357.

Boerjan, W., Ralph, J., & Baucher, M. (2003). Lignin Biosynthesis. *Annual Review of Plant Biology*, *54*(1), 519–546. https://doi.org/10.1146/annurev.arplant.54.031902.134938

Bos, C. J., Debets, A. J. M., Swart, K., Huybers, A., Kobus, G., & Slakhorst, S. M. (1988). Genetic analysis and the construction of master strains for assignment of genes to six linkage groups in Aspergillus niger. *Current genetics*, *14*(5), 437-443.

Boudet, A. M., Kajita, S., Grima-Pettenati, J., & Goffner, D. (2003). Lignins and lignocellulosics: a better control of synthesis for new and improved uses. *Trends in plant science*, *8*(12), 576-581.

Cain, R.B., Bilton, R.F., & Darrah, J.A. (1968). The metabolism of aromatic acids by micro-organisms. Metabolic pathways in the fungi. *Biochem J* 108, 797–828. https://doi.org/10.1042/bj1080797.

Cao, B., Nagarajan, K., & Loh, K. C. (2009). Biodegradation of aromatic compounds: current status and opportunities for biomolecular approaches. *Applied microbiology and biotechnology*, *85*(2), 207-228.

Cánovas, J.L., & Stanier, R.Y. (1967). Regulation of the enzymes of the beta-ketoadipate pathway in *Moraxella calcoacetica*. 1. General aspects. *Eur J Biochem. 1(3)*, 289-300.

Debets, A. J. M., & Bos, C. J. (1986). Isolation of small protoplasts from *Aspergillus niger*. *Fungal Genetics Reports*, *33*(1), 24.

Diaz, E., Jimenez J.I., & Nogales J. (2013). Aerobic degradation of aromatic compounds. *Curr Opin Biotechnol,* 24, 431– 442. https://doi.org/10.1016/j.copbio.2012.10.010

Doench, J. G., Hartenian, E., Graham, D. B., Tothova, Z., Hegde, M., Smith, I., ... & Root, D. E. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation. *Nature biotechnology*, *32*(12), 1262. https://doi.org/10.1038/nbt.3026

Dykes, L., & Rooney, L. W. (2007). Phenolic compounds in cereal grains and their health benefits. *Cereal foods world*, *52*(3), 105-111.

Ellouze, M., & Sayadi, S. (2016). White-rot fungi and their enzymes as a biotechnological tool for xenobiotic bioremediation. *Management of hazardous wastes. InTech, London*, 103-120.

Giles, N. H., Geever, R. F., Asch, D. K., Avalos, J., & Case, M. E. (1991). Organization and regulation of the qa (quinic acid) genes in Neurospora crassa and other fungi. *Journal of Heredity*, *82*(1), 1-7.

Goosen, T., Bloemheuvel, G., Gysler, C., de Bie, D. A., van den Broek, H. W., & Swart, K. (1987). Transformation of *Aspergillus niger* using the homologous orotidine-5′-phosphate-decarboxylase gene. *Current genetics*, *11*(6-7), 499-503.

Gross, G. G. (2008). From lignins to tannins: Forty years of enzyme studies on the biosynthesis of phenolic compounds. *Phytochemistry*, *69*(18), 3018-3031.

Grund, E., & Kutzner, H. J. (1998). Utilization of quinate and p-hydroxybenzoate by actinomycetes: key enzymes and taxonomic relevance. *Journal of Basic Microbiology: An International Journal on Biochemistry, Physiology, Genetics, Morphology, and Ecology of Microorganisms*, *38*(4), 241-255.

Guo J., Carrington Y., Alber A., and Ehlting J. (2014). Molecular Characterization of Quinate and Shikimate Metabolism in *Populus trichocarpa*. J. Biol. Chem. 289(34): 23846-23858.

Hawkins, A. R., Lamb, H. K., Moore, J. D., Charles, I. G., & Roberts, C. F. (1993). The pre-chorismate (shikimate) and quinate pathways in filamentous fungi: theoretical and practical aspects. *Microbiology*, *139*(12), 2891–2899.

Harwood, C. S., & Parales, R. E. (1996). The β-ketoadipate pathway and the biology of self-identity. *Annual Reviews in Microbiology*, *50*(1), 553–590.

Hatakka, A., & Hammel, K. E. (2011). Fungal biodegradation of lignocelluloses. In *Industrial applications* (pp. 319-340). Springer, Berlin, Heidelberg.

Herrmann KM., and Weaver LM. (1999). The shikimate pathway. Annu. Rev. Plant Physiol. Plant Mol. Biol. 50: 473–503.

Herrmann, K. M. (1995). The shikimate pathway as an entry to aromatic secondary metabolism. *Plant Physiology*, *107*(1), 7.

Hill, T. W., & Kafer, E. (2001). Improved protocols for Aspergillus minimal medium: trace element and minimal medium salt stock solutions. *Fungal Genetics Reports*, *48*(1), 20-21.

Holesova, Z., Jakubkova, M., Zavadiakova, I., Zeman, I., Tomaska, L., & Nosek, J. (2011). Gentisate and 3-oxoadipate pathways in the yeast *Candida parapsilosis:* identification and functional analysis of the genes coding for 3-hydroxybenzoate 6-hydroxylase and 4-hydroxybenzoate 1-hydroxylase. *Microbiology*, *157*(7), 2152–2163. https://doi.org/10.1099/mic.0.048215-0

Huiet, L. (1984). Molecular analysis of the Neurospora qa-1 regulatory region indicates that two interacting genes control qa gene expression. *Proceedings of the National Academy of Sciences*, *81*(4), 1174-1178.

Hutner, S. H., Provasoli, L., Schatz, A., & Haskins, C. P. (1950). Some approaches to the study of the role of metals in the metabolism of microorganisms. *Proceedings of the American Philosophical Society*, *94*(2), 152-170.

Jiang, W., Bikard, D., Cox, D., Zhang, F., & Marraffini, L. A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature biotechnology*, *31*(3), 233.

Jung, Y. H., & Kim, K. H. (2015). Acidic pretreatment. In *Pretreatment of Biomass* (pp. 27-50). Elsevier.

Käfer, E. (1977). Meiotic and Mitotic Recombination in Aspergillus and Its Chromosomal Aberrations, in *Advances in Genetics Volume 19*. Elsevier BV. p. 33-131.

Kallscheuer, N., Gätgens, J., Lübcke, M., Pietruszka, J., Bott, M., & Polen, T. (2017). Improved production of adipate with Escherichia coli by reversal of β-oxidation. *Applied microbiology and biotechnology*, *101*(6), 2371-2382.

Kamimura, N., Takahashi, K., Mori, K., Araki, T., Fujita, M., Higuchi, Y., & Masai, E. (2017). Bacterial catabolism of lignin-derived aromatics: New findings in a recent decade: Update on bacterial lignin catabolism. *Environmental microbiology reports*, *9*(6), 679-705.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., & Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*(12), 1647-1649.

Kelly, J. M., Drysdale, M. R., Sealy-Lewis, H. M., Jones, I. G., & Lockington, R. A. (1990). Alcohol dehydrogenase III inAspergillus nidulans is anaerobically induced and post-transcriptionally regulated. *Molecular and General Genetics MGG*, *222*(2-3), 323-328.

Kun, R. S., Meng, J., Salazar-Cerezo, S., Mäkelä, M. R., de Vries, R. P., & Garrigues, S. (2020). CRISPR/Cas9 facilitates rapid generation of constitutive forms of transcription factors in Aspergillus niger through specific on-site genomic mutations resulting in increased saccharification of plant biomass. *Enzyme and Microbial Technology*, *136*, 109508.

Lamb, H. K., Bagshaw, C. R., & Hawkins, A. R. (1991). In vivo overproduction of the pentafunctional arom polypeptide in Aspergillus nidulans affects metabolic flux in the quinate pathway. *Molecular and General Genetics MGG*, *227*(2), 187-196.

Lamb, H. K., Roberts, C. F., & Hawkins, A. R. (1992). A second gene (qutH) within the Aspergillus nidulans-quinic-acid utilisation gene cluster encodes a protein with a putative zinc-cluster motif. Gene, 112(2), 219-224.

Larsson, S., Quintana-Sáinz, A., Reimann, A., Nilvebrant, N. O., & Jönsson, L. J. (2000). Influence of lignocellulose-derived aromatic compounds on oxygen-limited growth and ethanolic fermentation by Saccharomyces cerevisiae. In *Twenty-First Symposium on Biotechnology for Fuels and Chemicals* (pp. 617-632). Humana Press, Totowa, NJ.

Levesley I. (2013). Molecular characterisation of mutations in the qutA activator gene for quinic acid utilisation in *Aspergillus nidulans* (Doctoral thesis). ProQuest/UMI Number: U098061.

Liu, T. T., & Zhou, N. Y. (2012). Novel L-cysteine-dependent maleylpyruvate isomerase in the gentisate pathway of Paenibacillus sp. strain NyZ101. *Journal of bacteriology*, *194*(15), 3987-3994.

Lubbers, R. J. M., Dilokpimol, A., Visser, J., Mäkelä, M. R., Hildén, K. S., & de Vries, R. P. (2019). A comparison between the homocyclic aromatic metabolic pathways from plant-derived compounds by bacteria and fungi. *Biotechnology Advances*. https://doi.org/10.1016/j.biotechadv.2019.05.002

Mäkelä, M. R., Marinović, M., Nousiainen, P., Liwanag, A. J. M., Benoit, I., Sipilä, J., … Hildén, K. S. (2015). Aromatic Metabolism of Filamentous Fungi in Relation to the Presence of Aromatic Compounds in Plant Biomass. In *Advances in Applied Microbiology* (Vol. 91, pp. 63–137). https://doi.org/10.1016/bs.aambs.2014.12.001

Marín, M., Pérez-Pantoja, D., Donoso, R., Wray, V., González, B., & Pieper, D. H. (2010). Modified 3-oxoadipate pathway for the biodegradation of methylaromatics in Pseudomonas reinekei MT1. *Journal of bacteriology*, *192*(6), 1543-1552.

Martins, T. M., Hartmann, D. O., Planchon, S., Martins, I., Renaut, J., & Silva Pereira, C. (2015). The old 3-oxoadipate pathway revisited: New insights in the catabolism of aromatics in the saprophytic fungus Aspergillus nidulans. *Fungal Genetics and Biology*, *74*, 32–44. https://doi.org/10.1016/j.fgb.2014.11.002

Master, E. R., Zheng, Y., Storms, R., Tsang, A., & Powlowski, J. (2008). A xyloglucan-specific family 12 glycosyl hydrolase from Aspergillus niger: recombinant expression, purification and characterization. *Biochemical Journal*, *411*(1), 161-170.

Mazur, P., Henzel, W.J., Mattoo, S., & Kozarich, J.W. (1994) 3-Carboxy-*cis,cis*-muconate lactonizing enzyme from *Neurospora crassa*: an alternate cycloisomerase motif. *J Bacteriol. 176(6)*,1718-28.

Metcalf, R. L., & Kogan, M. (1987). Plant volatiles as insect attractants. *Critical reviews in plant sciences*, *5*(3), 251-301.

Meyer, V., Ram, A. F., & Punt, P. J. (2010). Genetics, genetic manipulation, and approaches to strain improvement of filamentous fungi. In *Manual of Industrial Microbiology and Biotechnology, Third Edition* (pp. 318-329). American Society of Microbiology.

Milstein, O., Trojanowski, J., Hüttermann, A., & Gressel, J. (1988). Catabolism of single ring aromatic acids by four Aspergillus species. *Microbios*, *55*(222), 7-16.

Miura, Y., Wake, H., & Kato, T. (1999). TBE, or not TBE; that is the question: Beneficial usage of tris-borate for obtaining a higher resolution of small DNA fragments by agarose gel electrophoresis. *Nagoya medical journal*, *43*(1), 1-6.

New England Biolabs. High Efficiency Transformation Protocol (C2987H/C2987I). Retrieved March 3, 2020 from: https://international.neb.com/protocols/0001/01/01/high-efficiency-transformation-protocol-c2987

Nie L., and Shi X. (2009). A novel asymmetric synthesis of oseltamivir phosphate (Tamiflu) from (−)-shikimic acid. Tetrahedron: Asymmetry 20: 124–129.

Niku-Paavola, M. L., Karhunen, E., Salola, P., & Raunio, V. (1988). Ligninolytic enzymes of the white-rot fungus Phlebia radiata. *Biochemical Journal*, *254*(3), 877-884.

Niu, J., Alazi, E., Reid, I. D., Arentshorst, M., Punt, P. J., Visser, J., ... & Ram, A. F. (2016). An evolutionarily conserved transcriptional activator-repressor module controls expression of genes for d-galacturonic acid utilization in Aspergillus niger. *Genetics*, genetics-116.

Nødvig, C. S., Nielsen, J. B., Kogle, M. E., & Mortensen, U. H. (2015). A CRISPR-Cas9 system for genetic engineering of filamentous fungi. *PloS one*, *10*(7).

Nødvig, C. S., Hoof, J. B., Kogle, M. E., Jarczynska, Z. D., Lehmbeck, J., Klitgaard, D. K., & Mortensen, U. H. (2018). Efficient oligo nucleotide mediated CRISPR-Cas9 gene editing in Aspergilli. *Fungal Genetics and Biology*, *115*, 78-89.

Pagare, S., Bhatia, M., Tripathi, N., Pagare, S., & Bansal, Y. K. (2015). Secondary metabolites of plants and their role: Overview. *Current Trends in Biotechnology and Pharmacy*, *9*(3), 293-304.

Pao, S. S., Paulsen, I. T., & Saier, M. H. (1998). Major facilitator superfamily. *Microbiology and molecular biology reviews*, *62*(1), 1-34.

Pel, H. J., de Winde, J. H., Archer, D. B., Dyer, P. S., Hofmann, G., Schaap, P. J., ... & Andersen, M. R. (2007). Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. *Nature biotechnology*, *25*(2), 221.

Perdana, A. T., Arianata, M., & Larasati, T. R. D. (2019). Optimization of Benzene and Toluene Biodegradation by Aspergillus niger and Phanerochaete chrysosporium. *Jurnal Al-Azhar Indonesia Seri Sains dan Teknologi*, *5*(2), 87-91.

Prior, R. L., & Gu, L. (2005). Occurrence and biological significance of proanthocyanidins in the American diet q. *Phytochemistry*, *66*, 2264-2280.

Ragauskas, A. J., Beckham, G. T., Biddy, M. J., Chandra, R., Chen, F., Davis, M. F., … Wyman, C. E. (2014). Lignin Valorization: Improving Lignin Processing in the Biorefinery. *Science*, *344*(6185), 1246843. https://doi.org/10.1126/science.1246843

Sahasrabudhe, S. R., Amin, A. R., & Modi, V. V. (1985). Transformation of chlorinated benzoates and other benzene derivatives by Aspergillus niger and Aspergillus japonicus. *Applied microbiology and biotechnology*, *21*(6), 365-367.

Sander, J. D., & Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature biotechnology*, *32*(4), 347.

Schoenherr, S., Ebrahimi, M., & Czermak, P. (2018). Lignin Degradation Processes and the Purification of Valuable Products. In *Lignin-Trends and Applications*, pp. 29-63.

Semana, P., & Powlowski, J., (2019) Four aromatic intradiol ring cleavage dioxygenases from *Aspergillus niger*. *Appl Environ Microbiol*, 85:e01786-19. https://doi.org/10.1128/AEM.01786-19.

Shailubhai, K., Somayaji, R., Rao, N. N., & Modi, V. V. (1983a). Metabolism of resorcinol and salicylate in Aspergillus niger. *Experientia*, *39*(1), 70-72.

Shailubhai, K., Sahasrabudhe, S. R., Vora, K. A., & Modi, V. V. (1983b). Degradation of chlorinated derivatives of phenoxyacetic acid and benzoic acid by Aspergillus niger. *FEMS Microbiology Letters*, *18*(3), 279-282.

Singh, O. V., & Jain, R. K. (2003). Phytoremediation of toxic aromatic pollutants from soil. *Applied microbiology and biotechnology*, *63*(2), 128-135.

Sista Kameshwar, A. K., & Qin, W. (2018). Comparative study of genome-wide plant biomass-degrading CAZymes in white rot, brown rot and soft rot fungi. *Mycology*, *9*(2), 93-105.

Song, L., Ouedraogo, J. P., Kolbusz, M., Nguyen, T. T. M., & Tsang, A. (2018). Efficient genome editing using tRNA promoter-driven CRISPR/Cas9 gRNA in *Aspergillus niger*. *PloS one*, *13*(8), e0202868

Song, L., Tsang, A. (2017). *Protocol of CRISPR Plasmid Construction*; Version 1.1; Concordia University: Center for Structural and Functional Genomics. Unpublished protocol.

Sugiyama, A., Linley, P. J., Sasaki, K., Kumano, T., Yamamoto, H., Shitan, N., ... & Terakawa, T. (2011). Metabolic engineering for the production of prenylated polyphenols in transgenic legume plants using bacterial and plant prenyltransferases. *Metabolic engineering*, *13*(6), 629-637.

Sun, Z., Fridrich, B., de Santi, A., Elangovan, S., & Barta, K. (2018). Bright Side of Lignin Depolymerization: Toward New Platform Chemicals. *Chemical Reviews*, *118*(2), 614–678. https://doi.org/10.1021/acs.chemrev.7b00588

Tschech, A., & Schink, B. (1985). Fermentative degradation of resorcinol and resorcylic acids. *Archives of microbiology*, *143*(1), 52-59.

Thatcher D.R., & Cain, R.B., (1970). Metabolism of aromatic compounds by fungi: conversion of beta-carboxymuconolactone into 3-oxoadipate in *Aspergillus niger*. *Biochem J 120*:28P–29P. https://doi.org/10.1042/bj1200028pb.

Thatcher, D.R., Cain, R.B., (1974). Metabolism of aromatic compounds by fungi. 1. Purification and physical properties of 3-carboxy-cis-cis-muconate cyclase from *Aspergillus niger*. *Eur J Biochem. 48*(2), 549-56. https://doi.org/10.1111/j.1432-1033.1974.tb03796.x

Thatcher, D.R., Cain, R.B., (1975). Metabolism of aromatic compounds by fungi. Kinetic properties and mechanism of 3-carboxy-*cis,cis*-muconate cyclase from *Aspergillus niger*. *Eur J Biochem* 56(1),193-204. https://doi.org/10.1111/j.1432-1033.1975.tb02222.x

Tong, Z., Zheng, X., Tong, Y., Shi, Y. C., & Sun, J. (2019). Systems metabolic engineering for citric acid production by Aspergillus niger in the post-genomic era. *Microbial cell factories*, *18*(1), 28.

Tzin, V., & Galili, G. (2010). New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants. *Molecular plant*, *3*(6), 956-972.

Vaillancourt, F.H., Bolin, J.T., & Eltis, L.D. (2006). The ins and outs of ring-cleaving dioxygenases. *Crit Rev Biochem Mol Biol* 41, 241–267. https://doi.org/10.1080/10409230600817422.

Valone, JA; Case, ME; and Giles, NH (1971). Constitutive Mutants in a Regulatory Gene Exerting Positive Control of Quinic Acid Catabolism in *Neurospora crassa*. *Proc. Nat. Acad. Sci. USA*, 68/7, 1555-1559.

Vardon, D. R., Franden, M. A., Johnson, C. W., Karp, E. M., Guarnieri, M. T., Linger, J. G., ... & Beckham, G. T. (2015). Adipic acid production from lignin. *Energy & Environmental Science*, *8*(2), 617-628.

Veiga, T., Solis-Escalante, D., Romagnoli, G., ten Pierick, A., Hanemaaijer, M., Deshmuhk, A., … Daran, J.-M. (2012). Resolving Phenylalanine Metabolism Sheds Light on Natural Synthesis of Penicillin G in Penicillium chrysogenum. *Eukaryotic Cell*, *11*(2), 238–249. https://doi.org/10.1128/EC.05285-11

Ververidis, F., Trantas, E., Douglas, C., Vollmer, G., Kretzschmar, G., & Panopoulos, N. (2007). Biotechnology of flavonoids and other phenylpropanoid-derived natural products. Part I: Chemical diversity, impacts on plant biology and human health. *Biotechnology Journal: Healthcare Nutrition Technology*, *2*(10), 1214-1234.

Vesth, T. C., Nybo, J. L., Theobald, S., Frisvad, J. C., Larsen, T. O., Nielsen, K. F., ... & Gladden, J. M. (2018). Investigation of inter-and intraspecies variation through genome sequencing of Aspergillus section Nigri. *Nature Genetics*, 1.

Vrchotová, B., Macková, M., Macek, T., & Demnerová, K. (2013). Bioremediation of chlorobenzoic acids. In *Applied Bioremediation: Active and Passive Approaches*. BoD–Books on Demand.

Waldner, R., Leisola, M. S., & Fiechter, A. (1988). Comparison of ligninolytic activities of selected white-rot fungi. *Applied microbiology and biotechnology*, *29*(4), 400-407.

Wang, H., Pu, Y., Ragauskas, A., & Yang, B. (2019). From lignin to valuable products–strategies, challenges, and prospects. *Bioresource Technology*, *271*, 449–461. https://doi.org/10.1016/j.biortech.2018.09.072

Wenger, J., & Stern, T. (2019). Reflection on the research on and implementation of biorefinery systems–a systematic literature review with a focus on feedstock. *Biofuels, Bioproducts and Biorefining*. Advance online publication. https://doi.org/10.1002/bbb.2021

Whetten, R., & Sederoff, R. (1995). Lignin biosynthesis. *The plant cell*, *7* (7), 1001-1013.

Whittington, H. A., Grant, S., Roberts, C. F., Lamb, H., & Hawkins, A. R. (1987). Identification and isolation of a putative permease gene in the quinic acid utilization (QUT) gene cluster of Aspergillus nidulans. *Current genetics*, *12*(2), 135-139.

Wright, J. D. (1993). Fungal degradation of benzoic acid and related compounds. *World Journal of Microbiology and Biotechnology*, *9*(1), 9–16.

Xu, C., D. Arancon, R. A., Labidi, J., & Luque, R. (2014). Lignin depolymerisation strategies: towards valuable chemicals and fuels. *Chemical Society Reviews*, *43*(22), 7485–7500. https://doi.org/10.1039/C4CS00235K

Yazaki, K., Sasaki, K., & Tsurumaru, Y. (2009). Prenylation of aromatic compounds, a key diversification of plant secondary metabolites. *Phytochemistry*, *70*(15–16), 1739–1745. https://doi.org/10.1016/j.phytochem.2009.08.023

York, J. D., Ponder, J. W., & Majerus, P. W. (1995). Definition of a metal-dependent/Li (+)-inhibited phosphomonoesterase protein family based upon a conserved three-dimensional core structure. *Proceedings of the National Academy of Sciences*, *92*(11), 5149-5153.