Methyltransferase Specificity Toward Secondary Metabolite Production in *Aspergillus niger*

Susannah Selber-Hnatiw

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

July 2022

Concordia University

School of Graduate Studies

This is to certify that the thesis prepared

By: Susannah Selber-Hnatiw

Entitled: Methyltransferase Specificity Toward Secondary Metabolite Production 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. Jennifer Geddes-McAlister

_____ External Examiner

Dr. Elena Kuzmin

_____ Examiner

Dr. William Zerges

_____ Examiner

Dr. Isabelle Benoit-Gelber

_____ Thesis Supervisor

Approved by

_____

Dr. Robert Bertrand Weladji, Graduate Program Director

August 4th 2022

_____

Dr. Pascale Sicotte, Dean of Faculty

Abstract for Masters

Methyltransferase Specificity Toward Secondary Metabolite Production in *Aspergillus niger*

Susannah Selber-Hnatiw

Secondary metabolites are bioactive compounds that demonstrate a broad chemical diversity and represent a source of pharmacologically and industrially relevant compounds. Bioactivities of secondary metabolites can be improved through chemical modification, or *in vivo* using an organism as a cell factory. As such, secondary metabolite modifications were examined *in vivo* using the *Aspergillus niger* secondary metabolite TAN-1612/BMS-192548 (TAN/BMS). The genes involved in the biosynthesis of TAN/BMS are organized in a biosynthetic gene cluster, and consist of a polyketide synthase backbone, tailoring enzymes including an *O*-methyltransferase, a transporter, and a fungal-specific transcription factor. In general, methyl group addition represents a strategy used to modify and optimize drugs, and methyltransferases represent an opportunity to explore methylation *in vivo*. Selective methylation was examined in a TAN/BMS overexpression and knockout strain through the homologous insertion of genes. To this end, an intraspecies methyltransferase gene library was designed and overexpressed in *A. niger* at the glucoamylase A (*gla*A) locus which is often exploited for enzyme overproduction. Screening methods were used to examine the expression profiles of the gene-edited strains. The overexpression of the native TAN/BMS methyltransferase gene in *gla*A was shown to partially restore TAN/BMS methylation. The methyltransferase mutants did not recapitulate nor modify the methylation pattern of TAN/BMS. Results indicated that the localization of genes may play an important role in selective methylation.

Acknowledgements

I would like to give special thanks to those who supported me throughout my science journey.

Thank you to Dr. Isabelle Benoit-Gelber for taking me on as a student and elevating me as a scientist. Thank you for taking an interest in my education and for teaching me beyond the scope of the lab.

Thank you to Dr. Adrian Tsang for your advice, suggestions, and ideas. Working in your presence has been rewarding and motivating.

Thank you to Dr. Jennifer Geddes-McAlister who hosted me as a visiting scientist. Thank you for your virtue to collaborate, and for sharing your work with those who are interested to learn.

Thank you to Dr. Marcos di Falco for your extensive knowledge and capacity to share it with me.

Thank you to Marie-Claude Moisan for your company and interest in my work. Thank you for always listening, helping, and teaching me.

Thank you to Dr. Thi Thanh My Pham who was an excellent teacher. Thank you for your time and patience helping teach me essential concepts in science.

Thank you to Thi Truc Minh Nguyen and Dr. Victoria Aguilar for an informative and thorough introduction to bioinformatics and concepts in computational biology.

Thank you to Vanessa Blandford, Susan Sillaots, Maria Isabel Casanova, Veronica Trozzo, Ayat Farahat, Amrita Roka Magar, Mathieu Lavallee, Tri Minh Vo, Gregorio Budkewitsch, Nicholas Geoffrion, and Andrei Wasylyk who graciously helped me daily, passing along science wisdom and technical skills.

Thank you to Dr. Chiara Gamberi for helping launch my path and passion in science. I would not be here today without you.

Thank you for grant support from the NSERC-CREATE EvoFunPath fellowship program 2021-2022, and the Concordia University merit incentive 2020-2021.

Dedications

I dedicate this thesis to the people I love, my partner of seven years, Samuel Babity, my friends, Molly Carroll, Katherine Levasseur, Kim Raymond, Zarin Arshy, and Alexandra Baird, as well as my family, Jo-Anne Selber, Steven Hnatiw, Caroline Selber-Hnatiw, and Justin Selber-Hnatiw.

Contributions of Authors

Statement of Contribution	Contributor
• Preparation of this thesis. • Phylogenetic tree construction. • Extensive literature search on methyltransferase gene candidates. • DNA sequencing analysis on NRRL3_09545^{OE}09546^{KO}. • NMR extraction of NRRL3_09545^{OE} and NRRL3_09545^{OE}09546^{KO} strains. • Transformation of 13 methyltransferase genes into NRRL3 at the glucoamylase A gene locus. • Spectral analysis and comparison of NRRL 2270, NRRL3_09545^{OE}, NRRL3_09545^{OE}09546^{KO}, and mutant strains. • Validation of transformants through PCR amplification, mass spectrometry analysis of extracellular metabolites (sample preparation, instrument operation, and data analysis), and mass spectrometry based proteomics of intracellular proteins (sample preparation and data analysis). • Design of NRRL3_09546 gene locus methyltransferase insertions.	Susannah Selber-Hnatiw
• Plasmid construction of methyltransferase library genes. • NRRL3_09545^{OE} strain construction. • NRRL3_09545^{OE}09546^{KO} strain construction. • Design of NRRL3_09546 CRISPR-Cas9 guide.	Marie Beigas
• Mass spectrometer operation. • Metabolomic data analysis. • Proteomic data analysis.	Marcos di Falco
• Creator of mass spectrometry based proteomics protocol and workflow. • Mass spectrometry based proteomics data analysis.	Jennifer Geddes-McAlister
• Mass spectrometry based proteomics sample preparation.	Samantha Pladwig
• NMR data analysis.	Thi Thanh My Pham
• Phylogenetic tree and pairwise matrix assembly.	Thi Truc Minh Nguyen

All authors reviewed the final manuscript and approved the contents.

Table of Contents

List of Figures

List of Tables

List of Equations

List of Appendices:

List of Abbreviations

Item	Abbreviation
5-fluoro-orotic acid	5-FOA
ATP-dependent DNA helicase Ku70, involved in non-homologous end joining DNA double strand break repair	*kus*A
Base pair	bp
Chromosome	Chr
Clustered Regularly Interspaced Short Palindromic Repeats	CRISPR
Coenzyme	CoA,
CRISPR associated protein 9	Cas9
Da	Dalton
Deoxyribonucleic acid	DNA
Dilution factor	*df*
Dinucleotide triphosphate	dNTP
Genomic DNA	gDNA
Glucoamylase A	*gla*A
Glucoamylase A promoter	P*gla*A
Gravity	G
Kilodalton	kDA
Knockout	KO
Ligation Independent Cloning	LIC
Litre	L
Microgram	µg
Microlitre	µL
Micromolar	µM
Millilitre	mL
Millimetre	mm
Millimolar	mM
Minimal media	MM
Minutes	min
Molar	M
Non-Homologous End Joining	NHEJ
Nuclear Magnetic Resonance	NMR
CRISPR RNA	crRNA
Number	*n*
Overexpression	OE
Orotidine-5'-phosphate decarboxylase gene	*pyr*G
Phenol:chloroform:isoamyl alcohol	PCI
Polymerase Chain Reaction	PCR
Revolutions Per Minute	RPM
Ribonucleic acid	RNA
Short guide RNA	sgRNA
TAN-1612 BMS-192548	TAN/BMS
Transactivating crRNA	tracrRNA

Transfer RNA	tRNA
Ultraviolet	UV
Volume	V
Weight per volume	w/v
α-glucosidase promoter	P*agdA*

Purpose

The purpose of this thesis is to increase the chemical diversity of bioactive secondary metabolites and investigate methods to augment molecule optimization *in vivo* using fungal genomics and genetics. Secondary metabolites are natural products mainly produced by plant, fungi, and bacteria. They are molecules with low molecular weights that demonstrate a broad chemical diversity. Secondary metabolites are of great interest in health and agri-food sectors due to their bioactive properties. These include pharmaceutically relevant properties such as antimicrobial, antitumor, and cholesterol-lowering medications, and food colouring, flavouring, and texturing agents. In addition, some secondary metabolites represent important plant growth promoters. The biosynthesis of the secondary metabolite, neuropeptide Y receptor antagonist TAN-1612/BMS-192548 (TAN/BMS) was selected as a proof-of-concept to examine molecule optimization through the replacement and characterization of methyltransferases. The methyl group is one of the well-known high-impact substituents in drug design, where methyltransferases are enzymes that transfer methyl groups to their substrates resulting in selective methylation. In doing so, the study of methyltransferase specificity and secondary metabolite production was examined using *Aspergillus niger* as a cell factory. The objectives are as follows: 1) construction of a methyltransferase library, 2) production of modified secondary metabolites, and 3) characterization of non-methylated and methylated TAN/BMS molecules.

1. Introduction

1.1 Production and function of secondary metabolites

Secondary metabolites, also known as natural products, are organic compounds produced by organisms that are not required for growth, development, or reproduction, but can provide an ecological advantage that stimulates competition and fitness (Firn and Jones 2000). Secondary metabolites are low-molecular-weight compounds with a vast degree of structural and functional diversity. Notably, many identified secondary metabolites have bioactive effects and represent important pharmaceuticals and other commercially relevant organic compounds derived from plants, bacteria, and fungi.

1.1.1 Plant and bacterial secondary metabolites

Plant secondary metabolites comprise three major chemical classes, including polyphenols, terpenes, and alkaloids, with polyphenols primarily involved in processes such as UV damage protection, defence against herbivores, pathogens, and allelopathy, as well as facilitating the attraction of pollinators (Bourgaud *et al*. 2001; Stevenson *et al*. 2017). Pharmaceutically relevant plant secondary metabolites include the terpenoid chemotherapeutic agent paclitaxel, originally obtained from Pacific yew (Zhong 2002), as well as alkaloids such as morphine and codeine derived from opium poppies (Ziegler *et al*. 2009). In addition, polyphenols (*e.g.,* flavonoids, phenolic acids, and tannins) represent clinically significant antioxidants which have been investigated as treatments against cardiovascular disease and cancer (Morton *et al*. 2000; Dai and Mumper 2010; Abotaleb *et al*. 2020). Bacterial secondary metabolites demonstrate an even greater chemical diversity, including polyketides, ribosomal peptides and non-ribosomal peptides, and can include structural motifs such as isoprenes, oligosaccharides, and β-lactam rings (Bibb 2005; Ruiz *et al*. 2010). Bacterial secondary metabolites include an abundance of antibacterial, antifungal, and anticancer agents, with *Streptomyces* spp. producing a plurality of commonplace antibiotics (*e.g.,* tetracycline, erythromycin, and kanamycin) (Challis and Hopwood 2003; Ruiz *et al*. 2010).

1.1.2 Fungal secondary metabolites

Fungi are prolific producers of secondary metabolites, facilitating their diverse ecological roles and imparting bioactive effects. Fungi can grow in ecologically diverse regions as a result of melanin biosynthesis, which is involved in UV protection and has been shown to be involved in protection against pathogens (Jacobson 2000; Eisenman and Casadevall 2012; Keller 2019). In a similar manner, the antifungal properties of secondary metabolites — for example gliotoxin, produced by *Aspergillus fumigatus* — stimulate fungus-fungus competition and help shape ecological niches (Coleman *et al*. 2011). In terms of their pharmaceutical and commercial relevance to human health, fungal secondary metabolites have broad applications owing to their robust chemical diversity and broad antibacterial properties.

The bioactive properties of fungal secondary metabolites were first highlighted in the late 1920s following the discovery of the antibiotic penicillin, originally derived from *Penicillium notatum* (Fleming 1932; Fleming 2001). Since then, screening programs have identified thousands of secondary metabolites from a variety of chemical classes, including polyketides, non-ribosomal peptides, ribosomal peptides, and terpenes. Examples of other clinically relevant fungal secondary

metabolites include polyketide-derived lovastatin, which functions as a cholesterol lowering agent (Kennedy *et al*. 1999), and non-ribosomal peptide synthetase-derived cyclosporin, which can be used as an immunosuppressant during organ transplantation (Hoppert *et al*. 2001). A number of organisms produce chemically related secondary metabolites, including structurally similar compounds which are produced by evolutionarily related genera, or a range of chemical analogues produced by the same species. For instance, the β-lactam antibiotic penicillin is produced by several species of the so-called *Penicillium notatum-chrysogenum* group (Waksman and Reilly 1944). Considering their benefit to human health, a plethora of penicillin-producing species were identified and scrutinized at the onset of World War II for their ability to grow in high yields, shaping the concept of strain improvement (Gailey *et al*. 1946).

Genome sequencing has revealed metabolic pathways and expanded the identification of secondary metabolites (Baker 2006; Andersen *et al*. 2011). *A. niger* genome sequencing and examination of the metabolome has revealed extensive secondary metabolite production, as exemplified by the diverse chemical classes of bioactive molecules (Nielsen *et al*. 2009; Andersen *et al*. 2011; McDonnell *et al*. 2018). Furthermore, many secondary metabolites have bioactive effects or yet-to-be-defined clinical and/or commercial relevance (**Table 1,** adapted from Nielsen *et al.,* 2009).

Table 1: Bioactive secondary metabolite production in *Aspergillus niger*

Secondary Metabolite	Bioactive effect or industrial relevance	Reference
Asperenone	Lipoxygenase and human platelet aggregation inhibitor	(Jefferson 1967)
Azanigerone A	Anti-tumour, antifungal, and antiviral activity	(Zabala *et al*. 2012)
Nafuredin	Anthelminthic	(Ōmura *et al*. 2001)
Ochratoxin α	Mycotoxin	(Nielsen *et al*. 2009)
Fumonisin (B_2, B_4)	Mycotoxin	(Frisvad *et al.,* 2007; Noonim *et al.,* 2009)
Siderophore	Iron chelator	(Osman *et al*. 2019)
Hexylitaconic acid	Inhibitor of p53-induced tumorigenesis, plant growth regulator	(Isogai *et al*. 1984; Almassi *et al*. 1994; Tsukamoto *et al*. 2006)
Pyranonigrin (A, B, C, D, S)	Antioxidant	(Riko *et al*. 2014; Tang *et al*. 2018)
Tensidol A, B	Anti-Candida miconazole activity	(Fukuda *et al*. 2006)
2-Methylene-3-hexylbutanedioic acid	Improved crop germination	(Almassi *et al*. 1994)
Tensyuic acid A to F	Antibiotic	(Hasegawa *et al*. 2007)

Secondary Metabolite	Bioactive effect or industrial relevance	Reference
TAN-1612/BMS-192548	Neuropeptide Y inhibitor, and neurokinin-1 receptor antagonist	(Kodukula *et al*. 1995; Shu *et al*. 1995)
Nigerazine A and B	Plant growth inhibitor	(Iwamoto *et al*. 1983)
Nigerloxin	Lipoxygenase inhibitor	(Rao *et al*. 2005)
TMC-256A1, TMC-256C1	Naphthopyrone antibiotics	(Sakurai *et al*. 2002)
Pyrophen	Antimicrobial activity	(Padhi *et al*. 2020)
22-Deacetylyanuthone A	Anti-Candida activity	(Holm *et al*. 2014)
Yanuthone A-E	Anti-Candida activity	(Bugni *et al*. 2000; Petersen *et al*. 2015)
Aspernigrin B	Cytotoxicity to human cancer cells	(Hiort *et al*. 2004)
Bicoumanigrin	Neuroprotective effects against glutamic acid-induced cell death	(Hiort *et al*. 2004)
Flaviolin	Novel 3-chymotrypsin-like protease targeting SARS-CoV-2	(Wang *et al*. 2021; Rao *et al*. 2022)
Orlandin	Plant growth inhibitor	(Cutler *et al*. 1979; Wang *et al*. 2021)
Aspernigerin	Tumour cytotoxicity	(Shen *et al*. 2006)
Malformin (A, A_1, A_2, B, C,)	Mycotoxin	(Yoshizawa *et al*. 1975; Anderegg *et al*. 1976; Sugawara *et al*. 1990; Kim *et al*. 1993)

1.2 Biosynthetic gene clusters

Biosynthetic gene clusters coordinate the production of secondary metabolites, which in turn facilitates competition between other organisms in the environment and provides protection from abiotic factors. However, under laboratory conditions, secondary metabolite production is silent and requires activation through external stimulus such as culture conditions or manipulation at the genetic level.

1.2.1 Biosynthetic gene cluster architecture

In general, a biosynthetic gene cluster is composed of locally clustered genes that encode one or more backbone enzymes, tailoring enzymes, transporters, and in some cases regulatory genes such as transcription factors. Secondary metabolite biosynthesis is initiated through the polymerization of primary metabolites by a backbone enzyme that assemble the core structures of secondary metabolites (Keller 2019). A variety of these enzymes exist in filamentous fungi, including polyketide synthases (Ferracin *et al*. 2012), non-ribosomal peptide synthases (Cramer *et al*. 2006), prenyltransferases (Inglis *et al*. 2013), and terpene cyclases (Inglis *et al*. 2013).

The synthesis of polyketides and terpenes is derived from acetyl-CoA, propionyl-CoA, and malonyl-CoA, while the synthesis of non-ribosomal peptides and other derivatives is derived from amino acid substrates (*e.g.,* L-tryptophan, and L-alanine) (Keller 2019). Tailoring enzymes then decorate the secondary metabolite backbone by catalyzing reactions that join molecules (*i.e.,* synthetase reactions), add functional groups (*i.e.,* synthase reactions), or convert linear molecules into rings (*i.e.,* cyclase reactions). Some biosynthetic gene clusters also encode transporters that remove toxic by-products (*e.g.,* gliotoxin efflux pumps (Dolan *et al*. 2015)), and shuttle secondary metabolites into the cytosol for further backbone modifications or direct use by the organism (*e.g.,* citrate transport protein (Kirimura *et al*. 2016)).

Remarkably, some secondary metabolites are encoded by genes which are not locally clustered in the same chromosomal region. The environmental iron-chelating compound siderophore is produced by several *Aspergillus* species and is distributed over multiple biosynthetic gene clusters, with two clusters involved in the production of siderophore in *A. niger*, and three in *A. fumigatus* as well as *A. nidulans* (Franken *et al*. 2014). This is also exemplified in the biosynthesis of melanin and naphtho-γ-pyrones, in which *alb*A initiates the production of both secondary metabolites in distinct pathways that are not encoded by co-localized genes (Chiang *et al*. 2011).

1.2.2 Regulation of biosynthetic gene clusters

Certain biosynthetic gene clusters contain regulatory genes that function to induce the transcription of genes encoding secondary metabolite pathways, or to relax repression mechanisms. Fungal-specific transcription factors comprise three superfamilies including $Zn(II)_2Cys_6$ motif transcription factors (Chang *et al*. 1995), the DNA binding domain of Mlu1-box-binding protein MBP1, and the Zinc domain conserved in yeast copper-regulated transcription factor. The latter two superfamilies are specifically involved in cell cycle regulation (Jungmann *et al*. 1993; Ayté *et al*. 1997; Machado *et al*. 1997; Keller *et al*. 2001; Keller *et al*. 2005), while $Zn(II)_2Cys_6$ is essential for survival, metabolism, development, and reproduction (MacPherson *et al*. 2006; Shelest 2008). In general, transcription factors functionally activate transcription by interacting with pathway-specific regulation mechanisms, or through the concurrent and coordinated regulation of several clusters. For example, in *A. nidulans*, the master regulator McrA has been shown to coordinate the production of numerous secondary metabolites through negative regulation mechanisms in which the deletion of velvet complex proteins (implicated in secondary metabolite production), resulted in the upregulation of different biosynthetic gene clusters (Oakley *et al*. 2017). In contrast, the methyltransferase-domain protein LeaA functions as a transcriptional activator, and together with the velvet complex proteins, demonstrated a key role in secondary metabolite production in *P. chrysogenum* (Bayram and Braus 2012). Transcription factors can therefore be used as a tool to activate the regulation of clustered genes, resulting in the biosynthesis of secondary metabolites. A variety of inducible and constitutive promoters have been utilized in *A. niger* biosynthetic gene cluster pathway engineering — including the glucoamylase promoter (P*gla*A) (Fowler *et al*. 1990), and α-glucosidase promoter (P*agdA*) (Minetoki *et al*. 1996) that serve to replace or supplement native promoter genes and activate transcription in individual biosynthetic gene clusters.

Other regulatory mechanisms rely on DNA modifications or environmental stimuli. For instance, epigenetic transcriptional regulators control chromatin remodelling in which nucleosome assembly and disassembly is modified by the repression of enzymes (such as histone deacetylases) to regulate the activity of RNA polymerase II transcription machinery (Albright *et al*. 2015;

Macheleidt *et al*. 2016). As such, the deletion of histone deacetylase enzymes modulates epigenetic regulation networks and relaxes chromatin remodeling by removing repression mechanisms (Guzman-Chavez *et al*. 2018; Nie *et al*. 2018). Environmental stimuli can also induce the transcription of biosynthetic gene clusters through signal transduction pathways (*e.g.,* cyclic AMP-dependent protein kinase A) (Macheleidt *et al*. 2015), and by manipulating culture conditions to induce stress (*e.g.,* UV radiation, or altered pH, temperature, osmotic pressure, and oxygen content). Indeed, culture conditions can be used to replicate some factors that induce secondary metabolite production in nature (Buchanan and Ayres 1975; Espeso *et al*. 1993; Duran *et al*. 2010; Amare and Keller 2014; Fountain *et al*. 2016; Blachowicz *et al*. 2020), including changes in media composition (J.C. Frisvad *et al*. 2007) (*i.e.,* carbon sources) and co-culture with other organisms (Rateb *et al*. 2013).

1.3 Small-molecule modifications to secondary metabolites

The bioactive properties of secondary metabolites can be improved in a manner similar to lead drug optimization, which includes randomized strategies (*i.e.,* high-throughput and fragment screening), as well as rational optimization. In either case, these avenues of drug optimization presuppose that a parent drug is bioactive and can be redesigned using strategies to further improve its function. Rational approaches aim to improve the compatibility of a molecule with known biological targets using small-molecule modifications (Mandal *et al*. 2009). In randomized approaches, in contrast, hundreds to millions of drugs and chemical entities are tested *in silico* or in miniaturized assays for their ability to act on biological targets (Bajorath 2002; Mayr and Bojanic 2009). However, target identification using this method can generate a plethora of undesirable molecules (*e.g.,* toxic to cells, off-target effects) or false positives (Thorne *et al*. 2010), and ultimately requires some level of rational design or consideration of targetable biological mechanisms. As such, drug discovery approaches that focus on rational optimization may be more direct and robust (**Figure 1**).

A number of small-molecule modifications, including methyl (Feng *et al*. 2020; Aynetdinova *et al*. 2021), hydroxyl (Cramer *et al*. 2019), and acetyl addition (Adessi and Soto 2002), are used to conduct molecule optimization. In general, these modifications are applied to lead compounds and are used to alter steric hinderance at the active site, increase potency (*e.g.,* the magic methyl effect) and bioactivity, as well as modifying the lipophilicity and/or solubility of a molecule (Sun and Fu 2018). Depending on technical factors (*e.g.,* feasibility, cost, *etc*.) different small-molecule modifications, and combinations of modifications, are applied to a lead compound in order to generate an optimal structure. Due to their ability to augment small-molecule modifications, and the variety of successes related to their use, methyl group addition will be emphasized. The addition of methyl groups is a common strategy used in medicinal chemistry to modify the pharmacological properties of drugs. Examples of methyl group addition include the increased bioactivity of the anti-angiogenic drug Cilengitide through the modification of the receptor-binding domain (Mas-Moruno *et al*. 2010), and improvements in the potency of the anti-cancer drug Crizotinib (Cui *et al*. 2011). Furthermore, recent advances in organic synthesis have also enabled the regioselective addition of methyl groups (Sun and Fu 2018; Feng *et al*. 2020; Aynetdinova *et al*. 2021) to augment the rational redesign of bioactive molecules.

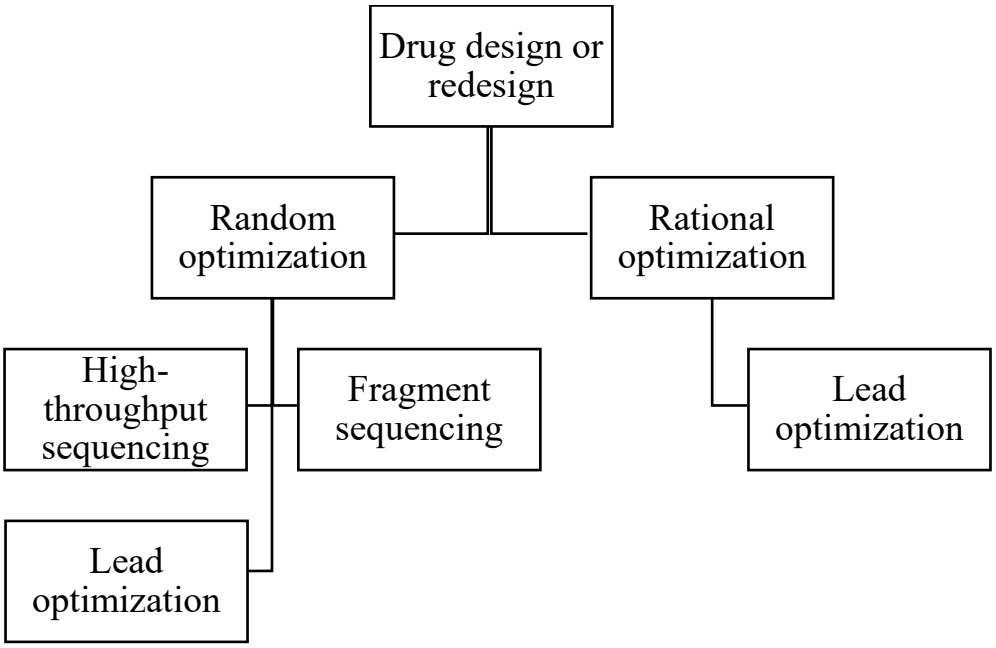


Figure 1: Drug design roadmap. Lead optimization is the objective for different approaches to drug design. Random optimization involves searches for drug hits using pharmacological screening and relies on a trial-and-error method of problem solving. Rational optimization posits strategies to fine-tune bioactive compounds through target-based pharmacology.

Secondary metabolites are functionally diverse molecules which can be improved through small-molecule modifications. For instance, the chemical addition of methyl groups has been explored in lovastatin to form simvastatin. As mentioned, lovastatin is a polyketide-derived cholesterol-lowering agent that is produced by *Penicillium* and *Aspergillus* species using various carbon sources and induction mechanisms (Ruiz *et al*. 2010). Lovastatin can be chemically modified through the addition of a methyl group to form the semisynthetic drug simvastatin, which displays a more robust safety and efficacy profile (Tobert 2003). Other chemically related secondary metabolites — such as morphine and codeine — are chemical derivatives that differ in terms of methylation patterning. While morphine and codeine both function as analgesics, the selective addition of a methyl group to C3 on the morphine chemical backbone results in codeine, which binds opioid receptors less strongly than the former, contributing to its milder pain-relieving properties (Vardanyan and Hruby 2006). This chemical modification attenuates the potency of morphine, in which lower potency results in a drug that is more pharmaceutically desirable under certain circumstances (Drewes *et al*. 2013). As such, the identification and modification of secondary metabolites though small molecule changes is of interest. In addition to the methylation of drugs through chemical interventions, biocatalytic methylation is a natural mechanism observed in the biosynthesis of secondary metabolites. As such, it is possible to leverage rational optimization *in vivo* using a model organism such as *A. niger* and its associated molecular toolbox to manipulate modular tailoring enzymes and modify secondary metabolites.

1.3.1 Methyltransferases

Methyltransferases are tailoring enzymes that catalyse the transfer of methyl groups from a donor molecule to a substrate, modifying -*O*, -*N*, -*S*, and -*C* atoms primarily using S-Adenosyl

methionine as a methyl donor. A variety of biosynthetic gene clusters contain a methyltransferase gene to enable secondary metabolite modifications, however few studies have been conducted to explore biocatalytic methylation as strategy to modify methyl patterning in secondary metabolites.

Combinatorial biosynthesis is an approach used to modify the genes in a biosynthetic gene cluster in order to modify the product of secondary metabolite biosynthesis, and has been explored in a variety of biosynthetic gene clusters encoding non-ribosomal peptide synthetase and polyketide synthase backbones (Weissman and Leadlay 2005; Zhang and Tang 2008; Ang *et al*. 2015 Feb). One study examined non-ribosomal peptide synthetase-derived daptomycin modifications in *Streptomyces roseosporus*, in which the deletion and rearrangement of genes — including the deletion of a methyltransferase — resulted in the successful formation of novel daptomycin analogues (Nguyen *et al*. 2006). Another study examined polyketide synthases and *O*-methyltransferase activity in the spinosyn biosynthetic gene cluster of *Saccharopolyspora spinosai* (Gaisser *et al*. 2008). Ultimately, it was found that novel macrolide antibiotics could be produced in *S. erythraea* following replacement of native methyltransferases with methyltransferases from *S. spinosai* exhibiting similar substrate specificity toward rhamnose sugar analogues (Gaisser *et al*. 2008). Methyltransferases can also target and modify DNA during transcription, which can be used to modulate epigenetic gene regulation. For example, fungal DNA methyltransferases coordinate post-translational modifications that modify the methylation pattern of histones, and in turn affect gene transcription and secondary metabolite production. As such, the modification of fungal-derived DNA methyltransferases using inhibitors has been used as a strategy to modulate the production of secondary metabolites and form novel bioactive compounds (Williams *et al*. 2008; Nie *et al*. 2018). While the specificity of the methyltransferases to their substrates is not well-described in the literature, the use of methyltransferase gene replacement to generate alternative methyl patterning could reveal new methods for the rational optimization of secondary metabolites. Therefore, using synthetic biological approaches including combinatorial biosynthesis, late-stage functionalization through methylation can be used to modify secondary metabolites and generate novel compounds with altered bioactivity.

1.4 *Aspergillus niger*

To prepare *A. niger* as a cell factory for the production of organic acids, enzymes, and secondary metabolites, as well as establishing robust gene-editing protocols, the cultivation of different parental strains and the development of molecular tools must be thoughtfully considered. In this section, a brief background of historical and modern applications of the *A. niger* toolbox will be discussed, in addition to an overview of established industrial strains.

1.4.1 *A. niger* in nature

A. niger is a common filamentous ascomycete globally distributed in soils. In the environment, *A. niger* excretes various polysaccharide-degrading enzymes (*e.g.,* galactosidases, esterases, and cellulases (Hurst *et al*. 1977; Manzanares *et al*. 1998; de Vries and Visser 1999) to allow the extracellular digestion of plant cell walls, thereby releasing nutrients for its growth (Garcia-Conesa *et al*. 1999; de Vries 2003; Rehman *et al*. 2014; Šimonovičová *et al*. 2021). *A. niger* can effectively grow under the influence of a variety of abiotic factors — including a range of pH values, temperatures, as well as UV radiation, aeration, and moisture levels — enabling its extreme

adaptability in nature (Šimonovičová *et al*. 2021), and contributing to its status as a common food contaminant and agricultural pest (Perrone *et al*. 2007).

The structure of an *A. niger* colony consists of elongated conidiophores supported by a network of vegetative hyphae. Each conidiophore supports a vesicle that differentiates into a basal structure called the metule, which is further differentiated into another supporting structure known as the sterigmata, and ultimately results in the formation of hundreds to thousands of conidia (Anderson and Smith 1971). Conidia are the asexual spores of *A. niger*, producing its characteristic black pigment. Under favourable growth and nutrient conditions, *A. niger* hyphae expand into a network of mycelia that can support millions of spore-forming conidiophores.

1.4.2 *A. niger* as a cell factory

Due to the range of conditions in which filamentous fungi can grow in nature, *A. niger* was probed experimentally for its potential to grow under different pH, salt, and nutrient conditions. In the early 20th century, the production of oxalic acid (Raistrick and Clark 1919) and citric acid (Currie 1917), as well as malic and fumaric acids (Challenger and Klein 1929) was identified in *A. niger*, and readily used commercially and in global trade (Wells and Herrick 1938).

A. niger represents a pillar of biotechnology, producing organic acids and numerous enzymes (Cairns et al. 2018),to the pharmaceutical production of secondary metabolites, including antibiotics (*e.g.,* penicillin (Foster and Karow 1945)), immunosuppressants (*e.g.,* cyclosporine (Hornbogen *et al*. 1992; Derkx and Madrid 2001)), cholesterol-lowering drugs (*e.g.,* statins (Mouafi *et al*. 2016)), and anti-cancer drugs (*e.g.,* malformins (Kobbe *et al*. 1977; Frisvad *et al*. 2018)).

In general, *A. niger* exhibits an extensive history of safe use (Frisvad *et al.,* 2018; Schuster *et al.,* 2002) culminating in the production of multiple enzymes used in food-related applications that are regarded as Generally Recognized as Safe (GRAS) by the US Food and Drug Administration (FDA, 2018) Nevertheless, some strains of *A. niger* are responsible for the production of mycotoxins, including ochratoxin A, fumonisin B_2, and fumonisin B_4 (Frisvad *et al.,* 2007; Noonim *et al.,* 2009), and acts as an opportunistic pathogen in individuals with compromised immune health, potentially leading to the development of pulmonary aspergillosis (Segal 2009).

1.4.3 *A. niger* strains

More than 100 years of *A. niger* research has helped shape the organic acid, enzyme, and secondary metabolite pipeline. Advances in molecular biology and genetics — including whole-genome sequencing (Pel *et al*. 2007), bioinformatics and increases in computing power (Andersen *et al*. 2011), as well as the modern use of the clustered regularly interspaced short palindromic repeats (CRISPR) system (Song, Ouedraogo, *et al*. 2018) — have particularly accelerated the discovery of industrially and pharmaceutically relevant molecules. Different genome projects established in the 21st century have sequenced three *A. niger* genomes, including the 35.25 Mb NRRL3 Wild type (CBS 120.49, ATCC 9029) (Integrated Genomics, US), the historic 34.85 Mb Wild-type ATCC 1015 citric acid-producing strain (CBS 113.46, NRRL_328) (US Department of Energy Joint Genome Initiative, US), and the 33.98 Mb CBS 513.88 mutant strain derived from

NRRL_3122 (ATCC 22343, CBS 115989) (DSM, Netherlands) which was engineered to display elevated glucoamylase A production (Pel *et al*. 2007) (**Table 2**).

Table 2: Comparison of *A. niger* strains

Strain Identification (CBS, ATCC model names)	Lineage (Phenotype)	Commercial Relevance	Reference
NRRL_328 (CBS 113.46, ATCC 1015)	Wild type	Citric acid production	(Currie 1917; Semova *et al*. 2006; Andersen *et al*. 2011)
NRRL3 or N400 (CBS 120.49, ATCC 9029)	Wild type (long conidiophores)	Enzymes, organic acids, and secondary metabolite production	(Bos *et al*. 1988; Debets *et al*. 1993; Andersen *et al*. 2011)
NRRL_3122 (Wild-type progenitor of CBS 513.88, ATCC 22343)	Derived from ATCC 1015 (Requires arginine supplementation)	Glucoamylase A production	(van Lanen JM 1968; Pel *et al*. 2007; Andersen *et al*. 2011)
(CBS 513.88)	Derived from ATCC 22343	Enzyme production	(Andersen *et al*. 2011)

Whole-genome shotgun sequencing has revealed sequence similarities and functional genomic comparisons between the ATCC 1015 and CBS 513.88 *A. niger* genomes, demonstrating *A. niger* evolution over the course of development between these two industrial strains (Andersen *et al*. 2011). The CBS 513.88 genome (van Lanen JM 1968) is closely related to NRRL3 (Aguilar-Pontes *et al*. 2018), as it is derived from a glucoamylase A-producing mutant and differs based on indel variations (Andersen *et al*. 2011).

A comparison of different *A. niger* parental strains is presented in **Table 2**, which includes the glucoamylase A-producing strain, as well as other strains that have been identified for commercial relevance; additional gene-edited *A. niger* strains are found in **Table 3**. *A. niger* sequenced genomes have enabled the identification of essential genes (Andersen *et al*. 2011; McDonnell *et al*. 2018), and the optimization of the *A. niger* genetic toolbox; including robust transformation protocols, the preparation of selection markers and gene repair mechanisms, as well as the engineering of inducible promoters (Debets and Bos 1986; Song, Ouedraogo, *et al*. 2018; Meyer *et al*., 2006; van Hartingsveldt *et al*. 1987). Overall, this genetic toolbox represents a system for gene activation, site-specific gene-editing, and includes various auxotrophy controls that can be used to confer artificial selection.

Table 3: *A. niger* strains used in synthetic organic acid and enzyme production

	Strain Identification (ATCC model names)	Lineage (Phenotype)	Product	Reference
Classic mutagenesis	N401	Derived from NRRL 3 though UV mutation (short conidiophores. Poor growth on starch)	N/A	(Demirci *et al*. 2021)
	N402 (ATCC 64974)	Derived from NRRL 3 through UV mutation (short conidiophores)	N/A	(Bos *et al*. 1988; Debets *et al*. 1993)
	N593 (ATCC 64973)	Derived from NRRL 3 through UV mutation (*pyr*G deletion mutant)	N/A	(Goosen *et al*. 1987)
	NRRL 2270 (ATCC 11414)	Derived from NRRL 328 (ATCC 1015)	Hyperproduction of citric acid	(Perlman *et al*. 1946)
	(ATCC 13496)	Derived from N402 through UV mutation	Hyperproduction of glucoamylase A	(Fiedler *et al*. 2018; Vesth *et al*. 2018)
Modern mutagenesis	N/A	Derived from NRRL 328 through CRISPR-Cas9 gene-editing	Oxalic acid	(Xu *et al*. 2019)
	N/A	Derived from NRRL 328 through CRISPR-Cas9 gene-editing	Malic acid	(Xu *et al*. 2019)
	N/A	Derived from NRRL 328 through CRISPR-Cas9 gene-editing	Succinic acid	(Yang *et al*. 2020)
	Computational model	N/A	Enzymes involved in central carbon metabolism	(David *et al*. 2003)

To induce the production of enzymes and organic acids, *A. niger* wild-type strains (**Table 2**) have been evolved through traditional mutagenesis (Currie 1917; Bos *et al*. 1988; Pel *et al*. 2007; Andersen *et al*. 2011) and synthetic modifications *via* site-specific gene-editing (**Table 3**). The availability of manually curated and annotated strains has enabled the identification of relevant genes and the potential to perform site-directed mutagenesis. The NRRL3 genome (https://genome.fungalgenomics.ca/) was manually curated and assembled based on pre-existing structural annotations, as well as *de novo* transcripts, RNA sequencing data (*i.e.,* transcriptomic information), mass spectrometry-derived protein data, and intron positions (Aguilar-Pontes *et al*. 2018; McDonnell *et al*. 2018).

Derivative strains of NRRL3 are used to produce a variety of enzymes and organic acids, including the citric acid-producing strain NRRL 2270 (Perlman *et al*. 1946), the glucoamylase A-producing strain NRRL_3122 (van Lanen JM 1968; Pel *et al*. 2007; Andersen *et al*. 2011), and the

glucoamylase A hyper-producing strain ATCC 13496 (Fiedler *et al*. 2018; Vesth *et al*. 2018). A variety of industrially and medicinally relevant enzymes related to central carbon metabolism (*e.g.,* oxalic, succinic, and malic acids) are also produced using NRRL3 (Xu *et al*. 2019). In addition, *in silico* reconstructions of *A. niger* carbon metabolism have revealed theoretical maximal enzyme and organic acid yields (David *et al*. 2003), as well as predicted protein production (Tsang *et al*. 2009), demonstrating the robust *A. niger* cell factory, and enzyme and organic acid pipeline. The manual curation of *A. niger* NRRL3 has revealed 96 secondary metabolite backbone genes assembled in 84 biosynthetic gene clusters (Palys *et al*. 2017). Only 11 biosynthetic gene clusters have been fully characterized, as depicted in **Table 4** (adapted from https://genome.fungalgenomics.ca/).

Table 4: Characterized biosynthetic gene clusters in *Aspergillus* NRRL3

Biosynthetic gene cluster	**Gene architecture**			
	Regulatory genes	**Backbone enzyme**	**Tailoring enzymes**	**Transporter**
Aza*niger*one	fungal-specific transcription factor *AzaR*	polyketide synthase *AzaA*, polyketide synthase *AzaB*	esterase/lipase *AzaC*, acyltransferase *AzaD*, ketoreductase *AzaE*, AMP-dependent CoA ligase *AzaF*, FAD-dependent oxigenase *AzaG*, salicylate monooxygenase *AzaH*, cytochrome P450 *AzaI*, zinc-type alcohol dehydrogenase *AzaJ*, haem-degrading HbpS-like protein, FAD-dependent oxygenase *AzaL*	major facilitator superfamily efflux transporter *AzaK*
Fumonisin	fungal-specific transcription factor *Fum21*	polyketide synthase *Fum1*, non-ribosomal peptide synthetase-like protein *Fum14*	bifunctional P-450:NADPH-P450 reductase *Fum6*, short-chain dehydrogenase/ reductase family protein AMP-dependent synthetase/ligase *Fum10*, iron-containing alcohol dehydrogenase *Fum7*, phytanoyl-CoA dioxygenase *Fum9*, aminotransferase *Fum8*, NAD-dependent epimerase/dehydratase family protein *Fum13*, cytochrome P450 *Fum15*	ABC transporter *Fum19*, major facilitator superfamily protein

Biosynthetic gene cluster	Gene architecture			
	Regulatory genes	**Backbone enzyme**	**Tailoring enzymes**	**Transporter**
Yanuthone C	fungal-specific transcription factor YanR	prenyltransferase *YanG*	cytochrome P450 *YanC*, decarboxylase *YanB*, 6-methylsalicylic acid synthase *YanA*, cytochrome P450 *YanH*, O-mevalon transferase *YanI*, short-chain dehydrogenase/reductase *YanD*, RmlC-like cupin domain-containing protein *YanE*, oxidase *YanF*	
TAN-1612/ BMS-192548	fungal-specific transcription factor *AdaR*	polyketide synthase *AdaA*	O-methyltransferase *AdaD,* FAD-dependent monooxygenase *AdaC,* metallo-beta-lactamase-type thioesterase *AdaB*	major facilitator superfamily protein, efflux pump
Carlosic acid-agglomerin	fungal-specific transcription factor *CaaR*	hybrid polyketide synthase/non-ribosomal peptide synthetase *CaaA*	oxoglutarate/iron-dependent dioxygenase CaaD*,*trans enoylreductase *CaaB*, cytochrome P450 *CaaC*	major facilitator superfamily protein, nuclear transport factor 2 domain-containing protein
DHN Conidial pigmentation (Alb/Fwn) (dispersed)		polyketide synthase *AlbA/FwnA*	hydrolase *AygA/OlvA*, multicopper oxidase *BrnA*, multicopper oxidase *Mco1A*	
Kotanin	no in-cluster transcription factor	polyketide synthase *KtnS*	cell wall mannoprotein 1-like protein, flavin monooxygenase-like protein, cytochrome P450 *KtnC*, O-methyltransferase *KtnB*	major facilitator superfamily protein

Biosynthetic gene cluster	Gene architecture			
	Regulatory genes	**Backbone enzyme**	**Tailoring enzymes**	**Transporter**
Pyranonigrin	fungal-specific transcription factor	hybrid polyketide synthase/non-ribosomal peptide synthetase PynA	FAD-dependent oxidoreductase *PynB*, NAD(P)-binding protein *PynE*, thioesterase domain-containing protein, aspartic peptidase-like protein, FAD-binding domain-containing protein *PynR*, O-methyltransferase *PynC*, cytochrome P450 PynD	major facilitator superfamily transporter *PynF*
Ochratoxin	bAP-1 transcription factor, fungal-specific transcription factor	bnon-ribosomal peptide synthetase, polyketide synthase-like protein	bcytochrome P450, short-chain dehydrogenase /reductase family protein, alcohol dehydrogenase domain-containing protein	
Siderophore	no in-cluster transcription factor	hydroxyornithine transacylase *SidF*, non-ribosomal peptide synthetase *SidD*, non-ribosomal peptide synthetase *SidC*	helicase, C-terminal domain-containing protein, GNAT domain-containing protein, phospholipid methyltransferase-like protein, siderophore iron transporter, mevalonyl-CoA hydratase, alpha/beta hydrolase fold domain-containing protein, tetratricopeptide-like helical domain-containing protein, mevalonyl-CoA ligase *SidI*,	major facilitator superfamily transporter SitT

1.4.4 The *A. niger* genetic toolbox

A range of gene-editing strategies have been used to evolve hyper-producing *A. niger* strains (**Table 3**). For example, increasing the gene copy number of the oxaloacetate acetylhydrolase gene, *oah*A, was found to result in the induction and overexpression of oxalic acid (Xu *et al*. 2019). In addition, transcription can be activated through the insertion of inducible tetracycline promoters (Wanka *et al*. 2016) or Pol III promoters for spliceosomal U6 small nucleotide RNA transcription

using the CRISPR-editing system (Meyer *et al*. 2011; Zheng *et al*. 2018), or by modulating regulation through alterations of the 5' untranslated region (Mulder *et al*. 2004; Tamayo-Ramos *et al*. 2013). Morphological engineering of mycelia has also been shown to alter enzyme production (Walisko *et al*. 2015), and different products can be grown and differentiated (*e.g.,* succinate) using renewable carbon biomass as a substrate for the production of enzymes and organic acids (**Table 3**) (Yang *et al*. 2020).

Though a variety of methods have been used to perform genome editing in *A. niger*, the CRISPR-CRISPR associated protein 9 (Cas9) system has shown improved capacity to direct site-specific gene editing and currently represents a leading method (Ledford 2015). The CRISPR-Cas9 system functions to localize CRISPR single guide RNA (sgRNA) to a specific region in the host genome and produces double-stranded DNA breaks *via* the RNA-guided endonuclease known as Cas9 (Jinek *et al*. 2012). sgRNA consists of a user-defined 20 nucleotide CRISPR RNA (crRNA) region, as well as a 5' transactivating crRNA (tracrRNA) region that is recognized by Cas9 through nucleotide complementarity (Jinek *et al*. 2012;). Double-stranded DNA breaks occur directly upstream of a protospacer adjustment motif (*i.e.,* 5'–NGG), and as such, genome editing can occur at any N_{20}-NGG sequence (Jinek *et al*. 2012;). However, the selection of sgRNA is important in order to maintain selectivity (*i.e.,* on-target endonuclease activity) and specificity (*i.e.,* minimizing off-target effects) (Sander and Joung 2014). Overall, the CRISPR-Cas9 genome-editing system enables the discovery of genes and is strategically used to engineer strains for the production of commercially relevant molecules. The availability of fully annotated and manually curated fungal genomes enables highly specific gene-editing capabilities and allows *A. niger* to behave as a cell factory for the production of novel organic acids, enzymes, and secondary metabolites. In addition, the rational redesign of bioactive compounds can be conducted to increase the chemical diversity of potential secondary metabolites.

Other important components in the *A. niger* toolbox are selectable markers and repair mechanisms. A variety of selectable markers have been used in *A. niger*, including auxotrophy markers for nutritional requirements (van Hartingsveldt *et al*. 1987), antibiotic resistance (Punt *et al*. 1987), and organic compound utilization (Kelly and Hynes 1985). The deletion of the orotidine-5'-phosphate (OMP) decarboxylase gene, *pyr*G, is a common selection strategy used to positively select for uridine auxotrophs (van Hartingsveldt *et al*. 1987). The deletion of *pyr*G is achieved by growing *A. niger* in the presence of mutagenic 5-fluoro-orotic acid (5-FOA) without uracil, causing *A. niger* to produce the toxic metabolite fluorodeoxyuridine. However, co-culturing *A. niger* with 5-FOA in the presence of uracil allows cells to grow by gaining resistance to 5-FOA and maintaining prototrophy.

Promoters represent another class of selectable marker, as they can be used to activate the expression of genes under inducible conditions. In *A. niger*, several inducible protomers have been identified for the purposes of transcription activation (Fowler *et al*. 1990; Turnbull *et al*. 1990). For instance, the glucoamylase A gene (*gla*A) is an inducible promoter that is highly expressed in the presence of several carbon sources, including maltose, glucose, and xylose (Fowler *et al*. 1990). Several groups have used the *gla*A promoter region to activate the expression of genes for metabolite engineering (Song, Ouedraogo, *et al*. 2018), as well as using the *gla*A region for gene integration and enzyme production (Dong *et al*. 2019).

Lastly, repair mechanisms in *A. niger* greatly influence the transformation efficiency of genes. The dominant repair mechanism in *A. niger* is non-homologous end joining (NHEJ). Compared to homologous recombination, where DNA double-stranded breaks are repaired using a homologous repair template, in NHEJ, double-stranded DNA breaks are ligated through the random integration of nucleotides (Goldman *et al*. 2002; Kooistra *et al*. 2004; Chang *et al*. 2017). As such, NHEJ can limit recombination efficiency and the likelihood that genes of interest are transformed into the *A. niger* genome. NHEJ has been demonstrated in organisms with dominant homologous recombination repair mechanisms (*e.g., Saccharomyces cerevisiae*) following genetic manipulation in order to observe the consequence of NHEJ (Clikeman *et al*. 2001). Overall, DNA repair was observed to be mediated through a DNA-dependent protein kinase comprising a catalytic subunit, as well as a heterodimer with a 70 kDa subunit (*i.e.,* ku70) and an 80 kDa subunit (*i.e.,* ku80). The catalytic subunit was found to be directed to the site of double-stranded DNA breaks, where the ku heterodimer conducted DNA ligation, while the entire complex functioned to prevent exonuclease activity (Gottlieb and Jackson 1993; Critchlow and Jackson 1998; Clikeman *et al*. 2001). In *A. niger*, the ku70 subunit is encoded by the gene *kus*A, which predominantly regulates the NHEJ pathway (Meyer *et al*. 2006). As such, to augment homologous recombination in *A. niger*, a k*us*A deletion strain has been generated to inhibit the ku70 subunit from ligating random double-stranded DNA (Meyer *et al*. 2006). As demonstrated in other homologous recombination-pathway models (Kooistra *et al*. 2004; Ninomiya *et al*. 2004; Pöggeler and Kück 2006), the deletion of NHEJ pathway genes resulted in an increased frequency of homologous recombination (Meyer *et al*. 2006). In addition, the integration capacity of specific DNA fragments between 500 and 1500 bp was experimentally tested, where the recombination frequency of the k*us*A mutant ranged from 80 – 95%, compared to 7 – 29% for the wild type (Meyer *et al*. 2006). Deletion of the NHEJ pathway is particularly important when studying the function of unknown and essential genes, in which targeted gene replacement by homologous recombination can be used to precisely elucidate the function of genes.

1.5 Rationale: The biosynthesis and modification of the TAN-1612/BMS-192548 biosynthetic gene cluster in *A. niger* NRRL3

Using an approach such as combinatorial biosynthesis, the molecular optimization of secondary metabolites can be explored by replacement of native tailoring enzymes with other intraspecies NRRL3 tailoring enzymes capable of functionalizing different substrates. In doing so, I hypothesize that novel secondary metabolites can be generated and optimized using different methyltransferases as methyl donors.

Among the 84 biosynthetic gene clusters that have been annotated in NRRL3 (https://genome.fungalgenomics.ca/), 17 included a methyltransferase gene. However, fewer have been described in the literature, including the Kotanin, Pyranonigrin, and TAN-1612/BMS-192548 (TAN/BMS) biosynthetic gene clusters (Li *et al*., 2011; Nielsen *et al*., 2009). The TAN/BMS biosynthetic gene cluster was ultimately selected due to the following criteria: methyltransferase activity, the presence of a transcription factor, a characterized *m/z* profile that can be used to distinguish secondary metabolite production through mass spectrometry, the production of a yellow pigment useful for screening, and previous experiments demonstrated that the overexpression of the biosynthetic gene cluster led to over 1000-fold higher TAN/BMS production than in the wild type (unpublished, presented in this thesis).

The TAN/BMS biosynthetic gene cluster of *A. niger* is localized on chromosome 7 and is comprised of six genes including the fungal-specific transcription factor (AdaR) and the methyltransferase gene (AdaD) (**Figure 2**). During TAN/BMS synthesis, secondary metabolite production is initiated by the polymerization of various acyl groups, which is encoded by the backbone polyketide synthase (AdaA) and subsequently modified by tailoring enzymes, including an FAD-dependent mono-oxygenase (AdaC) that coordinates oxidation, a metallo-β-lactamase-type thioesterase that performs the cyclic reduction of the backbone (AdaB), and an *O*-methyltransferase (AdaD) that methylates the secondary metabolite at the C9 position using S-Adenosyl methionine as a methyl donor to complete the biosynthesis of TAN/BMS (**Table 4**) (Li *et al*. 2011). However, I theorize that methylation may also occur at C5 or C7 or possibly at C2 or C15 (Bauerle *et al*. 2015).

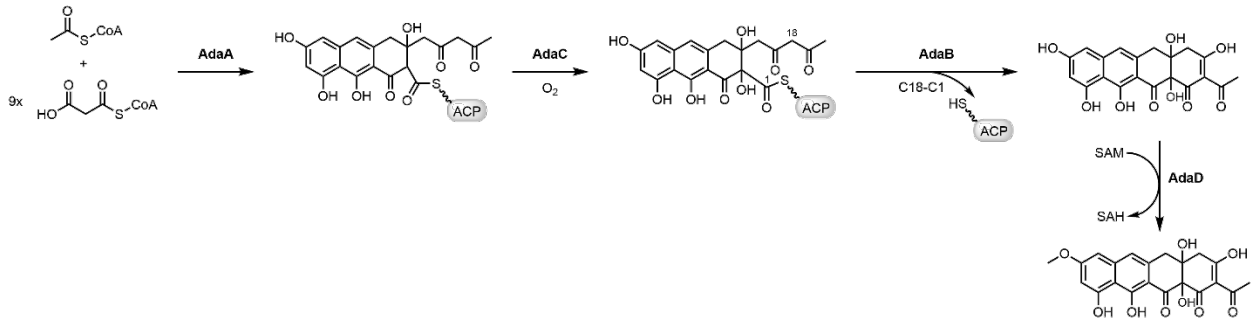


Figure 2: **TAN/BMS Biosynthesis**. TAN/BMS synthesis is initiated by polyketide synthase and is chemically modified by tailoring enzymes, including an *O*-methyltransferase that transfers a methyl group to C9. Figure adapted from Li et al., 2011.

To understand the substrate specificities of TAN/BMS and examine secondary metabolite optimization through methyltransferase activity, the replacement of the native TAN/BMS methyltransferase with other intraspecies NRRL3 methyltransferases was explored. To this end, the objectives of this thesis are as follows:

(1) Construct an intraspecies NRRL3 methyltransferase library;

Rationale: Within the NRRL3 genome are numerous methyltransferases that can be used to test the substrate specificity of TAN/BMS. Candidate methyltransferases will be selected based on bioinformatic and phylogenetic analyses, including the evaluation of amino acid sequence similarity and protein functional domains that conduct methyltransferase activity.

(2) Express the NRRL3 methyltransferase library in a different gene locus to test a pipeline for secondary metabolite production and optimization;

Rationale: Methyltransferase replacement in the overexpressed glucoamylase A locus will be tested for exploiting TAN/BMS production and altering methyltransferase patterning. Based on these findings, I could expand this proof of concept to test methylation in other biosynthetic gene clusters and modifications using other tailoring enzymes.

(3) Test methyltransferase selectivity by examining TAN/BMS methylation.

Rationale: Examining TAN/BMS production and methyltransferase activity through mass spectrometry and NRM will provide information about the substrate preferences of different intraspecies NRRL3 methyltransferases. Based on these findings, putative substrate preferences can be assigned to different methyltransferases.

2. Materials and Methods

2.1 Materials

Chemicals and solutions

The following were purchased from Sigma Aldrich (St. Louis, MO): 5 M betaine solution, polyethylene glycol-4000 (PEG-4000), 5 M ammonium acetate, uracil, sodium hydroxide, hydrogen chloride, acetic acid glacial, aurintricarboxylic acid ammonium salt, formic acid, trifluoroacetic acid, dithiothreitol, urea, HEPES, iodoacetamide.

The following were purchased from New England Biolabs (NEB) (Ipswich, MA): NEB Phusion® GC Buffer, 10mM dNTP solution mix, 2 × Phusion® High-Fidelity DNA Polymerase, NEBuffer 2.1™, 25 mM dCTP, 25 mM dGTP, PsiI-v2 NEB™ restriction enzyme, 10 × rCutSmart Buffer, NEB Quick CIP™ (5 U/ µL), 1 × NEB CutSmart Buffer™, T4 DNA polymerase.

The following were purchased from ThermoFisherScientific (Waltham, MA): acetonitrile, ethanol, acetone, HPLC grade 100% methanol, glycerol, tween 80, ampicillin sodium salt, dimethyl sulfoxide-d6 (DMSO-d6), ammonium bicarbonate.

The following were purchased from Bioshop Canada Inc. (Burlington, ON): tris, tris hydrochloride, D-glucose, SDS, agarose, potassium acetate, sodium chloride, sorbitol, calcium chloride salt, isopropanol, 5-fluoroorotic acid (5-FOA), magnesium sulfate, EDTA (disodium salt dihydrate)

The following were purchased from Pierce Biotechnology (Waltham, MA): Trypsin/Lys-C protease mix, MS grade.

The following were purchased from Roche (Basal, Switzerland): cOmplete protease inhibitor cocktail tablet.

The following were purchased from Lucigen Corporation (Middleton, WI): 10 × RNase, RNase I.

The following were purchased from Invitrogen (Waltham, MA): phenol:chloroform:isoamyl alcohol (PCI).

The following were purchased from J VWR International J.T. Baker (Radnor, PA): sodium acetate, ethyl acetate.

The following were purchased from Novozymes (Bagsværd, Denmark): VinoTaste® Pro (86.20% Maltodextrin, 4% water, 3% Polygalacturonase, 3% beta-glucanase, 2% Trisodium citrate dihydrate, 1% Potassium chloride, and 0.80% citric acid).

Culture and media

The following were purchased from Bioshop Canada Inc. (Burlington, ON): maltose, sodium chloride, 5-FOA, agar, D-glucose, magnesium sulfate, sucrose, sodium nitrate, bio-tryptone, magnesium sulfate.

The following were purchased from Becton, Dickinson and Company (Franklin Lakes, NJ): Yeast nitrogen base without amino acids and ammonium sulfate, vitamin assay casamino acid.

The following were purchased from Sigma Aldrich (St. Louis, MO): uracil, sodium hydroxide.

The following were purchased from ThermoFisherScientific (Waltham, MA): peptone, potassium chloride, magnesium chloride, sodium dihydrogen phosphate monohydrate.

2.2 Strains and controls

The reference wild type strain used in analyses was *A. niger* NRRL 2270 (ATCC 1015, CBS 513.88) (Pel *et al*. 2007). The strain NRRL3 v1.0 (ATCC 9029, CBS 120.49) was used to conduct gene-editing and perform bioinformatic analyses (Aguilar-Pontes *et al*. 2018; Vesth *et al*. 2018). The mutant strain NRRL3 Δ*pyr*GΔ*kus*A confers a uracil auxotrophy selection marker (*i.e., pyr*G, NRRL3_03466) (Carvalho *et al*. 2010; Arentshorst *et al*. 2015). This strain is also genetically manipulated to promote DNA repair by homologous recombination (*i.e., kus*A, NRRL3_03799) (Meyer *et al*. 2006). Gene coordinates and strain names used in experiments are indicated in **Table 5** (NRRL3 v2.0, https://www.fungalgenomics.ca/).

Table 5: *A. niger* strains and auxotrophy genes

Gene ID	Genotype	Phenotype	Reference
NRRL3	N/A	Wild type	(Aguilar-Pontes *et al*. 2018)
NRRL 2270	NRRL 2270 Δ*pyr*G Δ*kus*A	Orotidine 5'-phosphate decarboxylase PyrG (NRRL3_03466). ATP-dependent DNA helicase Ku70, involved in non-homologous end joining DNA double strand break repair (NRRL3_03799).	(van Hartingsveldt *et al*. 1987; Mattern *et al*. 1987)

Gene ID	Genotype	Phenotype	Reference
NRRL3_09545^{OE}	NRRL 2270 Δ*pyr*G Δ*kus*A NRRL3_09545^{OE}	TAN/BMS transcription factor overexpression strain, producing TAN/BMS.	This study
NRRL3_09545^{OE} 09546^{KO}	NRRL 2270 Δ*pyr*G Δ*kus*A NRRL3_09545^{OE}09546^{KO}	TAN/BMS transcription factor overexpression strain, native methyltransferase knockout strain producing TAN/BMS without C9 methylation.	This study
NRRL3_09545^{OE}09546^{KO} Δ*gla*A::09546	NRRL 2270 Δ*pyr*G Δ*kus*A NRRL3_09545^{OE}09546^{KO} Δ*gla*A::09546	TAN/BMS transcription factor overexpression strain, native methyltransferase knockout, overexpression of the native methyltransferase by gene replacement of the α-glucoamylase.	This study
NRRL3_09545^{OE}09546^{KO} Δ*gla*A::NRRL3 _methyltransferase library	NRRL 2270 Δ*pyr*G Δ*kus*A NRRL3_09545^{OE}09546^{KO} Δ*gla*A::MT library	TAN/BMS transcription factor overexpression strain, native methyltransferase knockout, overexpression of the native methyltransferase by gene replacement of the α-glucoamylase.	This study

The biosynthetic gene cluster producing the tautomeric secondary metabolite neuropeptide Y inhibitor TAN and neurokinin-1 receptor antagonist BMS encodes the NRRL3 genes listed in **Table 6**. (NRRL3 v2.0, https://www.fungalgenomics.ca/).

Overexpression of the transcription factor (NRRL 09545) resulted in the strain NRRL3_09545^{OE}. The subsequent knockout of the *O*-methyltransferase gene (NRRL 09546) encoded in the TAN/BS biosynthetic gene cluster resulted in the strain NRRL3_09545^{OE} 09546^{KO}. Gene coordinates and strain names used in experiments are indicated in **Table 5** (NRRL3 v2.0, https://www.fungalgenomics.ca/).

Table 6: TAN/BMS gene architecture

Gene ID	Gene Function	Reference
NRRL3_09545	Fungal-specific transcription factor	(Li *et al*. 2011)
NRRL3_09546	*O*- methyltransferase, AdaD	
NRRL3_09547	FAD-dependent monooxygenase	
NRRL3_09548	Metallo- betalactamase -typethioesterase	
NRRL3_09549	Polyketide synthase	
NRRL3_09550	Major facilitator superfamily protein, efflux pump	

Insertion of an *O*-methyltransferase gene at the *gla*A locus resulted in NRRL3_09545^{OE}09546^{KO} *gla*A::NRRL3_methyltransferase gene, where a total of 12 different *O*-methyltransferase genes were inserted at the *gla*A locus (**Table 7**) (NRRL3 v2.0, https://www.fungalgenomics.ca/).

Table 7: *O*-methyltransferase genes

NRRL3 gene ID	Chromosomal Location	NRRL3 gene function
NRRL3_06190	chr_4_2:3521382..3522800	O-methyltransferase
NRRL3_07487	chr_6_1:161659..163129	
NRRL3_08343	chr_6_1:2951908..2953354	
NRRL3_08383	chr_6_1:3061649..3063091	
NRRL3_08389	chr_6_1:3077764..3079484	
NRRL3_08972	chr_7_1:504491..505847	
NRRL3_10125	chr_7_2:2667459..2668550	
NRRL3_10146	chr_7_2:2732884..2734409	
NRRL3_00997	chr_1_1:261371..263287	Methyltransferase domain
NRRL3_06013	chr_4_2:2971860..2972986	
NRRL3_07488	chr_6_1:163483..164508	
NRRL3_10369.1	chr_7_2:2709316..2711475	

2.3 Preparation of Δ*pyr*G overexpression and knockout strain

The *A. niger* overexpression and knockout strain conferring the genotype NRRL3_09545^{OE}09546^{KO}Δ*pyr*G was prepared by ejecting the uracil auxotrophy gene, orotidine 5'-phosphate decarboxylase (*pyr*G) (Sun *et al*. 2013) from the strain NRRL3_09545^{OE}09546^{KO}*pyr*G+. Eviction of *pyr*G plasmid can be performed through successive plating on minimal media 1% maltose solid agar plates supplemented with mutagenic 5-FOA and 100 mM uridine. Following five generations of plating in the presence of 5-FOA, spores were grown on non-selective regeneration media (NSRM) containing uracil, and selective regeneration media (SRM) excluding uracil to examine positive and negative growth conditions, respectively. Spores were harvested from the NSRM plates using a 1× saline tween mixture, and counted using a BioRad TC20 automated cell counter (Bio-Rad Laboratories, Hercules, CA). Spore suspensions were adjusted to a final concentration of $1 \times 10^6 – 1 \times 10^7$ spores/mL. Spores were used immediately for protoplast digestion (**Section 2.8**), or either placed at 4 °C for short-term storage, or combined in a 1:1 ratio with 40% glycerol for long-term storage at -80 °C.

2.4 Media and culture conditions

Plates and media composition

An *A. niger* minimal media solution containing 1% maltose was used for *A. niger* growth and was prepared as a liquid media or solid agar plates. Minimal media 1% maltose was made using double distilled water, maltose, a 2.25M $MgSO_4$ solution, and 20X nitrate salts stock solution. An *A. niger* complete media solution was used for *A. niger* growth and was prepared as a liquid media. Complete media was made using double distilled water, a 20x nitrate salts stock solution, D-Glucose, peptone, yeast extract, casamino acids, L-Uracil, a 1000X vitamin stock solution , a 2.25M $MgSO_4$ solution, and a 100mM uridine solution.

NSRM and SRM was prepared as solid agar plates and used for *A. niger* growth. NSRM was made using double distilled water, sucrose, L-Uracil, sodium nitrate, potassium chloride, potassium phosphate monobasic, a 4M KOH solution, magnesium sulfate heptahydrate, a 100mM uridine solution, and agar. SRM was made with sucrose, sodium nitrate, potassium chloride, potassium phosphate monobasic, a 4M KOH solution, magnesium sulfate heptahydrate, and agar.

LB liquid media and 2YT liquid media were used for *E. coli* growth. LB and 2YT was made using double distilled water, Bacto-Tryptone, yeast extract, and NaCl.

Spore suspension preparation

Spores were grown on solid media and harvested using 1 – 5 mL of a 1× saline tween mixture. A cell spreader was used to gently rub the surface of the plate to make a spore suspension. The spore suspension was concentrated to $1 \times 10^6 – 1 \times 10^7$ spores/mL using a BioRad TC20 automated cell counter.

Standing cultures

Miniaturized assays were used to produce small volumes of extracellular metabolites (*e.g.,* for spectrophotometry or metabolomic analyses). Spores were grown in a 96-well culture plate in 200 µL of minimal media 1% maltose and combined with 15 µL of spores concentrated at $1 \times 10^6 – 1 \times 10^7$ spores/mL. Spores were incubated at 37 °C and grown for 5 days.

Larger assays were used to produce large volumes of extracellular metabolites (*e.g.,* NMR). Spores were grown in a 25 × 100 mm petri dish in 100 mL of minimal media 1% maltose and combined with 100 µL of spores concentrated at $1 \times 10^6 – 1 \times 10^7$ spores/mL. Spores were incubated at 37 °C and grown for 5 days.

Shaking cultures for protoplastation

Complete media was used to produce large volumes of *A. niger* mycelia in an overnight culture. Sealable 1 L culture flasks were prepared in parallel with each containing 100 mL of media and 2 – 10 mL of spores concentrated to $1 \times 10^6 – 1 \times 10^7$ spores/mL. Mycelia cultures were grown at 32 °C for 16 – 18 hours in an orbital shaker set to 220 RPM (Infors HT Multitron Standard, Anjou, QC).

Shaking cultures for harvesting gDNA

Complete media was used to produce small volumes of *A. niger* mycelia in an overnight culture. 50 mL tubes containing 5 mL of complete media was combined with 0.5 – 1 mL of spores concentrated to $1 \times 10^6 – 1 \times 10^7$ spores/mL. Mycelia cultures were grown at 32 °C in an orbital shaker set to 220 RPM.

Transfer cultures for MS based proteomics

A 100 mL mycelia culture was prepared in the same manner as *shaking cultures for protoplastation*. The mycelia culture was separated from the gelatinous homogenate by filtration using a 30 mL Büchner funnel lined with a 50 mm × 50 mm square of sterile Calbiochem® Miracloth (pore size 22 – 25 µm) (Millipore Sigma, Burlington, MA). Strained mycelia were then transferred to 100 mL of inducing media (minimal media 1% maltose) and incubated for 8 hours at 32 °C in an orbital shaker set to 220 RPM.

Bacterial culturing

E. coli plasmids were grown in LB or 2YT media. 13 mL culture tubes containing 2 mL of media were combined with 20 µL of a concentrated *E. coli* glycerol stock and 2 µL of 100× ampicillin. Primary cultures were grown for a minimum of 6 hours, and 1 mL of the solution was used to inoculate a secondary culture composed of an additional 2 mL media and 3 µL of 100× ampicillin solution. Secondary cultures were grown for 10 – 14 hours in an orbital shaker set to 220 RPM.

2.5 Phylogenetic and bioinformatic analyses

Song *et al.,* 2018 was used as a guideline for phylogenetic tree building and analysis (Song, Wu, *et al*. 2018). Amino acid sequences were obtained for each methyltransferase gene from the annotated NRRL3 genome database (NRRL3 v2.0, https://www.fungalgenomics.ca/) (Aguilar-Pontes *et al*. 2018; Vesth *et al*. 2018) and organized in a single file. The alignment of multiple amino acid sequences was performed using the free software MUltiple Sequence Comparison by Log- Expectation (MUSCLE) (Edgar, 2004). Sequence data was output as a FASTA file, and the free software Jalview Desktop (Waterhouse *et al*., 2009) was used to visualize amino acid sequence alignment. Jalview desktop was also used to perform gap-editing of amino acid sequences, where areas with poor or no amino acid overlap were manually removed from the alignment file. Tree construction was performed using the free software FastTree v2.1.11 (Price *et al*., 2010, Price *et al*., 2009), and phylogenetic tree visualization and editing was performed using FigTree (Rambaut, 2010). In FastTree, maximum-likelihood phylogenetic trees were constructed using a neighbor joining clustering algorithm that interprets distances between amino acid sequences to form branch lengths in the tree. Two additional branch length procedures (nearest-neighbor interchanges and subtree-prune-regraft moves) were used with the BLOSUM45 algorithm to improve clustering in the phylogenetic tree. In FigTree, phylogenetic tree rooting was first performed to visualize outliers, and to observe phylogenetic tree shape. Bootstrap values were calculated using FastTree were labeled directly on the tree to determine the confidence level for each node.

Clusters were assigned based on the appearance of clades with bootstrap values > 70% and confirmed using a pairwise comparison matrix that was scripted in Perl using BioPerl modules (Stajich et al., 2002). In doing so, protein sequences were obtained for each methyltransferase gene from the annotated NRRL3 genome database (NRRL3 v2.0, https://www.fungalgenomics.ca/) and BioPerl modules were run. The resulting matrix was organized manually based on a baseline cut-off of 20% sequence similarity. The average percent identity of sequences within and between the clusters was calculated to confirm the appearance of true clusters.

2.6 Primer design and polymerase chain reaction

Oligonucleotides were designed using the software Geneious Prime® 2021.2.2 and were based on complementary nucleotides from the manually curated and annotated NRRL3 genome (NRRL3 v2.0, https://www.fungalgenomics.ca/) (Aguilar-Pontes *et al*. 2018; Vesth *et al*. 2018).

Oligonucleotides were designed for polymerase chain reaction (PCR) as 16 – 20 base pair sequences in the direction 5’ – 3’ or 3’ – 5’, and contained < 50% GC content, and melting temperatures were within ± 5 °C of each other. CRISPR-Cas9 sgRNA were also designed using Geneious Prime® 2021.2.2, and were constructed to maximize activity and specificity scores (Hsu *et al*. 2013; Doench *et al*. 2016). Oligonucleotides and primers were ordered from Integrated DNA Technologies (Integrated DNA Technologies, Coralville, IA) and concentrated to 100 mM.

List of primers

A series of primers were designed to amplify regions of the *A. niger* genome and function as a tool for validating the insertion of genes. Based on the designs described in **Section 2.6** and **2.7**, primer pairs were made to amplify regions of DNA and visualize gene insertions in

NRRL3_09545^{OE}09546^{KO} and NRRL3_09545^{OE}09546^{KO} *gla*A::NRRL3 methyltransferase gene. Primer pairs are described in **Table 8**.

PCR setup

PCR was performed in a reaction volume of 50 µL using a PCR thermocycler (ThermoFischer Scientific, SimpliAmp™ Thermal Cycler, Waltham, MA). For each PCR, a mixture of 10 µL of NEB Phusion® GC Buffer, 10 µL 5 M betaine solution, 1 µL10 mM dNTP solution mix, 0.25 µL 100 mM forward and reverse primers (**Table 8**), 0.5 µL 2 × Phusion® High-Fidelity DNA Polymerase, and a mixture of nanopure water and 10 – 20 ng template DNA was combined in the remaining 28 µL volume. PCR cycling depended on the amplicon and template DNA. In general, an initial denaturing step of 5 minutes (98 °C) was used for both plasmid and genomic DNA (gDNA), followed by 35 cycles consisting of a secondary 30 second denaturing step (98 °C), annealing for 30 seconds (various temperature based on GC content, and length of primers), amplification of 1 minute per 1 kb of DNA (72 °C), and a final extension step of 6 minutes (72 °C).

Table 8: Primers used to amplify gDNA in *A. niger*

Primer ID	Sequence
NRRL3_09545^{OE} 09546^{KO} • Forward primer • Reverse primer	• 5'-CCAGGCCAGTGAGGCCAG-3' • 3'-CCCCTGGGGAAAGCAATG-5'
NRRL3_09545^{OE} 09546^{KO} *gla*A:: NRRL3 methyltransferase gene • Forward primer • Reverse primer	• 5'- TTGGTGTGTGGTATGACTAAATG -3' • 3'- ATCTGAATCTAGACACACTTGGC -5'

2.7 *O*-Methyltransferase gene library construction

O-methyltransferase gene plasmid preparation

O-methyltransferase genes of interest were individually cloned into the pjetC *E. coli* plasmid using the ligation independent cloning (LIC) principle (Aslanidis and de Jong 1990). Genomic DNA (gDNA) was extracted from the wild type strain and oligonucleotides corresponding to individual *O*-methyltransferase genes were used as a template to drive PCR. Oligonucleotides were ordered with the following sequences: 5'- GTACTTCCAATCCAATCCATTTTTA -3' and 3'- GTTATCCACTTCCAATCCATTTTTA -5' (*i.e.,* LIC regions), which are found flanking *O*-methyltransferase insertion genes and served as a homologous region for insertion into pjetC. The pjetC plasmid comprises the following genes and selection markers: *bla* to confer ampicillin resistance, lacUV5 to drive the transcription of genes, pMB1 for replication of the plasmid, *gla*A promoter and terminator sequences, an additional 100 bp forward and reverse *gla*A sequence.

Following PCR and the construction of individual *O*-methyltransferase genes containing the LIC region, PCR products and plasmid DNA were purified. 10 µL of each PCR product (30 ng/µL) was combined with 2 µL of 10 × NEBuffer 2.1™, 2 µL 25 mM dCTP, 0.8 µL 100 µM DTT, 1 µL 3 U/µL T4 DNA polymerase, and 4.2 µL nanopure water for a total reaction volume of 20 µL. Samples were incubated at 22 °C for 35 min and heat inactivated at 72 °C for 20 min.

In parallel, pjetC *E. coli* plasmids were linearized using 1 µL (10,000 units/mL) of PsiI-v2 NEB™ restriction enzyme, 5 µL 10 × rCutSmart Buffer, 1 µg plasmid DNA, and up to 50 µL of nanopure water. Samples were incubated for 1 hour at 37 °C and heat inactivated at 65 °C for 20 min. Dephosphorylation was performed on the linearized vector using 10 µg of plasmid DNA, 10 µL NEB Quick CIP™ (5 U/µL), 20 µL 1 × NEB CutSmart Buffer™, and 70 µL nanopure water for a total reaction volume of 100 µL. Samples were incubated 37 °C for 1 hour. 1 µL 40 ng/µL linearized plasmid DNA was combined with 2 µL of 10 × NEBuffer 2.1™, 2 µL 25 mM dGTP, 0.8 µL 100 µM DTT , 1 µL 3 U/µL T4 DNA polymerase, and 4.2 µL nanopure water for a total reaction volume of 20 µL. Samples were incubated at 22 °C for 35 min, and heat inactivated at 72 °C for 20 min.

Cloning was performed by combining up to 100 ng from an individual T4 DNA polymerase treated *O*-methyltransferase gene with 100 µg of the linearized and T4 DNA polymerase treated plasmid. Samples were incubated in a PCR thermocycler set to 60 °C with a ramp decreasing temperature at a rate of 0.1 °C/second. Transformation in DH5α competent cells was performed by incubating samples on ice for 30 minutes followed by heat shock at 42°C for 45 sec, and subsequently incubation on ice for 2 min. 800 µL of SOC was gently added to samples and incubated at 37 °C for 45 min and at 220 RPM. Samples were then centrifuged at 10,000 RPM and 500 µL of the supernatant was removed. The remaining supernatant was used to gently resuspend the pellet and 50 µL was used for plating on LB + ampicillin and incubated at 37 °C for 12 hours.

PjetC plasmid DNA harbouring *O*-methyltransferase genes was extracted via plasmid preparation to yield plasmid DNA. Culturing was performed as per *Bacterial culturing* (**Section 2.4**). 1 mL aliquots of culture were transferred to 1.5 mL Eppendorf tubes and centrifuged at 13,200 RPM at 4 °C for 5 min to pellet the plasmid. Pellets were vigorously resuspended in 100 µL tris buffer (100 mM Tris-HCl, pH 8.0, 20 mM EDTA, 50 mM D-glucose) and incubated at room temperature for 1 min. 250 µL of extraction solution (1 M NaOH, 10% SDS) was gently added to the sample and mixed. 150 µL of lysis solution (3 M potassium acetate, pH 4.8, 1.83 M acetic acid) was added to the sample and inverted several times, and samples were incubated on ice for 5 min. Samples were centrifuged at 13,200 RPM at 4 °C centrifuge for 5 min to pellet the digested cell components. The aqueous phase was transferred to a 1.5 mL Eppendorf tube and precipitated in 0.7 × 100% isopropanol. Samples were centrifuged at 13,200 RPM at 4 °C for 5 min. The supernatant was removed, and the pellet of plasmid DNA was washed with 70% ethanol and centrifuged at 13,200 RPM at 4 °C for 5 min. Ethanol was removed and the plasmid DNA pellet was dried and resuspended in elution buffer (10 mM Tris, pH 8.0).

To expand materials and increase the efficiency for *A. niger* transformation, plasmid DNA harboring *O*-methyltransferase genes was amplified using PCR. Primer pairs were used to amplify

the *gla*A promoter and terminus region that is found in the pjetC plasmid flanking inserted *O*-methyltransferase genes. Using the primers NRRL3_09545^{OE}09546^{KO}*gla*A::NRRL3 methyltransferase gene (**Table 8**), PCR was performed as per *PCR setup* (**Section 2.6**). For this primer pair, an annealing step of 63 °C was used. Elongation time differed for *O*-methyltransferase genes, ranging from 1 min to 2.30 min. PCR product was purified using a 1:10 volume of 3 M sodium acetate (pH 5.8) and precipitated in 1 volume of 100% isopropanol. Samples were incubated at -20 °C for a minimum of 10 min, and centrifuged at 13,200 RPM at 4 °C for 30 min. The supernatant was removed, and the pellet of DNA was washed with 70% ethanol and centrifuged at 13,200 RPM at 4 °C for 10 min. Ethanol was removed and the DNA pellet was dried and resuspended in nanopure water.

CRISPR-Cas9 sgRNA preparation

The homologous insertion of *O*-methyltransferase genes at the *gla*A locus was directed by CRISPR-Cas9 sgRNA using the shuttle vector ANEp8 (Song, Ouedraogo, *et al*. 2018). The ANEp8 vector comprises the following genes and selection markers: AMA1 to enable extra-chromosomal replication, a tRNA promoter that was experimentally verified to drive the transcription of genes (Song, Ouedraogo, *et al*. 2018), *bla* to confer ampicillin resistance, *pyr*G to retore uracil prototrophy in the Δ*pyr*G parental strain, the pyruvate kinase promoter gene *pki*A as a strong inducible promoter, a *gla*A terminator sequence, a nuclear localization signal (*i.e.,* SRADPKKKRKV) for nuclear transport, and a sgRNA cassette consisting of crRNA for gene targeting and tracrRNA for Cas9 binding. sgRNA sequences were designed using Geneious Prime® 2021.2.2 (**Section 2.6**), and ANEp8 plasmid construction and sgRNA insertion was ordered through integrated DNA technologies (Integrated DNA Technologies, Coralville, IA).

2.8 *A. niger* protoplast preparation

All steps of protoplast preparation were performed under sterile conditions — surfaces were cleaned with 70% isopropanol and a grid top gas burner was tuned on to maintain a sterile environment. Previously published methods were used as a guideline, including Debets and Bos, 1986, Storms *et al.,* 2006, and Werner *et al.,*.

2.8.1 Culturing protoplasts

A 2 – 10 mL aliquot of concentrated spores (**Section 2.4**) conferring the genotype NRRL3_09545^{OE}09546^{KO}Δ*pyr*G was combined with 100 mL of complete media in a sealed 1 L flask to permit adequate gas exchange, and was incubated for 16 – 18 hours in an orbital shaker set to 220 RPM and 32 °C.

2.8.2 Enzymatic digestion of protoplasts

Prior to mycelia harvesting, six pre-weighed 50 mL centrifuge tubes each containing 20 mL of an osmotic medium (1 mM phosphate buffer, pH 7.9, 1.2 M $MgSO_4$), used to maintain protoplast integrity, and 1 g of a commercially available enzymatic solution (VinoTaste® Pro), used to digest chitinous fungal cell walls, were prepared and filter sterilized though a 0.45 µm MilliporeSigma™ Millex™-HA Sterile Syringe Filter (MilliporeSigma, Burlington, MA).

2.8.3 Harvesting digested protoplasts

Following digestion, a 30 mL Büchner funnel lined with a 50 mm × 50 mm square of sterile Miracloth (pore size 22 – 25 μm) was assembled over a waste container to filter mycelia from the gelatinous homogenate. Miracloth was moistened with 5 mL of a 0.6 M $MgSO_4$ solution to promote adhesion to the funnel. The mycelia culture flask was opened and flamed, and slowly dispensed into the Miracloth-lined funnel. Mycelia were washed with 600 mL of 0.6 M MgSO4 and sterile tweezers were used to remove excess liquid. Using pre-weighed 50 mL centrifuge tubes containing the osmotic media and enzyme solution, 1 g of dried mycelia was added to each tube and mixed by gently inverting. Tubes were loaded into an orbital shaker set to 75 RPM and 37 °C to undergo digestion, and a 0.5 mL sample was aseptically removed from one tube and examined under a microscope (Nikon TMS Inverted Microscope, Nikon Instruments Inc., Melville, NY) at 30-minute intervals. Images were captured using an associated 11.8mm x 8.9mm camera (Diagnostic instruments Inc., Sterling Heights, MI). Mycelia undergoing digestion were examined at 20 × magnification to monitor the appearance of protoplasts. Mycelia digestion was continued for 2 – 3 hours until approximately 90% of protoplasts emerged.

Following incubation, protoplast tubes were cooled on ice for 10 min to stop digestion. For every 20 mL of protoplasts, a 50 mL centrifuge tube with 20 mL of ice-cold trapping buffer (1 M Tris-HCL, pH 7.5, 0.6 M sorbitol) was prepared. A 20 mL aliquot of protoplasts was slowly added to the bottom of the trapping buffer layer using a 50 mL serological pipette. Tubes were loaded in a preset 4 °C centrifuge and spun for 20 min at 3200 × g to augment the separation of protoplasts from the osmotic media layer. When a beige-coloured interface was visible between the layers, centrifugation was stopped, otherwise centrifugation was repeated for another 20 min.

Without aspirating the interface, protoplast layers were collected using a wide-tip serological pipette to avoid damaging fragile protoplasts. To increase the concentration of protoplasts, three individual layers were combined in a 50 mL centrifuge tube, and ice-cold sorbitol-calcium solution (1 mM sorbitol, 50 μM calcium chloride) was added to a final volume of 45 mL, and gently inverted. Tubes were loaded in a preset 4 °C centrifuge, and spun for 10 min at 2400 × g. The supernatant was removed, and the pellet of protoplasts — appearing as a mixture of dark brown debris and beige protoplasts — was further separated by gently pipetting up and discarding the dark debris. The remaining protoplasts were resuspended with additional sorbitol-calcium solution and centrifuged at 4 °C for 10 min at 2400 × g and was repeated until the protoplast pellet was free of debris. The supernatant was decanted, and the protoplast pellet was resuspended in 10 mL of sorbitol-calcium solution.

2.8.4 Counting protoplasts

Three *df* 10 serial dilutions were performed on a sample of protoplasts to achieve 10^{-1}, 10^{-2}, and 10^{-3} dilutions using sorbitol-calcium solution. Using the 10^{-3} diluted protoplast mixture, 10 μL was loaded into a 0.1000 MM deep hemocytometer (Propper Lumicyte® blood counting chamber (Long island, NY) and placed on the stage of a microscope. The average number of protoplasts were counted under 20 × magnification using the smallest grid located in the hemocytometer, where each inner square in the chamber contains a volume of 0.1 mm^3 / 25 squares. The average number of protoplasts was used to calculate the volume of protoplasts required to concentrate the 10 mL mixture to 1×10^8 protoplasts/mL and was estimated using **Equation 1**. Protoplasts were

used immediately for transformation (**Section 2.7**), or combined in a 2:1 ratio of sorbitol-calcium solution and 10 M glycerol (92% w/v) for long-term storage at -80°C.

Equation 1: Estimating protoplast concentration

$$Concentration\ of\ protoplasts\ (protoplasts\ per\ mL\ ) = n\ \times df\ x\ V$$

Where *n* = number of protoplasts; *df* = dilution factor; V = volume of hemocytometer (*i.e.,* 10^4).

2.9 Co-transformation of *A. niger*

2.9.1 Materials used for transformation

The transformation of DNA into the *gla*A region of the *A. niger* genome was performed by combining competent fungal protoplasts with a plasmid expressing a CRISPR-Cas9 guide-specific sequence, as well as a plasmid or linearized PCR fragment serving as a repair template for homologous recombination.

Protoplasts

Protoplasts conferring the genotype NRRL3_09545^{OE}09546^{KO}Δ*pyr*G were prepared by digesting and harvesting protoplasts (**Section 2.8**), and a 1 mL aliquot containing 1×10^8 protoplasts/mL was used for the *A. niger* transformation.

Plasmid DNA

The ANEp8-Cas9 shuttle plasmid (Song, Ouedraogo, *et al*. 2018) was used for the homologous transformation of DNA into the *gla*A region of the *A. niger* genome, where endonuclease activity was directed to *gla*A using a guide-specific sgRNA (ANEp8-Cas9-sgRNA(*gla*A). The ANEp8-Cas9 shuttle plasmid contains a gRNA cassette region flanked by *gla*A promoter and terminator sequence that served as the target region of homology. A variety of selection and auxotrophy markers are present in the ANEp8-Cas9 shuttle plasmid, including the *pyr*G selection marker that restored uracil auxotrophy in *Δpyr*G protoplasts (*i.e.,* NRRL3_09545^{OE}09546^{KO}Δ*pyr*G). An ANEp8-Cas9 shuttle plasmid that does not have specificity to a sgRNA sequence (*i.e.,* ANEp8-Cas9-empty) was used as a control for Cas9 functionality.

Repair Template DNA

A library of *O*-methyltransferase genes was designed in the plasmid pjetC (**Section 2.7**). As described in **Section 2.7**, *O*-methyltransferase gene inserts were amplified using PCR to serve as repair template fragments in transformation. In doing so, a 400-nucleotide flanking region was amplified to increasing annealing in the *gla*A promoter and terminator regions that are used to increase the efficiency for homologous recombination.

2.9.2 Transformation using *A. niger*

Protoplasts were defrosted and maintained on ice. To promote DNA function during transformation, a solution of 40 µL aurintricarboxylic acid, and 200 µL 20% PEG-4000 solution was directly mixed in a 1 mL solution of protoplasts. For each transformation, 5 – 10 µg of ANEp8-Cas9-sgRNA(*gla*A) plasmid DNA (or ANEp8-Cas9 empty) was combined with 3 – 5 µg of repair

template DNA (**section 2.7**) in 1.5 mL Eppendorf tubes and combined with 64 µL of the protoplast solution. Samples were incubated for 10 min at room temperature. 300 µL of 60% concentrated PEG-4000 solution was added to each sample and mixed by gently inverting sample tubes. Samples were incubated for 20 min at room temperature. 1.0 mL of 1.2 M sorbitol was added to each sample and centrifuged for 4 min at 12,000 × g. The supernatant was decanted, and the pellet was resuspended in 200 µL 1.2 M sorbitol. Controls (**Table 9**) and transformants were plated on SRM or NSRM to examine positive and negative growth conditions (**Section 2.3**), and incubated for 2 – 5 days at 37°C.

Table 9: Experimental controls

Positive Control	**Media**
• Test protoplast viability and serial dilution from 10^{-6} to 10^{-8} in media with uracil	NSRM
• Test protoplast competency (ANEp8-Cas9 empty) to observe uracil prototrophy in media without uracil	SRM
Negative Control	**Media**
• Test protoplast competency in media without uracil • Test CRISPR-Cas9 endonuclease activity	SRM
• Test mycelium fragments regeneration	dd H_2O

Calculating transformation efficiency

Transformation efficiency was calculated using

Equation 2 for controls in which ANEp8-Cas9- empty was used to measure the competency of protoplasts, and for *O*-methyltransferase mutants that underwent transformation using ANEp8-Cas9-sgRNA(*gla*A).

Equation 2: Transformation efficiency

$$transformation\ efficiency = \frac{n}{[DNA]} \times \frac{1}{df}$$

Where *n* = number of colonies; [DNA] = concentration of DNA in ng; *df* = dilution of DNA before plating

Purification of A. niger Colonies

Transformant plates were examined for colonies and novel morphology. Depending on the transformation efficiency, three colonies were selected from transformation plates for purification. Each colony was removed from the plate using a sterile toothpick and combined with 50 µL of a 1 × saline-tween mixture and re-plated on a fresh selection media to generate mycelia and isolate independent transformants. This process was repeated (*i.e.,* two rounds of purification). Plates were incubated for 5 days at 37 °C, or until mycelia sporulated. Spore suspensions were prepared (**Section 2.4**) for each transformant and maintained at 4°C.

2.10 Validating *A. niger* transformation

2.10.1 Genomic DNA extraction and PCR amplification

The extraction of gDNA was prepared using PCI. Spore suspensions from purified colonies were grown using *Shaking cultures for harvesting gDNA*. Mycelia was separated by filtration using a 30 mL Büchner funnel lined with a 50 mm × 50 mm square of sterile Miracloth (pore size 22-25 μm) and washed with double-distilled water. Washed mycelia were added to a mortar filled with 100 mL of liquid nitrogen and ground to fine powder using a pestle. 100 mg of frozen and ground mycelia were combined with 500 μL of gDNA extraction buffer (100 mM Tris-HCl, pH 8.0, 20 mM Na_2EDTA, 0.5 M NaCl, 1% SDS) in a 1.5 mL Eppendorf tube and vortexed vigorously for 30 seconds. 250 μL of PCI was added to each tube and vortexed vigorously for 30 seconds. Tubes were centrifuged for 5 minutes at 16,000 × g. The organic phase was transferred to a 1.5 mL Eppendorf tube and combined with 250 μL of PCI. Vortexing and centrifugation was repeated. The organic phase was transferred to a 1.5 mL Eppendorf tube and the DNA mixture was precipitated in 0.6 × volume of 100% isopropanol. Samples were placed at -20 °C for a minimum of 10 minutes to improve the DNA yield. Samples were centrifuged at 16,000 × g for 15 minutes. White coloured DNA pellets were identified, and the supernatant was gently discarded. Pellets were rinsed in 750 μL 70% ethanol and samples were centrifuged at 16,000 × g for 15 minutes. The supernatant was discarded, and remaining ethanol was evaporated by placing samples at room temperature until pellets were dry (approximately 20 minutes). Dried pellets were resuspended in 85 μL elution buffer (10 mM Tris, pH 8.0), 10 uL 10 × RNase, 5 uL RNase I, and 2 uL RNase A. Samples were incubated for 60 minutes at 37 °C, and enzyme solutions were heat inactivated by incubating for 15 minutes at 70 °C. 300 μL of PCI and 300 uL elution buffer was added to each sample, and samples were vigorously vortexed for 30 seconds. Samples were centrifuged at 16,000 × g for 5 minutes. The organic phase was transferred to a 1.5 mL Eppendorf tube and precipitated in 0.5 × volume of 7.5 M ammonium acetate and 2.5 × volume of 100% ethanol. Samples were vortexed vigorously for 30 seconds and placed at -20 °C for overnight incubation. Samples were centrifuged at 16,000 × g for 15 mins and the supernatant was gently discarded. Pellets were rinsed in 750 μL 70% ethanol and samples were centrifuged at 16,000 × g for 5 minutes. The supernatant was gently discarded, and remaining ethanol was evaporated by placing samples at room temperature until pellets were dry (approximately 20 minutes). Pellets were resuspended in a minimum of 20 μL nanopure water. gDNA quality was tested using gel electrophoresis using a 1% agarose gel. Extracted gDNA was stored at -20 °C.

To validate the insertion of *O*-methyltransferase genes at the *gla*A locus, PCR amplification was pursued using the same primer pairs and conditions listed in *O-methyltransferase gene plasmid preparation* (**Section 2.7**).

2.10.2 DNA sequencing

20 μL unpurified PCR product (approximately 0.8 – 1.0 μg/μL) was used as the DNA template for Sanger sequencing. Primers were diluted using 5 μL 100 mM forward or reverse primers and 95 μL nanopure water. Samples were processed at Génome Québec (Génome Québec, Montreal, QC).

The analysis and comparison of DNA sequences to theoretical constructs was performed manually. First, the length of sequences was trimmed by removing unidentified nucleotides. Trimmed sequences representing forward and reverse constructs were manually aligned using MUSCLE (Edgar, 2004) and compared to the theoretical DNA sequence. Theoretical and DNA sequence alignment files were visualized using the free software Jalview Desktop (Waterhouse *et al*., 2009). Differences in the composition of sequences were noted.

2.10.3 Extracellular spectrophotometry analysis

Spores were cultured in a miniaturized assay as per the directions described in *Standing cultures* (**Section 2.4**). An aliquot of 100 µL of the supernatant was transferred to a 1.5 mL Eppendorf tube and diluted between 6 – 10-fold. Samples were loaded in a 96-well plate (such as a Greiner 96 Flat Bottom Transparent Polystyrene plate) and loaded into a spectrophotometer (Tecan Infinite® 200 PRO, Männedorf, Switzerland). The absorbance of DNA was measured between 250 – 500 nm with a wavelength step of 2 nm. Data was plotted and the maximum wavelength of each sample was noted and compared to a theoretical maximum.

2.10.4 Extracellular metabolomic analysis

Spores were cultured in a miniaturized assay as per the directions described in *Standing cultures* (**Section 2.4**). An aliquot of 100 µL of the supernatant was transferred to a 1.5 mL Eppendorf tube and centrifuged at 13,000 × g at 4 °C for 30 min to pellet debris. The supernatant was transferred to a new 1.5 mL Eppendorf tube and vortexed. An aliquot of 40 µL was transferred to a 1.5 mL Eppendorf tube and combined with 80 µL of chilled HPLC grade 100% methanol. Samples were incubated on ice for 10 min to increase protein precipitation. Samples were then centrifuged at 13,000 × g at 4 °C centrifuge for 20 min to pellet precipitated protein. The supernatant was transferred to a new 1.5 mL Eppendorf tube and vortexed. An aliquot of 50 µL was transferred to a 96-well culture plate and combined with 50 µL of 0.1 M formic acid in nanopure water. 96-well culture plates were stored at -20 °C.

Analysis of metabolites was conducted using a 7-Tesla Finnigan Linear Ion Trap Mass spectrometer (LTQ) combined with a Fourier Transform Ion Cyclotron Resonance Mass Spectrometer (FT) (Thermo Electron Corporation, San Jose, CA). A volume of 10 µL was injected into a Kinetex 150 × 2.1 mm, 5 µm, C18 column (Phenomenex, Torrence, CA, USA) to separate compounds using an Agilent 1260 Infinity II HPLC system (Agilent Technologies, Santa Clara, CA, USA). The following solvents were used to generate a reversed-phase separation gradient: 0.1% formic acid in water (Solvent A), 0.1% formic acid in acetonitrile (Solvent B). Solvent B flow rate (250 nL/min) was applied over the following linear gradient: 3% (1 min), 3 – 80% (10 min), 80 – 95% (0.1 min), held at 95% (1 min), 95 – 3% (0.1 min), and held at 3% (4.8 min). Analysis was performed in the LTQ-FT using an ionization voltage of 4900 V (positive, in electrospray ionization mode). Metabolites were analyzed in a scan range of 100 to 1400 *m/z* at 100000 resolution at 200 *m/z*. Data was visualized using Thermo Xcalibur (Thermo Electron Corporation, San Jose, CA). Chromatographs were adjusted to display mass ranges corresponding to the theoretical mass range of TAN/BMS when C9 is methylated (414.0951 Da), or when the methyl group on C9 is absent (400.0794).

Data analysis

t-tests were performed using GraphPad Prism v7.4 for Windows.

2.10.5 Intracellular Proteomic analysis

Culturing mycelia

Spores were cultured as per the directions described in *Shaking cultures for protoplastation* and *Transfer cultures for MS based proteomics* (**Section 2.4**). Mycelia was frozen in liquid nitrogen and ground using the mortar and pestle method as described in *Genomic DNA extraction and PCR amplification* (**Section 2.9.1**).

Lysing mycelia

Intracellular protein digestion and analysis was performed as per Ball *et al*, 2019, and Ball and Geddes-McAlister 2019 (Ball *et al*. 2019; Ball and Geddes-McAlister 2019). Approximately 0.5 mL ground mycelia was transferred to a 15 mL tube and combined with 600 µL tris buffer HCl (100 mM Tris-HCl, pH 8.5), as well as a protease inhibitor tablet. The sample was vigorously vortexed and 20% SDS was added to reach a final concentration of 2%. A probe sonicator (Fisherbrand™ Model 120 Sonic Dismembrator, Waltham, MA) was used to disrupt tissue for 15 cycles each lasting 30 seconds.

Samples were transferred to a 2 mL Eppendorf tube and 1:100 1 M DTT was added and vortexed. Samples were incubated at 95 °C for 10 min and at 800 RPM. Samples were chilled on ice and 1:10 volume of 0.55 M IAA was added and vortexed. Samples were incubated in the dark at room temperature for 20 min. 100% acetone was added to each sample to achieve a final concentration of 80%. Samples were incubated overnight at -20 °C.

Trypsin digestion

Samples were centrifuged at 13,500 rpm at 4 °C for 10 min to pellet the protein precipitate. The supernatant was removed and the pellet was washed with 500 µL of 80% chilled acetone. Centrifugation and washing in acetone was repeated. The supernatant was removed and samples were air dried at room temperature. Samples were resuspended in 200 µL of a 8 M Urea and 40 mM HEPES solution. Samples were vortexed vigorously to dissolve the pellet. Protein concentration was measured using a BSA tryptophan assay. For each 50 µg of protein, a 1:50 Trypsin/Lys-C protease mixture was used for digestion. Samples were gently mixed in low sample-to-surface binding 1.5 mL Eppendorf tubes (Eppendorf LoBind microcentrifuge tubes, cat. no. 022431081) and incubated overnight at room temperature.

Stopping digestion

Digestion was ceased using a 0.1 volume of stopping solution (20% acetonitrile, 6% trifluoroacetic acid). Samples were centrifuged for 5 min at 10,000 × g to pellet debris and the supernatant was transferred to a 1.5 mL Eppendorf tube.

Peptide purification using STop And Go Extraction (STAGE) tipping

Prior to purification, STAGE tips were packed with C18 resin (3M Empore, cat. no. 3M2215). Three successive washing cycles were performed on STAGE tips using a Sonation, cat. no. STC-V2 centrifuge: 100 µL 100% acetonitrile (centrifuged 2 min at 1000 × g), 50 µL Buffer B (centrifuged 2 min at 1000 × g), and 200 µL Buffer A (centrifuged 3 min at 1000 × g), where Buffer A consists of 2% acetonitrile, 0.1% trifluoroacetic acid, 0.5% acetic acid, and water, and Buffer B consists of 80% acetonitrile, 0.5% (v/v) acetic acid, and water. A volume equivalent to 50 µg of protein was loaded into the C18 STAGE tip, and centrifuged for 3 – 5 minutes at 1000 × g. C18 STAGE tips were then washed with 200 µL Buffer A, and centrifuged for 3 – 5 minutes at 1000 × g. Peptides were eluted from C18 STAGE tips using 50 µL Buffer B, and centrifuged for 2 min at 1000 × g directly into 0.2 mL PCR tubes. Buffer B was then evaporated from the sample using a vacuum concentrator for 30 – 40 min. Finally, peptides were resuspended in 10 µL Buffer A and pipetted up and down 3 – 5 times.

Protein identification using LC-MS/MS

Analysis of proteins was conducted using a Q Exactive™ Plus Hybrid Quadrupole-Orbitrap™ Mass Spectrometer (Thermo Electron Corporation, San Jose, CA). A volume corresponding to 1.5 µg was injected into a 50-cm Easy-Spray column heated to 50 °C. Peptides were separated over a 60 min gradient using 5% – 60% acetonitrile in 0.5% acetic acid, followed by a washing step over a 10 min gradient using 95% acetonitrile. A flow rate of 300 nL/min was applied over the linear gradient using an EASY-nLC™ 1200 System. Proteins were analyzed in a scan range of 300 to 1650 *m/z* and with a resolution of 60,000 at 100 *m/z*. Chromatographs were visualized using Thermo Xcalibur (Thermo Electron Corporation, San Jose, CA), and data analysis was conducted using Proteome Discoverer™ Software v2.4.

Data analysis

Proteome Discoverer software v2.4 (ThermoFisherScientific, Waltham, MA) was used to compute chromatographs and spectra, and provide analyses of gene relative abundance. One-way ANOVA followed by Dunnett's multiple comparisons test was performed using GraphPad Prism v7.4 for Windows.

2.10.6 NMR

Spores were cultured in a larger assay as per the directions described in *Standing cultures* (**Section 2.4**). 100 mL of supernatant was transferred from each 25 × 100 mm petri dish and loaded into a sterile separatory funnel along with 20 mL ethyl acetate. The separatory funnel was closed and shaken vigorously to extract metabolites into the organic phase. Once settled, the organic layer

was decanted into a sterile container and placed at -80 °C for 10 min to separate residual water. Water droplets were removed using a pipette. The organic layer was evaporated under airflow using nitrogen gas. The dried samples were prepared by dissolution in DMSO-d6.

^{1}H and ^{13}C NMR spectra were recorded on a Varian VNMRS-500 MHz NMR (Varian Inc., Palo Alto, CA) equipped with a 5 mm AutoX DB (Dual Broadband) probe ^{1}H-$^{19}F/X[^{15}N$-$^{31}P]$, with z-PFG and automatic tuning for all nuclei by the ProTune accessory. ^{1}H NMR spectra were obtained by running 8 scans with a relaxation delay of 1 second, and ^{13}C NMR spectra were obtained by running 4096 scans with a relaxation delay of 1 second. The system operates with VNMRJ 3.2 software under LINUX Red Hat 5.

3. Results

3.1 Phylogenetic analysis of methyltransferases in *A. niger*

Methyltransferases encoded in NRRL3

To test the hypothesis that different methyltransferases can perform alternative methyl patterning using TAN/BMS as a substrate, the construction of a library of intraspecies NRRL3 methyltransferases was pursued. I began by analyzing the total number of methyltransferases present in the NRRL3 genome and organizing them into a phylogenetic tree to examine trends. The NRRL3 genome consists of 11,846 genes (https://genome.fungalgenomics.ca/), with 147 genes encoding methyltransferases. Among these, 34 genes encode *O*-methyltransferases involved in secondary metabolite modifications (https://genome.fungalgenomics.ca/). Epigenetic methyltransferases that modify histones, tRNA, RNA, DNA and N-terminals of DNA represent 41 methyltransferase genes in NRRL3, and secondary metabolite methyltransferases such as *N*-methyltransferases are encoded by one gene (https://genome.fungalgenomics.ca/). Other methyltransferases (*e.g.,* methyltransferase domain-containing proteins and methyltransferase-like proteins) represent 71 methyltransferase genes in NRRL3 (https://genome.fungalgenomics.ca/) (**Table 10**).

Table 10: Methyltransferases in *A. niger* NRRL3

Methyltransferases in NRRL3 based on protein family domain function	n
Methyltransferase	**147**
• **Epigenetic methyltransferases**	**41**
Histone methyltransferase	15
tRNA methyltransferase	11
RNA methyltransferase	7
DNA methyltransferase	4
N-terminal methyltransferase	4
• **Secondary metabolite methyltransferases**	**35**
O-methyltransferase	34
N-methyltransferase	1
• **Other methyltransferases**	**71**

Biosynthetic gene clusters encoding methyltransferases

I then identified methyltransferases that were encoded in biosynthetic gene clusters and involved in the production of a secondary metabolite. In the NRRL3 genome, 597 genes are organized into 39 different biosynthetic gene clusters that are involved in secondary metabolite production. Among these, 20 methyltransferase genes are encoded in 17 biosynthetic gene clusters (**Table 11**).

Next, a bioinformatic search using the NRRL3 database (https://genome.fungalgenomics.ca/) and the Protein Family Database (J. Mistry *et al*., 2021) revealed six principle protein families that corresponded to methyltransferase function. Specifically, *O*-methyltransferase function was encoded by the PF00891 and PF13847 protein domains (Kotanin cluster, Pyranonigrin E cluster,

NRRL3_06189 cluster, NRRL3_08341 cluster, NRRL3_08388 cluster, NRRL3_08969 cluster, and NRRL3_10148 cluster), while other types of methyltransferase function were encoded by PF04191, PF08242, PF13489, PF13649, and PF13847 (**Table 11**).

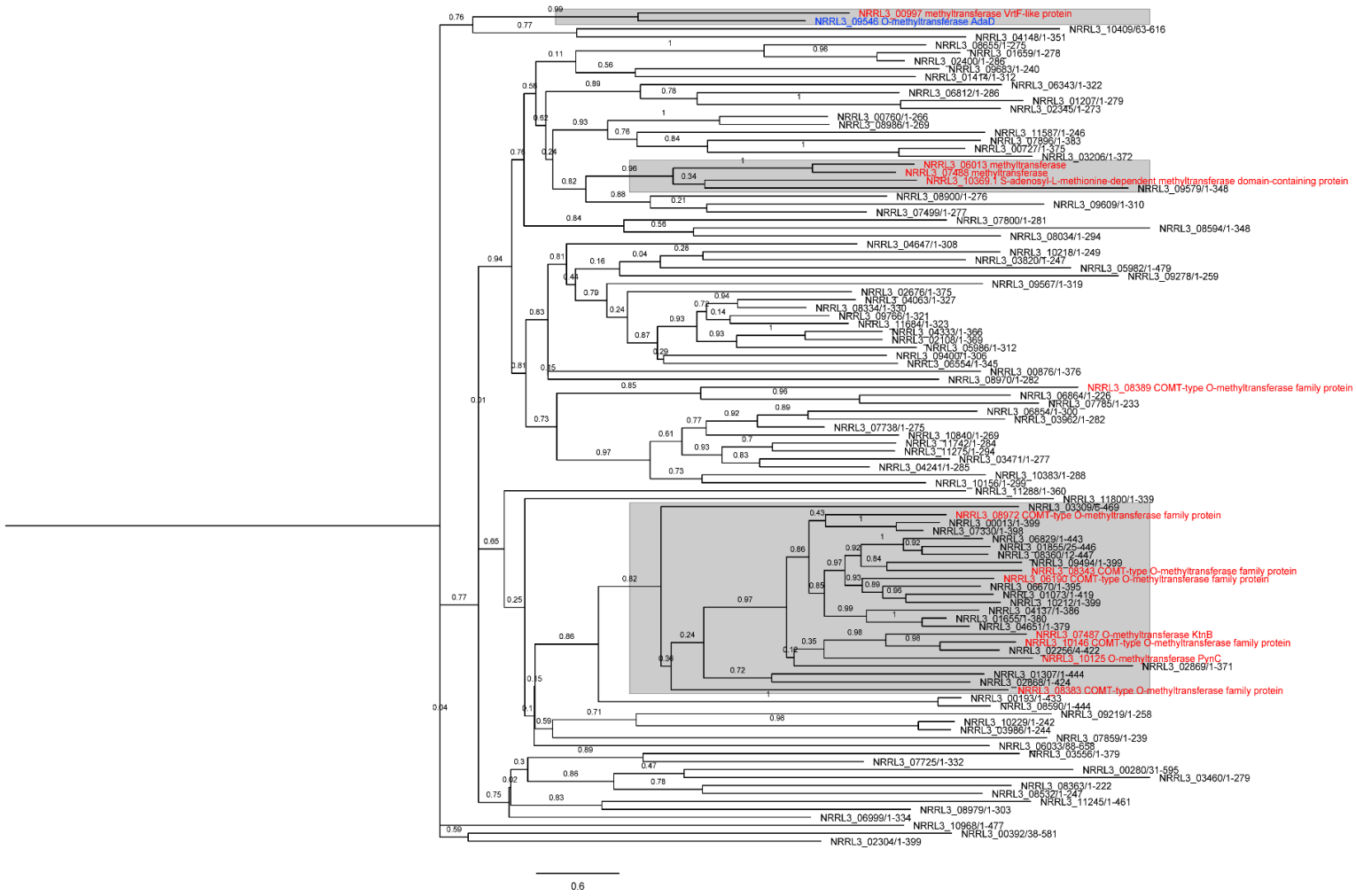
Table 11: Biosynthetic gene clusters in *A. niger* NRRL3

NRRL3 biosynthetic gene cluster ID	Protein family domain ID	Gene function	Reference
Kotanin cluster Pyranonigrin E cluster NRRL3_06189 cluster NRRL3_08341 cluster NRRL3_08388 cluster NRRL3_08969 cluster NRRL3_10148 cluster	PF00891	*O*-methyltransferase	(Keller *et al*. 1993)
Siderophore cluster	PF04191	Phospholipid methyltransferase	(Kodaki and Yamashita 1987)
NRRL3_00998 cluster	PF08242		
NRRL3_01418 cluster NRRL3_04226 cluster	PF13489		
NRRL3_06340/NRRL3_06341 cluster NRRL3_08969 cluster NRRL3_08978/NRRL3_08980/ NRRL3_08984.1 cluster NRRL3_10375 cluster	PF13649	Methyltransferase domain	
TAN-1612/BMS-192548 cluster		*O*-methyltransferase	
NRRL3_07739 cluster	PF13847	Methyltransferase domain	

Phylogenetic tree building

An intra-species phylogenetic tree was constructed based on methyltransferase gene sequence similarity and protein family domains involved in secondary metabolite production (Song *et al.,* 2018). This included the 105 NRRL3 methyltransferases that were identified in **Table 10**, however epigenetic and *N*-methyltransferases were excluded because they do not modify secondary metabolites, or are not involved in the modification of -OH groups on TAN/BMS.

The phylogenetic tree depicted in **Figure 3** demonstrates the correlation between the 105 intraspecies NRRL3 methyltransferase genes. A pairwise comparison matrix (presented in **Appendix 1**) was performed in parallel in order to identify clustering in the phylogenetic tree using threshold values of 20% protein sequence similarity between individual genes. In addition, threshold values of 0.70 – 0.95 were used to assess bootstrapping in clusters. Using these criteria, three clusters are apparent in the intraspecies methyltransferase phylogenetic tree.



Legend: = NRRL3 09546 (*O*-methyltransferase gene) native; = *O*-methyltransferase mutant library genes

Figure 3: **Intraspecies NRRL3 *O*-methyltransferase phylogenetic tree**. The correlation between 105 genes is reported among methyltransferase and *O*-methyltransferase domains.

Methyltransferase library selection and assembly in pjetC

Using the phylogenetic tree as a tool for identifying candidate methyltransferases, I selected a variety of methyltransferases based on the diversity of clusters present in the phylogenetic tree, in which clustering demonstrated the relationship between intraspecies NRRL3 methyltransferases. In addition, bootstrap values confirmed the likeliness of the algorithm used to construct the phylogenetic tree.

I selected a total of 13 methyltransferases based on the inclusion criteria (1) methyltransferases are allocated in a biosynthetic gene cluster involved in secondary metabolite production; (2) methyltransferase function is encoded by a relevant protein family; (3) genes are described in the literature.

The 13 methyltransferase gene candidates selected for gene replacement are listed in **Table 12** and are shown in red or blue in **Figure 3**. While most genes fulfil the selection criteria, NRRL3_06013 and NRRL3_07488 are not encoded in a biosynthetic gene cluster but match other criteria, and a number of genes encode methyltransferase domains rather than *O*-methyltransferases. Furthermore, NRRL3_07487 and NRRL3_10125 represent genes from the well-described Kotanin and Pyranonigrin E clusters, respectively.

Table 12: NRRL3 intraspecies methyltransferase library genes

NRRL3 gene ID	Biosynthetic gene cluster ID	Protein family domain ID	NRRL3 gene function
NRRL3_09546	TAN-1612/BMS-192548 cluster	PF13847	Methyltransferase domain
NRRL3_06190	NRRL3_06189 cluster	PF00891	*O*-methyltransferase
NRRL3_07487	Kotanin cluster		
NRRL3_08343	NRRL3_08341 cluster		
NRRL3_08383	NRRL3_08388 cluster		
NRRL3_08389	NRRL3_08388 cluster		
NRRL3_08972	NRRL3_08969 cluster		
NRRL3_10125	Pyranonigrin E cluster		
NRRL3_10146	NRRL3_10148 cluster		
NRRL3_00997	NRRL3_00998 cluster	PF08242	Methyltransferase domain
NRRL3_06013	N/A	PF13649	
NRRL3_07488	N/A		
NRRL3_10369.1	NRRL3_10375 cluster		

A plasmid library was designed and constructed in the *E.coli* plasmid pjetC. As described in **Table 5**, the genetic background of the strain used to perform transformations includes the deletion of *pyr*G and *kus*A, resulting in NRRL3_09545^{OE}09546^{KO}*pyr*G- (herein referred to as NRRL3_09545^{OE}09546^{KO}). To prepare the gene library for transformation, plasmids harboring methyltransferase genes were prepared as linear fragments using PCR. Sequences corresponding to the *gla*A promoter and terminator flank methyltransferase genes, and were amplified to later

use as the template for homologous recombination during transformation. Methyltransferase library genes are depicted in **Figure 4**.

Overall, all PCR products were consistent with the expected sizes of the methyltransferase genes. The lengths of each methyltransferase gene are as follows, and include an 820 bp sequence corresponding to the *gla*A promoter and terminator region of the pjetC plasmid: NRRL3_00997, 1592 bp; NRRL3_06013, 1966 bp; NRRL3_06190, 2258 bp; NRRL3_07487, 2310 bp; NRRL3_07488, 1865 bp; NRRL3_08343, 2286 bp; NRRL3_08383, 2282 bp; NRRL3_08389, 2573 bp; NRRL3_08972, 2196 bp; NRRL3_09546, 1537 bp; NRRL3_10125, 1931 bp; NRRL3_10146, 2365 bp; and NRRL3_10369.1, 1896 bp.

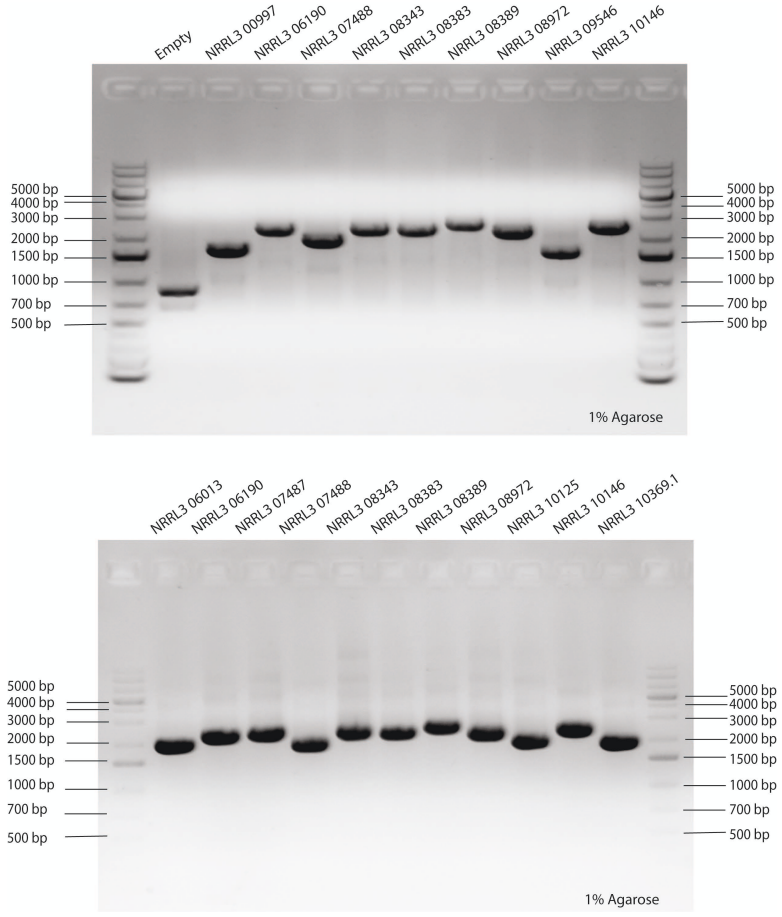


Figure 4: **Methyltransferase gene library**. Methyltransferase genes were amplified by PCR from the pjetC methyltransferase library as linear repair template fragments for transformation.

3.2 Generation and characterization of TAN/BMS overexpression and knockout

Generating a TAN/BMS gene knockout strain

As described in **Figure 5,** the overexpression of the TAN/BMS biosynthetic gene cluster was pursued through the insertion of the strong and inducible promoter gene, α-glucosidase, which was inserted upstream of the promoter region of the TAN/BMS transcription factor, NRRL3_09545.

Overexpression of the clustered genes resulted in the strain NRRL3_09545^{OE}, which exhibited a distinct phenotype including the production of a yellow pigment (**Figure 5b**). Knockout of the TAN/BMS *O*-methyltransferase gene was pursued using the CRISPR-Cas9 system, in which the ANEp8 plasmid harboring a guide specific to *gla*A underwent double-stranded breaks and was repaired using a NRRL3_09546 template that was designed outside of the gene START/STOP region. The knockout resulted in NRRL3_09545^{OE}09546^{KO}, which produced a phenotype similar in appearance to NRRL3_09545^{OE} (**Figure 5c**).

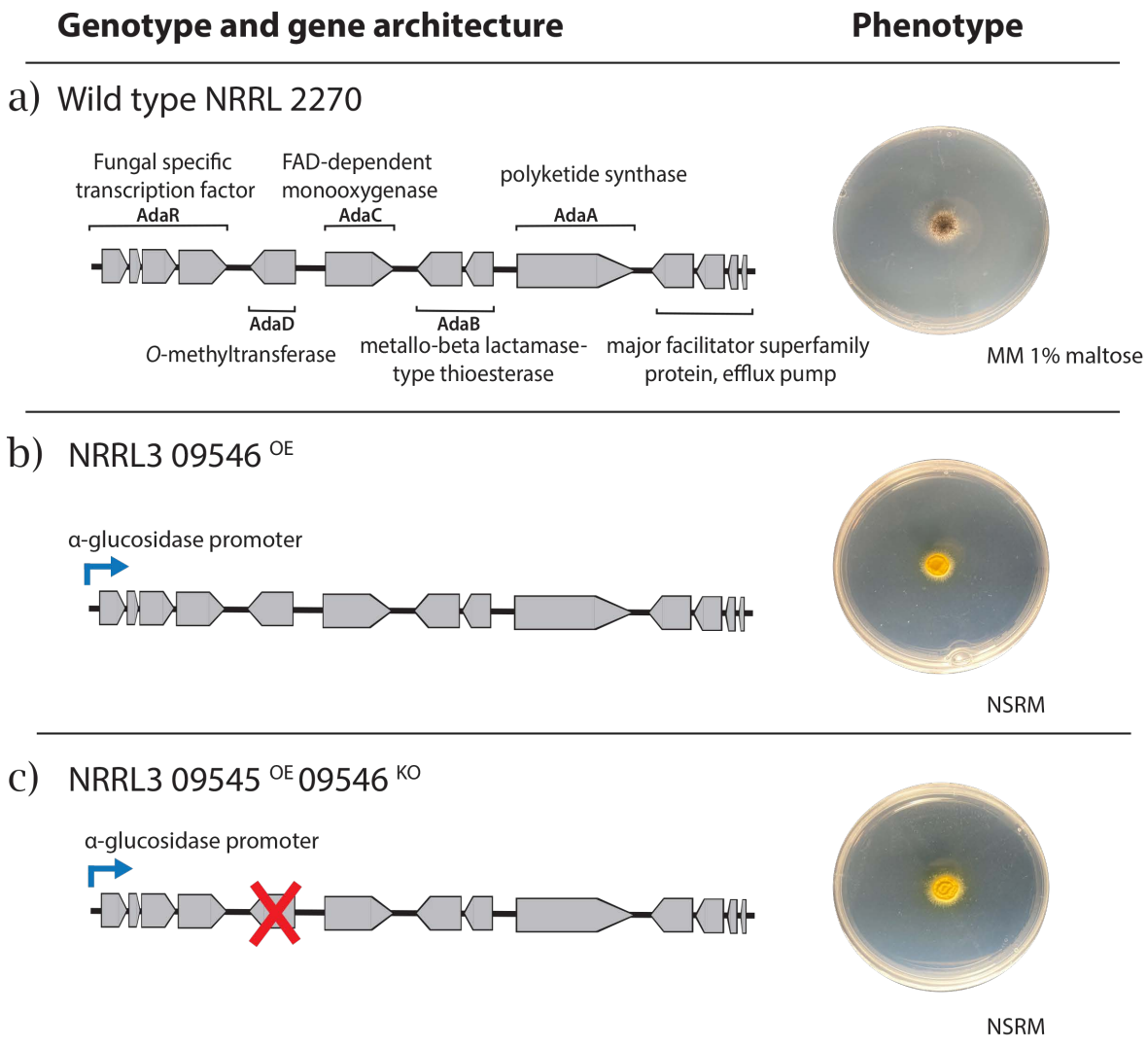


Figure 5: **TAN/BMS gene-edited strains**. **a)** Wild type TAN/BMS biosynthetic gene cluster. **b)** Overexpression of the TAN/BMS biosynthetic gene cluster was prepared by inserting the strong and inducible promoter gene α-glucosidase (depicted with a blue arrow) upstream of the TAN/BMS transcription factor gene NRRL3_09545. **c)** Knockout of the TAN/BMS *O*-methyltransferase NRRL3_09546 was subsequently pursued using the CRISPR-Cas9 system (gene knockout depicted with a red cross).

Considering the molecular mass of TAN/BMS is 415.1009 Da, the expected mass of TAN/BMS produced in the NRRL3_09545^{OE}09546^{KO} strain was calculated to be 401.0852 Da, in which methylation at C9 does not occur, and the mass will be equal to TAN/BMS minus a methyl group. The metabolomic profiles of NRRL3_09545^{OE} and NRRL3_09545^{OE}09546^{KO} were analyzed using mass spectrometry, and mass chromatograms and spectra were validated to compare literature *m/z* profiles corresponding to the gene-edited strains. In doing so, strains were grown for 5 days in inducible conditions and secreted metabolites were extracted and analyzed (**Figure 6**).

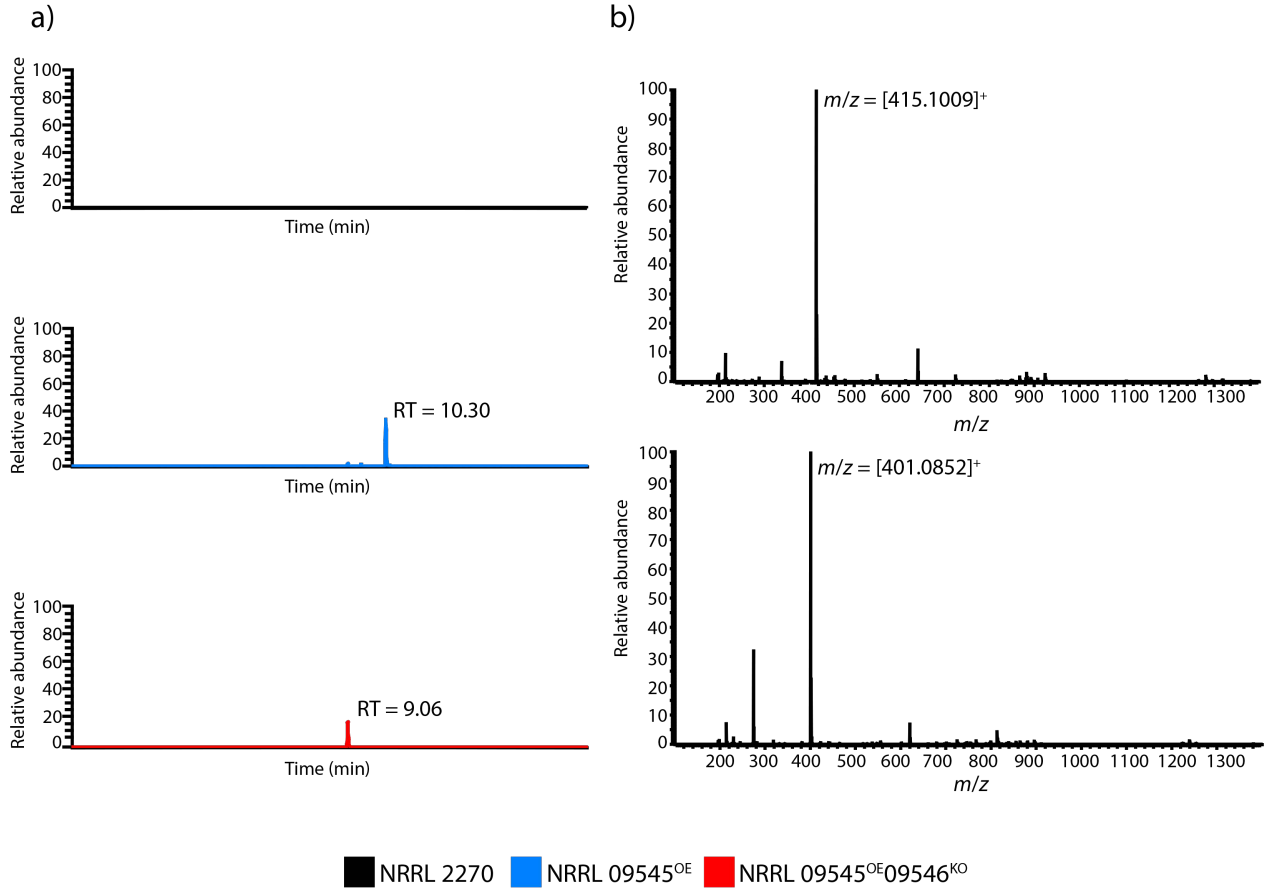


Figure 6: **Mass chromatograms of TAN/BMS**. **a)** Extracted-ion chromatograms displaying the retention times at which TAN/BMS and TAN/BMS analogues are present. **b)** Mass spectra displaying the *m/z* corresponding to TAN/BMS (methylated at C9 above) and TAN/BMS analogues (not methylated at C9 below).

Extracted ion chromatograms that demonstrate the retention times at which NRRL3_09545^{OE} (10.30 min) and NRRL3_09545^{OE}09546^{KO} (9.06 min) appear are depicted in **Figure 6a**. Mass spectra and the *m/z* corresponding to TAN/BMS and the TAN/BMS analogue are depicted in **Figure 6b**. It is evident that the *m/z* of TAN/BMS in NRRL3_09545^{OE} is equivalent to the literature value reported for TAN/BMS (415.1009 Da), and the *m/z* for NRRL3_09545^{OE}09546^{KO} is equivalent to the mass of TAN/BMS minus a methyl group (401.0852 Da). No peaks are present in the wild type strain. The areas under the curve (AUCs) for each peak are as follows (given in arbitrary units): NRRL3_09545^{OE} = 6.20×10^6 U, and NRRL3_09545^{OE}09546^{KO} = 4.25×10^6 U.

Characterization of the methyltransferase knockout strain

The molecular validation of the deletion of NRRL3_09546, was tested by PCR. Band sizes corresponding to the overexpression of the TAN/BMS cluster (*i.e.,* NRRL3_09545^{OE}) and deletion of NRRL3_09546 in NRRL3_09545^{OE}09546^{KO} are depicted in **Figure 7**. Expected and observed band sizes are presented in **Table 13**, in which the NRRL3_09545^{OE} strain amplifies a region of DNA that is 1450 bp in size, while the NRRL3_09545^{OE}09546^{KO} amplifies a region that is 700 bp.

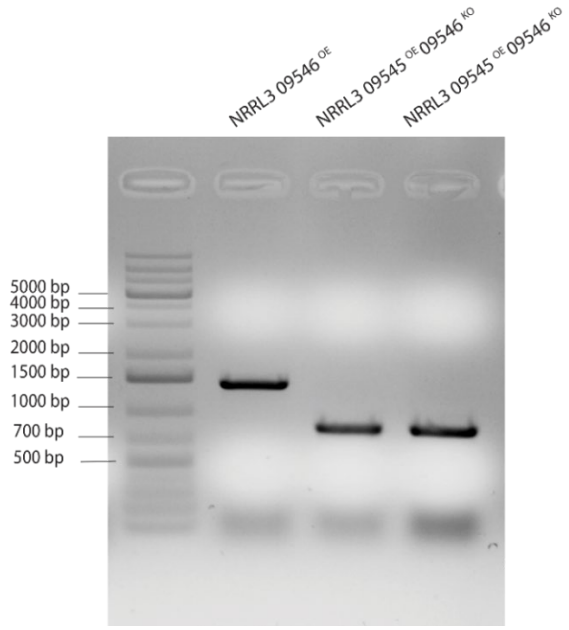


Figure 7: ***O*-methyltransferase knockout strains**. PCR amplification of gene regions corresponding to NRRL3_09545^{OE} and NRRL3_09545^{OE}09546^{KO}.

I observed band sizes in two biological replicates of the NRRL3_09545^{OE}09546^{KO} strain that depicted a slightly larger construct than 700 bp, suggesting the partial deletion of NRRL3_09546 (**Figure 7**). As such, I analyzed NRRL3_09545^{OE}09546^{KO} for complete or partial knockout using Sanger sequencing.

Table 13: *O*-methyltransferase parental knockout strain observed gene length

Sequence	Expectation (size of amplicon)	Outcome
NRRL3_09545 ^{OE}	• 1450 bp	• 1450 bp product on gel
NRRL3_09545^{OE}09546^{KO} Sanger sequencing results	• 700 bp	• 816 bp product on gel • 116 bp region was identified and confirmed by Sanger sequencing

NRRL3_09545^{OE}09546^{KO} was used as a matrix for Sanger sequencing, in which NRRL3_09546 forward and reverse primers were used to amplify the 5' and 3' ends of each gene segment. Using alignment software, I manually aligned the DNA matrix to the NRRL3_09545^{OE} strain in order to visualize the DNA fragment. I identified a region of DNA in the NRRL3_09545^{OE}09546^{KO} that indicated a partial knockout had taken place. However, examination of the NRRL3_09546 coding sequence confirmed that NRRL3_09545^{OE}09546^{KO} was indeed a functional knockout.

As shown in **Figure 8**, the overlap between the NRRL3_09545^{OE} and the reverse primer used to amplify NRRL3_09545^{OE}09546^{KO} indicated the presence of a 116 bp sequence that was not efficiently targeted by the CRISPR-Cas9 plasmid used to delete NRRL3_09546 (**Table 13**).

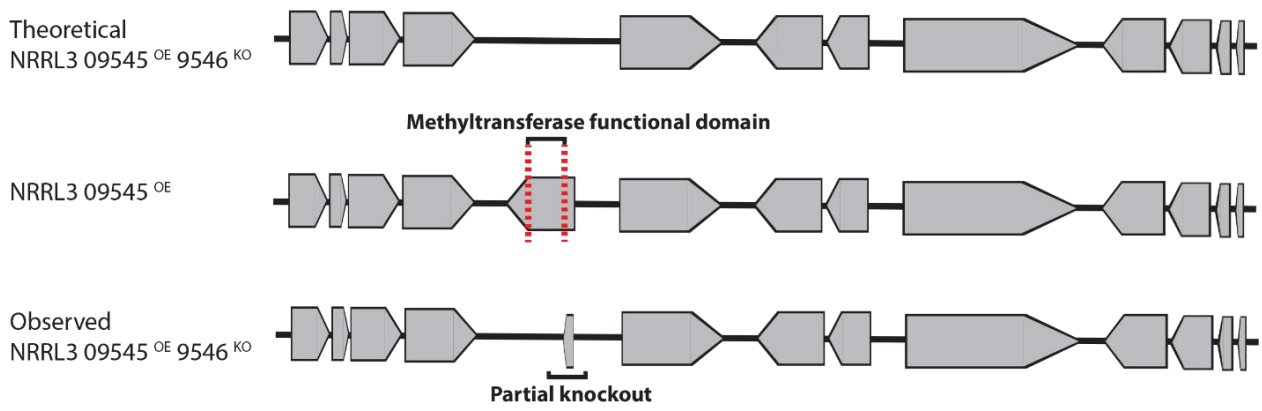


Figure 8: *O*-methyltransferase knockout analysis. *O*-methyltransferase partial knockout with deletion inside the functional domain of the methyltransferase gene. Compared to the theoretical knockout, the observed knockout maintained part of the NRRL3_09546 *O*-methyltransferase gene, as identified in the reverse amplicon.

3.3 NMR analysis of TAN/BMS wild type strains

To identify the molecular structure and methylation patterning of TAN/BMS in NRRL3_09545^{OE} and NRRL3_09545^{OE}09546^{KO}, I tested these gene-edited strains by ^{1}H NMR (^{13}C NMR shown in **Appendix 2**). Spores were grown in a 5-day culture of inducible media and were subsequently extracted using ethyl acetate, dried, and prepared for NMR by dissolution in DMSO-d6.

The spectra corresponding to each strain, including NRRL3_09545^{OE} (TAN/BMS with methylation at C9) and NRRL3_09545^{OE}09546^{KO} (TAN/BMS with no methylation at C9) are shown in **Figure 9**, and confirm the production of each respective product. NMR samples were not purified prior to their examination and, as such, other enzymes and organic acids are present in the NMR spectra.

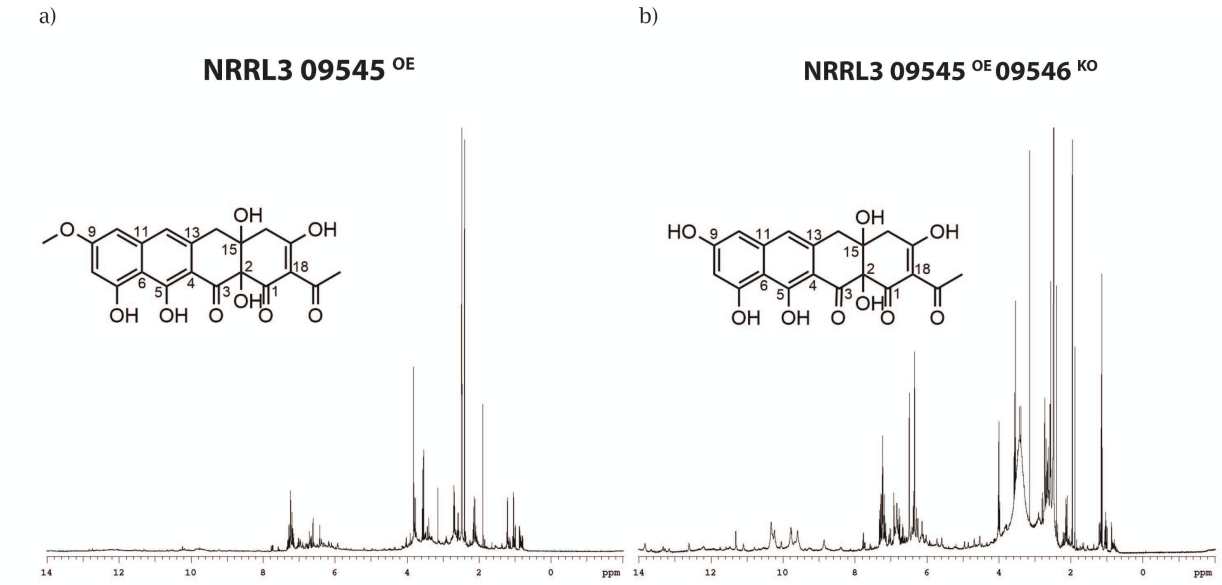


Figure 9: ^{1}H NMR analysis of TAN/BMS strains. **a)** NRRL3_09545^{OE} produced the secondary metabolite TAN/BMS, whereas **b)** NRRL3_09545^{OE}09546^{KO} produced of a modified TAN/BMS analog that does not have a methyl group at C9.

3.4 Protoplast preparation

Following the examination of TAN/BMS production in the gene-edited strain, I optimized the preparation of *A. niger* protoplasts. Protoplasts were harvested from the strain NRRL3_09545^{OE}09546^{KO} through the enzymatic digestion of mycelia, enabling the release of protoplasts over a 2.5 hour period. Protoplast digestion was monitored at 30 min – 1 hour intervals in order to assess the progress of protoplastation. Following digestion, protoplasts were purified by separating mycelia and cell debris.

As depicted in **Figure 10a-d**, the digestion of mycelia results in the release of a large quantity of protoplasts. To estimate the number of protoplasts, a 0.1000 mm deep hemocytometer was used and the protoplast solution was concentrated in the range 10^7 – 10^8 protoplasts/mL. This step was also used to visualize the quality of protoplasts, which should ideally appear as hollow circles that transmit light. As depicted in **Figure 10e**, a mixture of light and dark protoplasts are apparent, in which the latter may be less competent for transformation, owing to incomplete or excessive digestion.

The competency of protoplasts was tested by plating protoplasts on NSRM media, and gene auxotrophy was tested by plating protoplasts on SRM. Transformation with an empty version of the ANEp8 plasmid harboring *pyr*G, but no Cas9-sgRNA, was used as an additional control to observe protoplast competency and *pyr*G prototrophy. Serial dilutions of protoplasts were also prepared to determine the accuracy of the hemocytometer and to later establish transformation efficiency in the *A. niger* co-transformation.

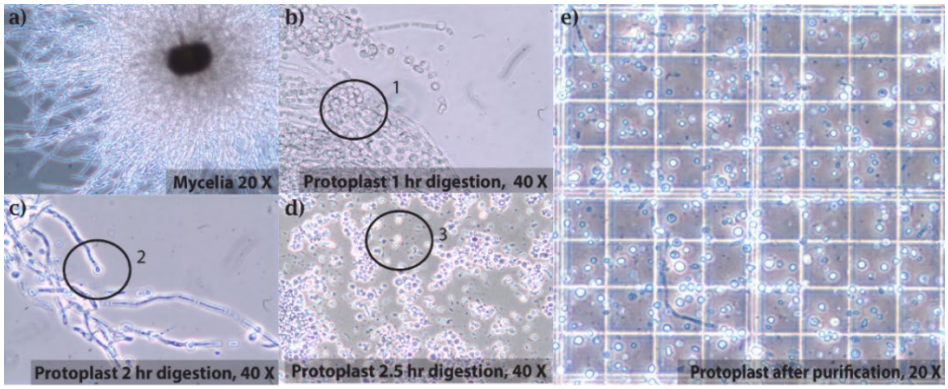


Figure 10: **Protoplastation**. Mycelia digestion occurred over an interval of 2.5 hours and was followed by the purification of protoplasts. **a)** mycelia were grown in an overnight culture and examined prior to enzymatic digestion; **b)** following 1 hour of digestion, protoplasts began to emerge through the hydrolysis of the cell wall and were found in masses trapped within the structure of mycelia (indicated in circle 1); **c)** following 2 hours of digestion, protoplasts emerged from the tips of hyphae (indicated in circle 2) and were abundantly found throughout the digestion solution; **d)** following 2.5 hours of digestion, nearly all mycelia were digested and protoplasts were released (indicated in circle 3); **e)** purification of the protoplast solution separated most mycelia fragments and cell debris, and protoplasts were counted in a hemocytometer in order to concentrate the solution.

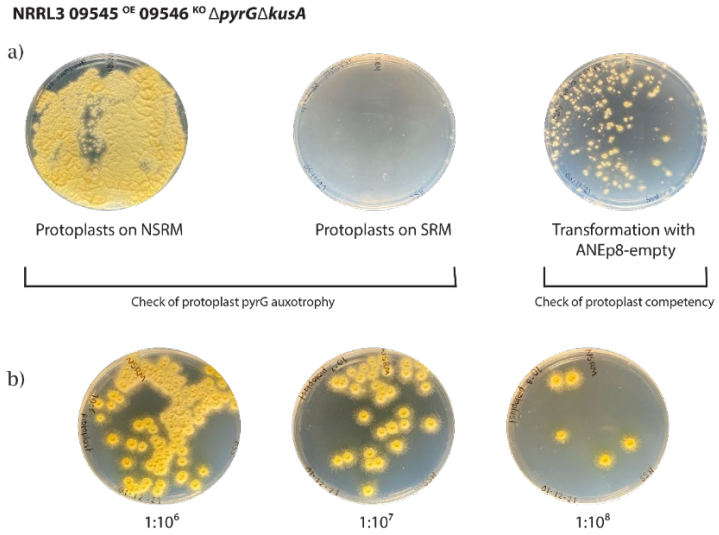


Figure 11: **TAN/BMS strains and controls**. The status of *pyr*G auxotrophy and restoration of *pyr*G gene ejection, as well as the dilution of protoplasts was tested. **a)** Protoplasts were grown on media with uracil (NSMR) and media without uracil (SRM) to test the ejection of *pyr*G and ensure that controls were correctly established for later transformation. Transformation with ANEp8-empty was performed to observe the competence of protoplasts; **b)** In addition to counting protoplasts in a 0.1000 mm deep hemocytometer, protoplasts were also serially diluted and plated to estimate the number of protoplasts.

3.5 *A. niger* co-transformation

Co-transformation was performed using *A. niger* protoplasts (*i.e.,* NRRL3_09545^{OE}09546^{KO}*pyr*G-), and the following: (1) ANEp8-Cas9-sgRNA(*gla*A)_*pyr*G, and (2) a sequence of DNA amplified from pjetC containing the *gla*A promoter, a methyltransferase gene, and the *gla*A terminator. During transformation, the ANEp8 plasmid harboring Cas9, augmented endonuclease activity and targeted the *gla*A gene, resulting in double-stranded DNA breaks. The PCR fragments harboring the *gla*A promoter and terminator served as a repair template and re-circularized cut DNA while replacing individual methyltransferase genes at *gla*A. As described in **Section 3.4**, ANEp8-empty was used as a control to examine protoplast competence and to observe the restoration of *pyr*G. The transformation strategy used herein, and the concentrations of materials used, are summarized in **Figure 12**.

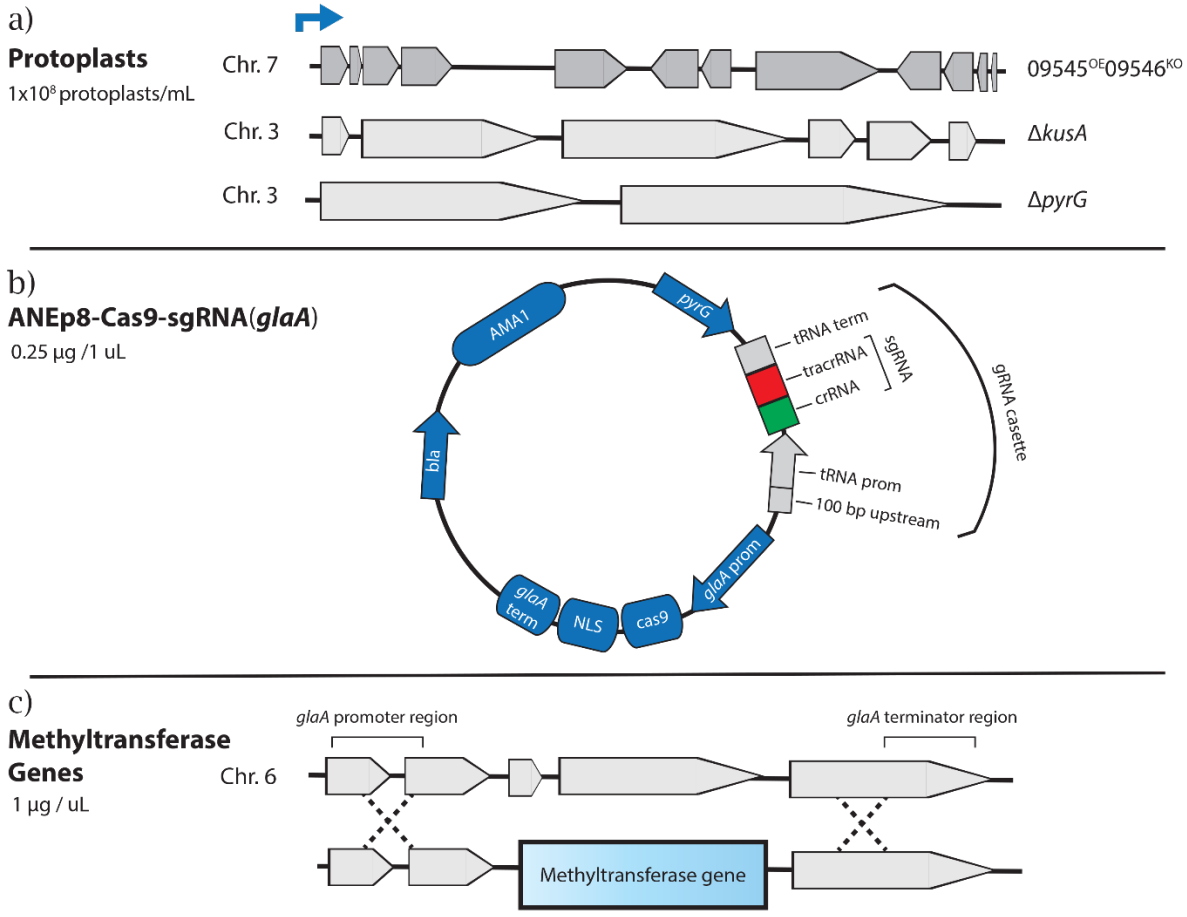
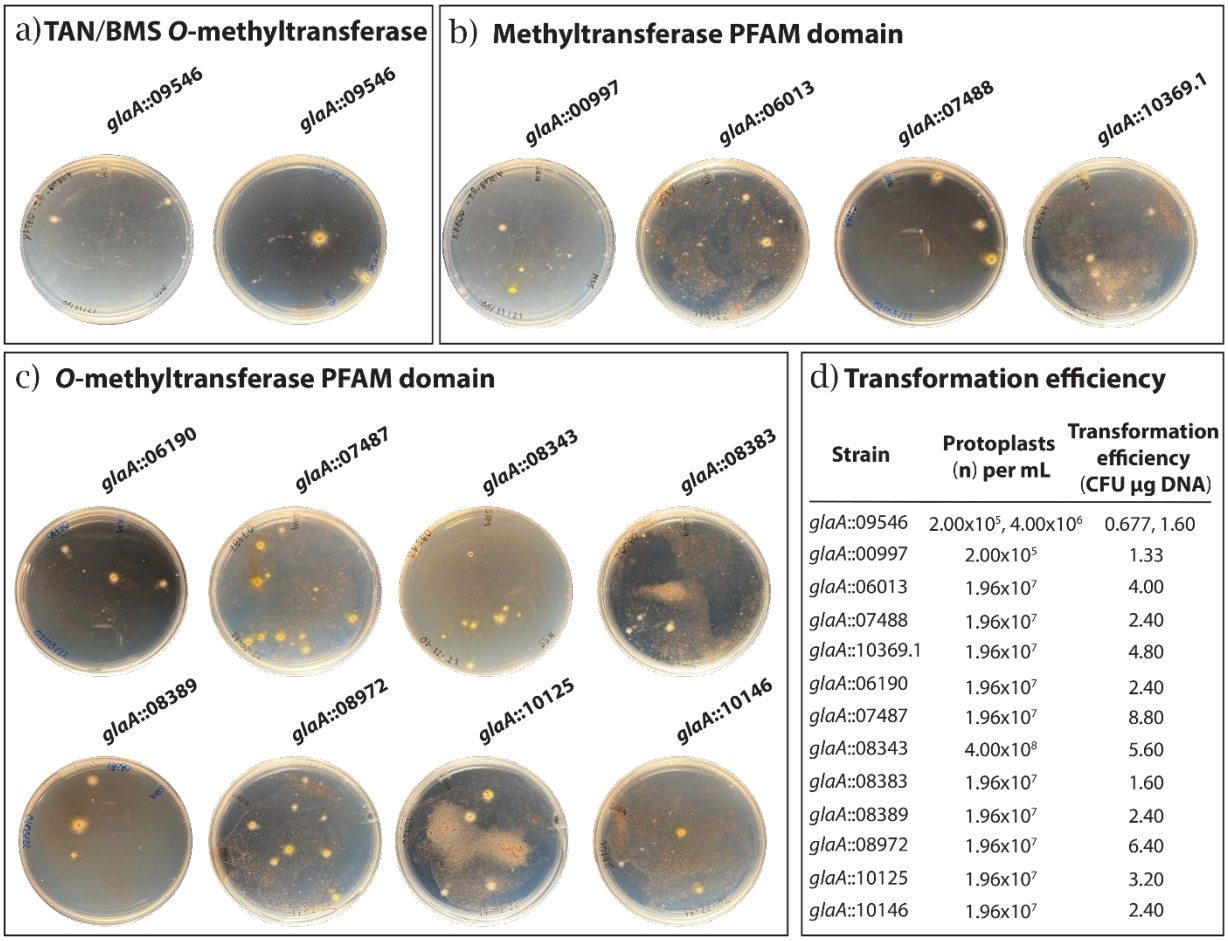


Figure 12: ***A. niger* co-transformation strategy**. **a)** *A. niger* co-transformation was performed using protoplasts that conferred *kus*A and *pyr*G gene deletions. **b)** Co-transformation was also performed using a plasmid that produced double-stranded DNA breaks targeted at the *gla*A gene and which also harbors a gene to restore *pyr*G auxotrophy. **c)** Double-stranded DNA breaks were repaired using a PCR fragment that can undergo homologous recombination with the *gla*A promoter and terminator region, and which augments the insertion of methyltransferase genes. Figure adapted from Song et al., 2018a.

Transformation using ANEp8-emtpy is depicted in **Figure 11**, demonstrating the competence of protoplasts and the restoration of *pyr*G. Transformation using ANEp8-Cas9-sgRNA(*gla*A) is depicted in **Figure 13**. The number of transformants was dependent on the concentration of protoplasts and DNA used in the transformation. On average, 0.25 µg/mL of the ANEp8 plasmid harboring the Cas9 endonuclease and 1 µg/µL of methyltransferase genes were used in each transformation reaction. Overall, the transformation efficiency ranged from 0.677 – 8.80 CFU/µg DNA.



d) **Transformation efficiency**

Strain	Protoplasts (n) per mL	Transformation efficiency (CFU µg DNA)
glaA::09546	$2.00x10^5$, $4.00x10^6$	0.677, 1.60
glaA::00997	$2.00x10^5$	1.33
glaA::06013	$1.96x10^7$	4.00
glaA::07488	$1.96x10^7$	2.40
glaA::10369.1	$1.96x10^7$	4.80
glaA::06190	$1.96x10^7$	2.40
glaA::07487	$1.96x10^7$	8.80
glaA::08343	$4.00x10^8$	5.60
glaA::08383	$1.96x10^7$	1.60
glaA::08389	$1.96x10^7$	2.40
glaA::08972	$1.96x10^7$	6.40
glaA::10125	$1.96x10^7$	3.20
glaA::10146	$1.96x10^7$	2.40

Figure 13: ***A. niger* co-transformation**. Transformation using the *A. niger* strain NRRL3_09545^{OE}09546^{KO}, the ANEp8 plasmid harboring *pyr*G and the Cas9 endonuclease sgRNA targeting *gla*A, as well as methyltransferase genes that served to repair linearized DNA at *gla*A, and included repair with the following methyltransferases: **a)** the TAN/BMS *O*-methyltransferase, **b)** methyltransferases in NRRL3, and **c)** *O*-methyltransferases in NRRL3. **d)** Co-transformation in *A. niger* resulted in a range of transformation efficiencies depending on the concentration of materials used per transformation reaction.

I tested three biological replicates from each transformation plate and purified colonies. Spores from purified colonies were then grown in an overnight culture of complete media, and gDNA was

extracted and purified using PCI. Using primer pairs that amplify sequences corresponding to the *gla*A promoter and terminator — the same primer pairs used to generate the methyltransferase library and served as the region for homologous recombination — gene products were amplified and visualized using gel electrophoresis. As a control, the size of the gene product following transformation using the ANEp8-empty plasmid was also tested. As expected, the size of the gene product was 2960 bp, corresponding to the size of *gla*A without gene editing. As described in **Section 3.1**, the expected length of each gene product, including an individual methyltransferase gene and sequences corresponding to the *gla*A promoter and terminator region (820 bp) are as follows, NRRL3_00997, 1592 bp; NRRL3_06013, 1966 bp; NRRL3_06190, 2258 bp; NRRL3_07487, 2310 bp; NRRL3_07488, 1865 bp; NRRL3_08343, 2286 bp; NRRL3_08383, 2282 bp; NRRL3_08389, 2573 bp; NRRL3_08972, 2196 bp; NRRL3_09546, 1537 bp; NRRL3_10125, 1931 bp; NRRL3_10146, 2365 bp; and NRRL3_10369.1, 1896 bp. All gene products appeared to amplify the correct sequence lengths, indicating successful transformation.

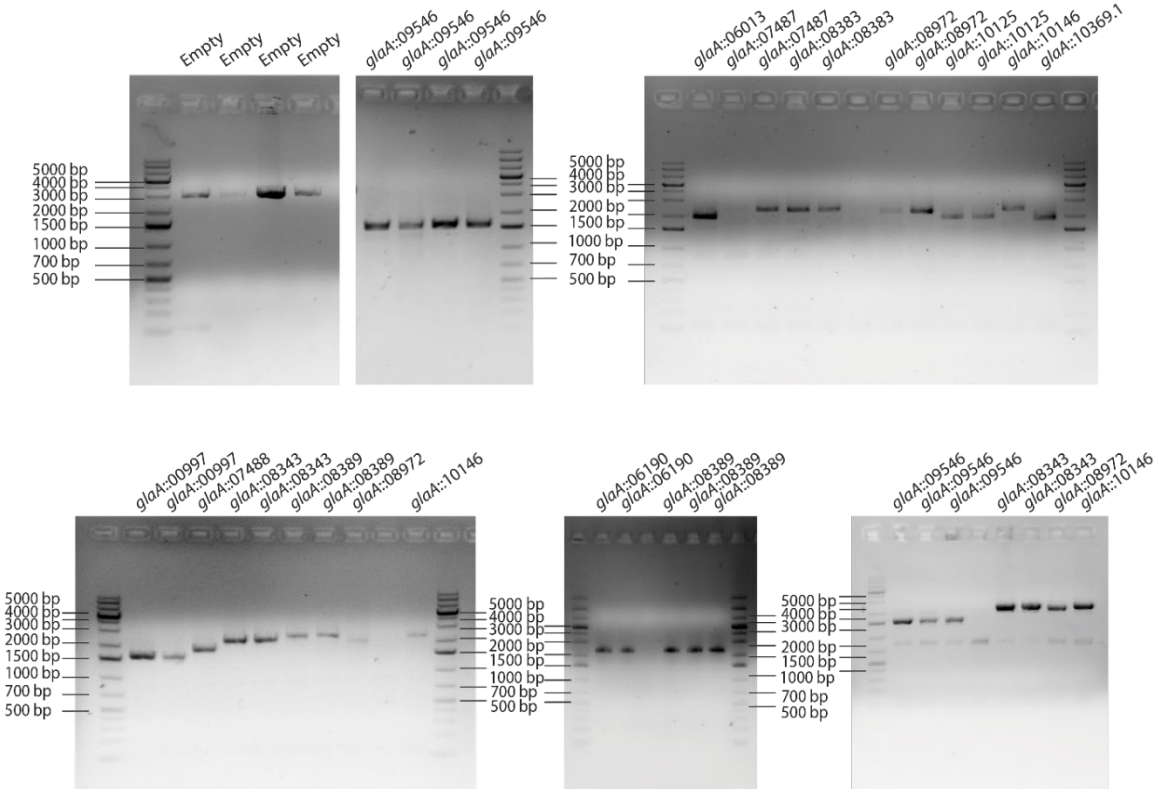


Figure 14: **Methyltransferase gene insertion validation**. Validation of methyltransferase gene insertions at *gla*A. Failed PCRs (*i.e.*, no band) are unlabelled and the transformant was checked in a subsequent reaction.

3.6 Spectrophotometry

I developed a screening method using spectrophotometry to distinguish differences between the absorbance peaks and spectral properties of the wild type strain (n = 3), parental strains NRRL3_09545^{OE} (n = 3) and NRRL3_09545^{OE}09546^{KO} (n = 3), as well as the methyltransferase

mutant NRRL3_09545^{OE}09546^{KO}*gla*A::09546 (n = 9). Considering the same materials used to perform mass spectrometry can be used to examine absorbance, this method represented a quick and low-cost screening method to potentially identify positive transformants. In doing so I prepared a 5-day culture of each strain in inducible media (minimal media 1% maltose) and tested the supernatant in a spectrophotometer.

Differences in absorbance were observed between the strains, and the maximum wavelength value for NRRL3_09545^{OE}, NRRL3_09545^{OE}09546^{KO}, and NRRL3_09545^{OE}09546^{KO}*gla*A::09546 corresponded to the value reported in the literature (*i.e.,* 414 nm) (Shu *et al.* 1995). As shown in **Figure 15**, the wild type strain displayed the lowest absorbance value at 350 nm, and NRRL3_09545^{OE} displayed the highest absorbance maxima at 421 nm. The absorbance maxima of NRRL3_09545^{OE}09546^{KO} was 412 nm, and NRRL3_09545^{OE}09546^{KO}*glaA*::09546 displayed an absorbance maxima of 416 nm.

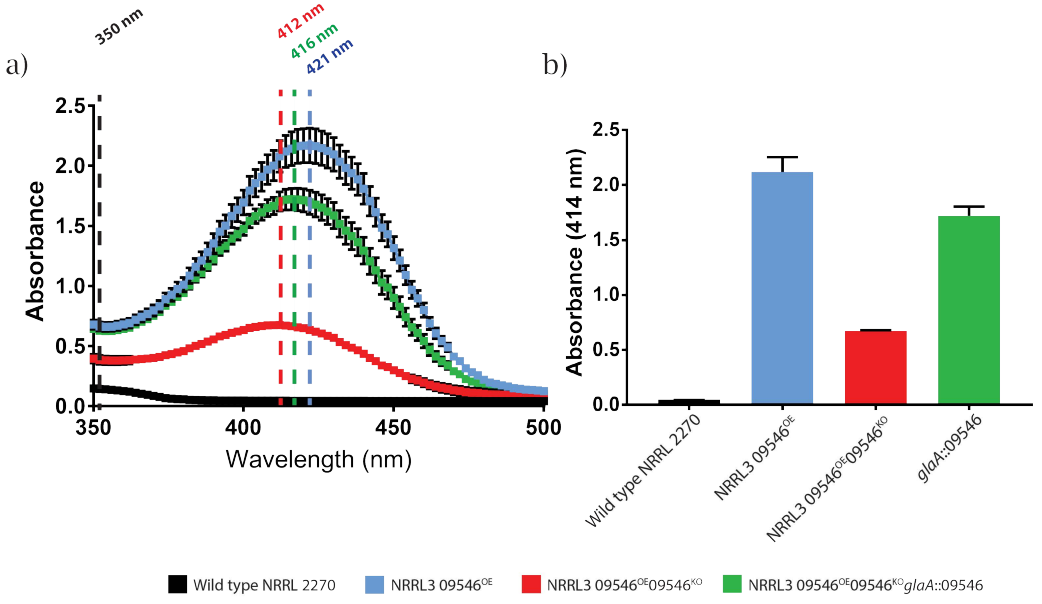


Figure 15: Spectral analysis of TAN/BMS strains. **a)** Wild type, NRRL3_09545^{OE}, NRRL3_09545^{OE}09546^{KO} and NRRL3_09545^{OE}09546^{KO}*gla*A::09546 displayed absorbance maxima at 350 nm, 421 nm, 412 nm, and 416 nm, respectively. **b)** wild-type and gene-edited strains have different absorbance maxima, with NRRL3_09545^{OE} most comparable to NRRL3_09545^{OE}09546^{KO}*gla*A::09546.

3.7 Comparative extracellular metabolomics of TAN/BMS mutant and controls strains

Methylation and relative abundance of wild type, NRRL3_09545^{OE}, NRRL3_09545^{OE}09546^{KO}, and NRRL3_09545^{OE}09546^{KO}glaA::09546

I tested secondary metabolite production using mass spectrometry to compare the expression profiles in the wild type strain, the parental strains NRRL3_09545^{OE} and NRRL3_09545^{OE}09546^{KO}, as well as the NRRL3_09545^{OE}09546^{KO}*gla*A::09546 strain. This was performed prior to testing all *gla*A mutants in order to assess the level of TAN/BMS production in gene-edited strains. As such, mass chromatograms and spectra were generated to confirm the

retention times and *m/z* corresponding to TAN/BMS and modified TAN/BMS analogues, and expression profiles were compared.

The expression profiles comparing the wild type strain, parental controls, and the NRRL3_09545^{OE}09546^{KO}*gla*A::09546 strain are depicted in **Figure 16**, in which extracted-ion chromatograms were adjusted to display the *m/z* of TAN/BMS and TAN/BMS analogues. I observed NRRL3_09545^{OE} at 10.30 min and a smaller peak at 9.05 min, NRRL3_09545^{OE}09546^{KO} at 9.06 min, and NRRL3_09545^{OE}09546^{KO}*gla*A::09546 at both 9.05 and 10.29 min (**Figure 16a**). NRRL3_09545^{OE} and NRRL3_09545^{OE}09546^{KO}*gla*A::09546 appeared to contain peaks corresponding to both the TAN/BMS and non-methylated TAN/BMS expression profiles, potentially indicating that methylation of TAN/BMS was partially restored in the *gla*A::09546 mutant, and that not all TAN/BMS was methylated in NRRL3_09545^{OE}. As previously observed, no peaks were present in the wild type strain.

The spectra displaying the *m/z* of TAN/BMS and the TAN/BMS analogue in which no methylation occurred are reported in **Figure 16b**. The *m/z* of each peak corresponded to the values reported in the literature for TAN/BMS (415.1009 Da) and the TAN/BMS analogue was found to be equivalent to the mass of TAN/BMS minus a methyl group (401.0852 Da).

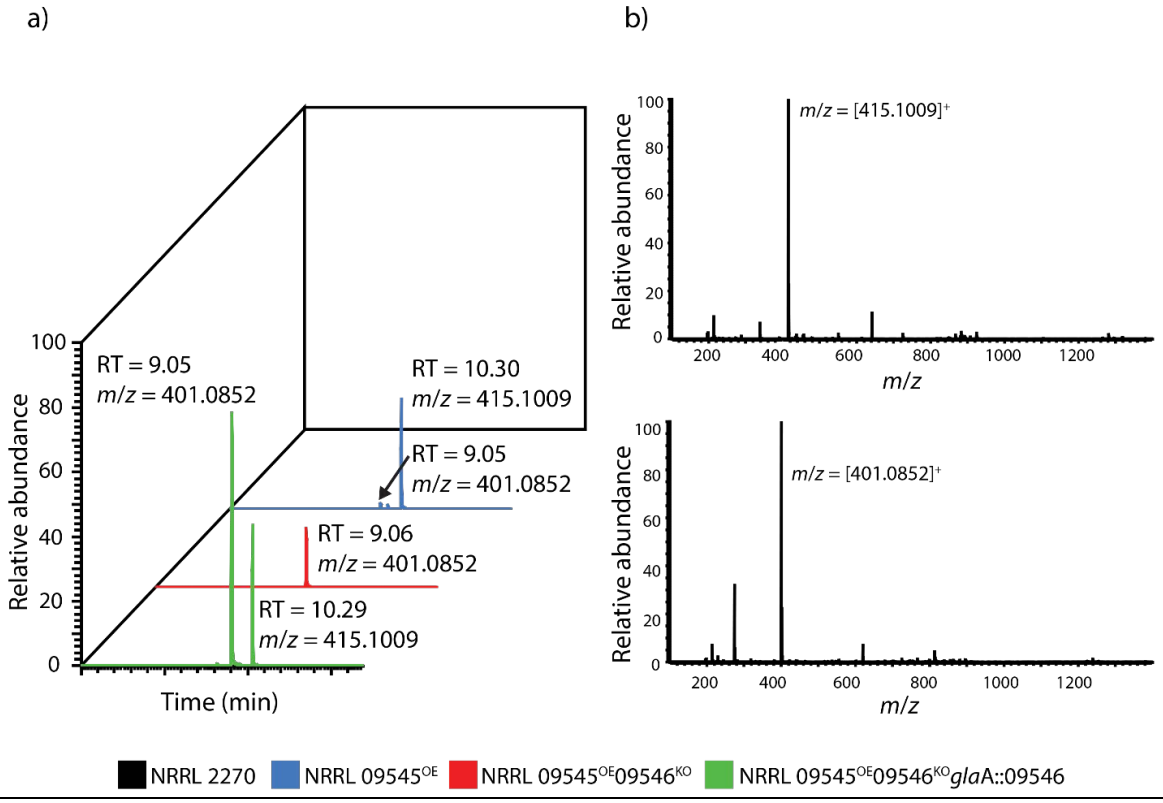


Figure 16: **Mass chromatograms of TAN/BMS and *gla*A edited strains**. **a)** Extracted-ion chromatograms displaying the retention times (RT) at which TAN/BMS and TAN/BMS analogues are present. **b)**: Mass spectra displaying the *m/z* corresponding to TAN/BMS (methylated at C9 above) and TAN/BMS analogues (not methylated at C9 below).

I examined the relative abundance of TAN/BMS production in each strain to predict the extent of methylation exhibited following *gla*A gene replacement. The AUCs for each peak are as follows (given in arbitrary units): NRRL3_09545^{OE} = 6.20 × 10^6 U; NRRL3_09545^{OE}09546^{KO} = 4.25 × 10^6 U; and NRRL3_09545^{OE}09546^{KO} 09546::*gla*A = 2.83 × 10^7 U (9.05 min) and 7.55 × 10^6 U (10.29 min). However, to compare the extent of methylated TAN/BMS to non-methylated TAN/BMS exhibited in NRRL3_09545^{OE}, the relative abundance of each peak was observed in three technical replicates of NRRL3_09545^{OE}.

As depicted in **Figure 17**, the relative abundance of the methylated expression profile was higher than the non-methylated expression profile, ranging from 1.80 × 10^7 U – 7.95 × 10^5 U, respectively. A statistical analysis using an unpaired t-test comparing the relative abundance of methylated TAN/BMS to unmethylated TAN/BMS revealed a significant difference between the two products generated in the strain NRRL3_09545^{OE} ($p < 0.001$). All statistical analyses are presented in **Appendix 3**.

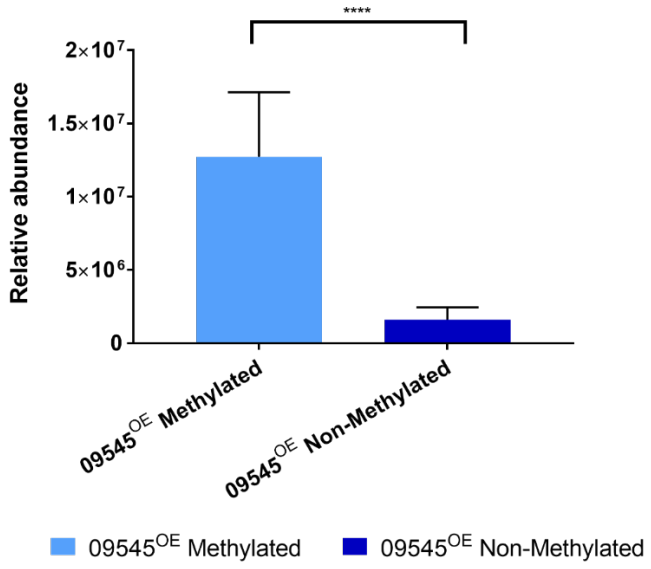


Figure 17: **C9 methylation in the strain NRRL3_09545^{OE}** (n = 3). The proportion of methylated TAN/BMS in the strain NRRL3_09545^{OE}.

In a similar manner, I tested the extent of methylation between TAN/BMS and the non-methylated TAN/BMS analog in the strain NRRL3_09545^{OE}09546^{KO}*gla*A::09546. Considering the relationship among TAN/BMS production in the *gla*A::09546 mutant was observably larger than in NRRL3_09545^{OE} (**Figure 16**), additional NRRL3_09545^{OE}09546^{KO}*gla*A::09546 transformations were prepared. As such, a total of six biological replicates (with three technical replicates per transformed strain) were compared.

As depicted in **Figure 18**, the extent of methylation and absence of methylation in the *gla*A::09546 mutants ranged from 1.68 × 10^7 U – 4.04 × 10^7 U, respectively. A statistical test was performed using a t-test comparing the difference and extent of methylation in each NRRL3_09545^{OE}09546^{KO}*gla*A::09546 replicate. Overall, the difference in methylation in the

strain NRRL3_09545^{OE}09546^{KO}*gla*A::09546 was significant when comparing methylation versus non-methylation, with p-values ranging from $p < 0.05$ to $p < 0.0001$. All statistical analyses are presented in **Appendix 3**.

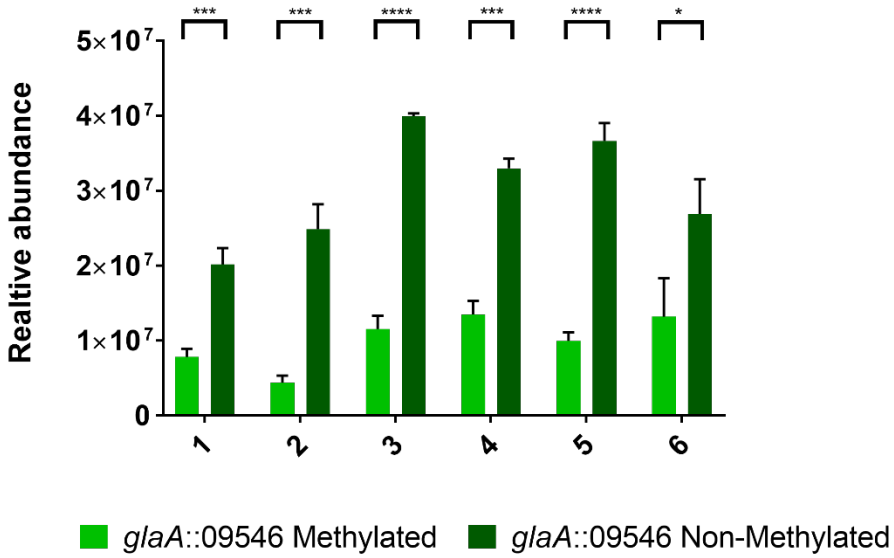


Figure 18: **C9 methylation in the strain NRRL3_09545^{OE}*gla*A::09456** ($n = 6$). The proportion of methylated TAN/BMS produced in the strain NRRL3_09545^{OE}09546^{KO} examined in six replicates.

Methylation and relative abundance of NRRL3_09545^{OE}09546^{KO} and methyltransferase library mutants

I tested secondary metabolite production using mass spectrometry to compare the expression profiles in the NRRL3_09545^{OE}09546^{KO} and *gla*A gene-edited strains, in which the knockout strain was used to the compare methylation status and observe the restoration of methylation in TAN/BMS production. **Figure 19** depicts comparisons between the knockout strain and seven gene-edited strains each comprising three biological replicates, including the following strains: (1) *gla*A::00997, (2) *gla*A::06013, (3) *gla*A::07487, (4) *gla*A::07488, (5) *gla*A::08343, (6) *gla*A::08389, and (7) *gla*A::08972. Extracted-ion chromatograms were adjusted to display the *m/z* of TAN/BMS and TAN/BMS analogues.

While the *gla*A::09546 led to the partial restoration of TAN/BMS, I observed the non-methylated TAN/BMS expression profile in all other *gla*A mutants. As such, all *gla*A mutants were equivalent to the knockout strain, demonstrating retention times of approximately 9.06 min. Spectra displaying the *m/z* of the TAN/BMS analogue is not shown, but is equivalent to the spectra reported in **Figure 16b**, in which the mass of the non-methylated TAN/BMS is equivalent to 401.0852 Da.

The relative abundance between NRRL3_09545^{OE}09546^{KO} and the gene-edited strains were observably different, however the relative abundance between each *gla*A mutant was consistant. The AUCs for each peak are as follows (given in arbitrary units): NRRL3_09545^{OE}09546^{KO} = 4.25

× 10^6 U; *gla*A::00997 = 2.48 × 10^7 U; *gla*A::06013 = 3.31 × 10^7 U; *gla*A::07487 = 3.57 × 10^7 U; *gla*A::07488 = 2.64 × 10^7 U; *gla*A::08343 = 2.75 × 10^7 U; *gla*A::08389 = 1.89 × 10^7 U; and *gla*A::08972 = 2.16 × 10^7 U. The remaining methyltransferase genes (not shown) depicted a similar pattern, with AUC values of *gla*A::06190 = 1.98 × 10^7 U, *gla*A::08383 = 1.76 × 10^7 U, *gla*A::10125 = 3.93 × 10^7 U, *gla*A:: 10146 = 8.85 × 10^6 U, and *gla*A::10369.1 = 3.63 × 10^7 U.

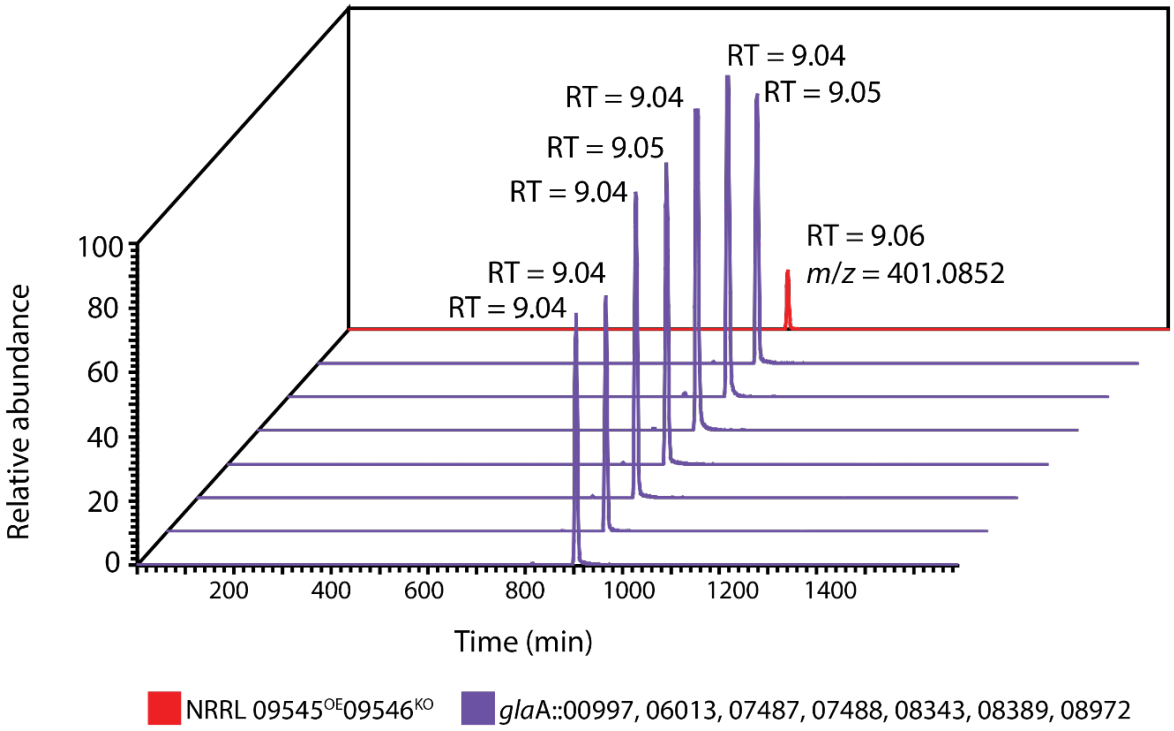


Figure 19: Mass chromatograms of TAN/BMS knockout and methyltransferase gene library edited strains. TAN/BMS production was examined following CRISPR-Cas9 gene editing with methyltransferase mutants. The retention time (RT) was examined corresponding to both the methylated and non-methylated forms of TAN/BMS, however only the presence of the latter was detected.

3.8 Mass-spectrometry based proteomic analysis of intracellular TAN/BMS production

To ensure TAN/BMS synthesis was occurring in the *gla*A mutant strains, including the expression of methyltransferases, I tested the level of protein expression in each gene of the TAN/BMS cluster. As such, protein expression was estimated in the wild type strain, parental strains NRRL3_09545^{OE} and NRRL3_09545^{OE}09546^{KO}, as well as the following gene-edited strains in which genes were transformed into *gla*A: *gla*A::09546, *gla*A::00997, *gla*A::07487, *gla*A::10125, and *gla*A::10369.1. I selected these strains to test a diversity of methyltransferases that were present in each cluster of the phylogenetic tree to estimate the effect on TAN/BMS synthesis and methyltransferase activity (**Figure 3**). The whole proteome was examined using LC-MS/MS, and mass chromatograms and spectra were processed computationally.

Protein expression in the TAN/BMS gene cluster was isolated from the proteome and statistical tests were performed comparing the level of protein expression in the wild type strain to the different gene-edited strains. The intensity of protein expression between genes in the TAN/BMS gene cluster were grouped by treatment type and are presented in **Figure 20a**, with the intensity of protein expression of the polyketide synthase gene NRRL3_09549, and *O*-methyltransferase gene NRRL_09546 presented in a larger view in **Figure 20b** and **Figure 20c**, respectively.

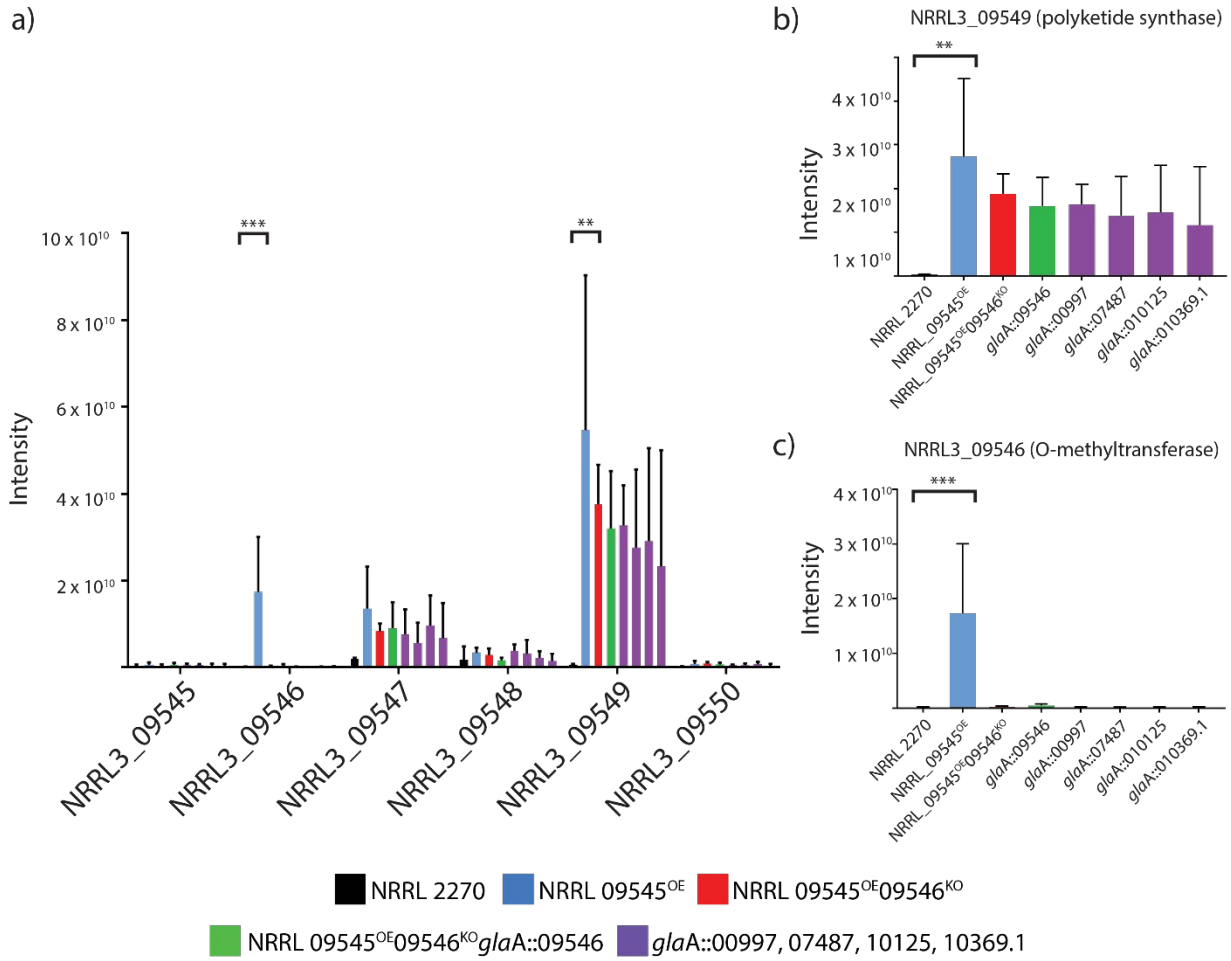


Figure 20: Protein abundance in the TAN/BMS biosynthetic gene cluster. a) The intensity of protein abundance in each gene from the TAN/BMS biosynthetic gene cluster was examined and grouped by treatment type, **b)** Intensity of protein abundance in the polyketide synthase gene NRRL3_09549, **c)** Intensity of protein abundance in the *O*-methyltransferase gene NRRL_09546.

While the transcription factor gene NRRL3_09545 appears to have a low level of expression, I observed a higher level of expression among the other TAN/BMS genes, indicating abundant TAN/BMS synthesis. Additionally, compared to the wild type, I anticipated that a form of TAN/BMS was generated in all strains based on the expression of the polyketide synthase gene (**Figure 20b**). However, the expression of the *O*-methyltransferase gene was predominantly expressed in NRRL3_09546^{OE} (**Figure 20c**). A small level of *O*-methyltransferase expression was also observed in NRRL3_09545^{OE}09546^{KO}*gla*A::09546 (**Figure 20c**), corresponding to the

metabolomic profile observed for this strain (**Figure 16**). Altogether, these results indicate that the methylated form of TAN/BMS was generated in NRRL3_09546^{OE} and to a lesser extent in NRRL3_09545^{OE}09546^{KO}*gla*A::09546, however the non-methylated form of TAN/BMS was produced in the *gla*A mutants due to the low expression of methyltransferases in these strains.

Statistics were computed using a One-way ANOVA followed by Dunnett's multiple comparisons test to examine the level of protein expression in each gene-edited strain compared to the wild type (**Appendix 4)**. Significant differences were observed in NRRL3_09545^{OE} corresponding to the expression of the *O*-methyltransferase ($p < 0.0005$), and the polyketide synthase ($p < 0.005$). No other significant relationships were observed, possibly suggesting methyltransferases must be locally clustered in order to be expressed.

The level of protein expression was also visualized using a heatmap, in which genes in the TAN/BMS cluster were compared to each gene-edited strain. **Figure 21** demonstrates that protein expression is largest in the polyketide synthase NRRL3_09549, and that the *O*-methyltransferase NRRL3_09546 is predominantly expressed in NRRL3_09545^{OE}. These trends validate the production of TAN/BMS and the non-methylated TAN/BMS observed in the metabolomic analyses (**Figure 6**, **Figure 16**, and **Figure 19**).

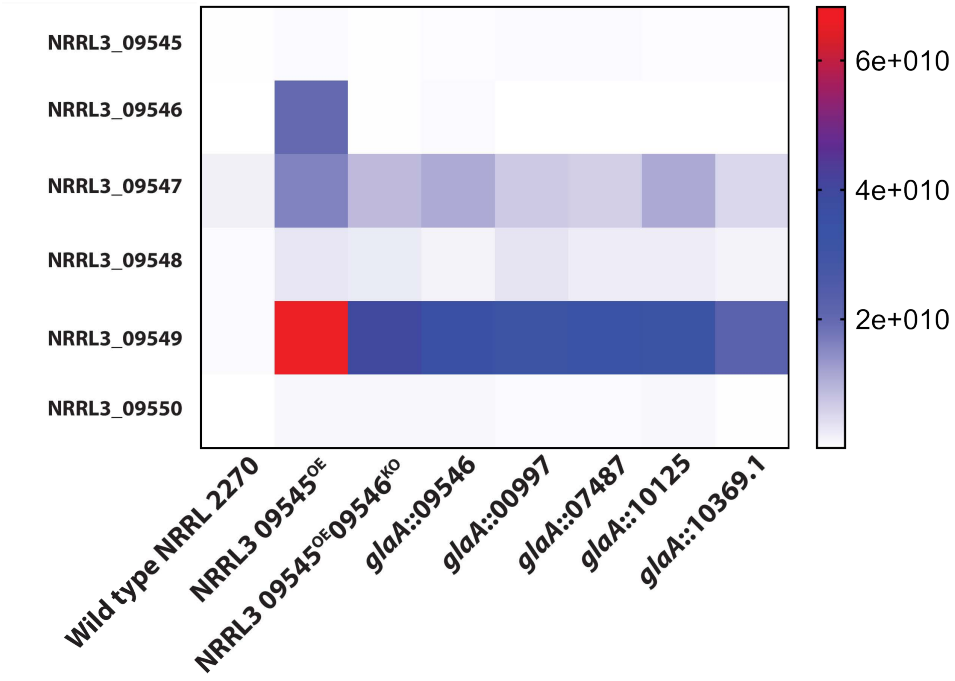


Figure 21: TAN/BMS biosynthetic gene cluster relative abundance versus treatment type. The expression level of genes in the TAN/BMS cluster compared to treatment type.

4. Discussion

Small-molecule chemical modifications, such as methyl group addition, is a relevant strategy to undergo molecule optimization in order to modify drug properties such as alleviating steric hinderance at the active site and improving bioavailability. Compared to other studies that aim to add methyl groups to molecules using chemical approaches (*e.g.,* the magic methyl effect), the combinatorial biosynthesis of TAN/BMS was optimized *in vivo* using *A. niger* as a cell factory. Within the *A. niger* genome are subclasses of methyltransferases that comprise DNA/RNA methyltransferases, protein methyltransferases, and secondary metabolite methyltransferases (https://genome.fungalgenomics.ca/). As such, I hypothesize that different methyltransferases can modify the methylation pattern of a molecule, and in doing so different methyltransferases can be tested for the production of novel secondary metabolites.

An *A. niger* intraspecies phylogenetic tree was made to gain insights into the functional specificity of methyltransferases, and to select those involved in secondary metabolite biosynthesis. Phylogenetic tree construction was performed based on the techniques described by Song *et al*., however a phylogenetic tree comparing the relatedness of secondary metabolite methyltransferases has not previously been constructed. In contrast, phylogenetic trees analyzing epigenetic methyltransferases have been previously used to identify the relatedness of enzymes controlling plant cell wall synthesis (Wu *et al*., 2013).The intraspecies NRRL3 phylogenetic tree contained 105 methyltransferase genes. Combining characterized methyltransferases from the literature and well-supported clades, a selection of methyltransferases with similar function, but diverse specificity, were identified as candidates. In parallel, a pairwise matrix was constructed, representing the amino acid sequence similarity between the 105 genes of the phylogenetic tree. Considering the pairwise matrix was generated using a list of intraspecies proteins with similar function, the percent identity was low and ranged from 4 – 72%. Overall, using a percent identity threshold score of 20%, three main clusters were identified and labeled on the phylogenetic tree. These clusters corresponded to clades with high bootstrap values, revealing methyltransferases that encoded similar functions. Genes with similar functional diversity were hypothesized to influence alternative methyl group patterning while maintaining selectivity. In total, the 13 methyltransferase genes (including the native TAN/BMS methyltransferase) were selected as candidates and constructed in an *E. coli* gene library.

To test the hypothesis that different methyltransferases can modify the methylation pattern of a molecule, a model biosynthetic gene cluster was chosen as a proof-of-concept. Among the 17 characterized NRRL3 biosynthetic gene clusters that included a methyltransferase gene, the TAN/BMS biosynthetic gene cluster was selected. In particular, the six genes encoding the biosynthesis of TAN/BMS have been identified and the product has been characterized by mass spectrometry, facilitating the analysis of the methylation pattern (Li *et al*. 2011). The *O*-methyltransferase gene NRRL3_09546 functions to methylate the final product at the C9 position, however TAN/BMS contains four additional sites in which methyl group addition could take place (*e.g.,* C5 or C7, or possibly at C2 or C15). The gene editing strategy used in this project involved 1) the in-cluster overexpression of the TAN/BMS biosynthetic gene cluster through the insertion of a strong and inducible promoter gene upstream of the transcription factor NRRL3_09545, and

2) the simultaneous out-of-cluster overexpression of individual methyltransferase library genes at *gla*A. In doing so, the specificity of methyltransferases was investigated in addition to examining the regulation and expression profiles of genes in an out-of-cluster design. Indeed, gene replacement at *gla*A has been successfully utilized for enzyme overproduction (van Lanen JM 1968; Pel *et al*. 2007; Andersen *et al*. 2011; Dong *et al*. 2019; Dong *et al*. 2020), suggesting that methyltransferase overexpression would also be elevated in this manner.

The overexpression of the TAN/BMS biosynthetic gene cluster and the subsequent overproduction of the secondary metabolite was successful with a 1500-fold increase compared to the wild type strain. In addition, the overproduction of TAN/BMS corresponded to a strong yellow pigmentation, facilitating the visual screening of successful transformants. The production of a yellow pigment is consistent with the findings reported by Li *et al*., in which overexpression of the transcription factor AdaR was achieved using the *A. nidulans* gpd*A* promoter, and the subsequent induction of the clustered genes on glucose rich media led to the production of TAN/BMS. Gene knockout of the native methyltransferase gene NRRL3_09546 was examined using mass spectrometry by comparing the expression profiles of the wild type strain, NRRL3_09545^{OE}, and NRRL3_09545^{OE}09546^{KO}, in which mass chromatograms were generated over a mass range expected to include the *m/z* values corresponding to TAN/BMS and TAN/BMS analogs. The peak corresponding to TAN/BMS was identified in NRRL3_09545^{OE} at a retention time of 10.30 min, and a TAN/BMS analog was identified at 9.06 min. The *m/z* of NRRL3_09545^{OE} corresponded to the literature value reported for TAN/BMS (415.1009 Da), while the *m/z* of the TAN/BMS analog produced by NRRL3_09545^{OE}09546^{KO} was equal to the mass of TAN/BMS minus a methyl group (401.0852 Da). The knockout strain also produced a yellow pigment, however no previous publication has characterized the knockout of NRRL3_09456. No peaks were evident in the mass chromatograms of the wild type strain, as secondary metabolite production is silent outside of ecological conditions or without a regulation mechanism.

While *A. niger* encompasses a large genetic toolbox that includes the Δ*kus*A modification to drive repair through homologous recombination, in addition to precise gene-editing using the CRISPR-Cas9 system, verification of the knockout through PCR yielded a product that was larger than the expected size. To examine the gene knockout, a repair template was designed to amplify the entire methyltransferase gene using primers upstream and downstream of the native methyltransferase coding sequence. Then, to confirm the status of the NRRL3_09546 gene deletion, Sanger sequencing was performed using the PCR products. Based on the sequence alignment of the knockout to the overexpression strain, it appeared that a 116 bp sequence remained, and resulted in the partial deletion of NRRL3_09546. However, examination of the *O*-methyltransferase functional domain (https://genome.fungalgenomics.ca/) as well as the mass spectrometry expression profile of NRRL3_09545^{OE}09546^{KO}, indicated that the functional domain was indeed deleted. The remaining 116 bp sequence is believed to be a region unrelated to methyltransferase function and which may function as a cytoplasmic domain based on its current bioinformatic annotation (https://genome.fungalgenomics.ca/). Considering not all publications report the systematic sequencing of their mutant strains, sequencing was performed to achieve a comprehensive understanding of the gene deletion. Future improvements can be made to redesign

the CRISPR guide RNA used to target the methyltransferase gene, however it was not immediately considered, as the knockout resulted in a non-functional gene. Additional screening was pursued by examining the molecular structures of TAN/BMS and the TAN/BMS analog produced by the strains NRRL3_09545^{OE} and NRRL3_09545^{OE}09546^{KO} using ^{1}H NMR ^{13}C NMR. Overall, the NRM profiles confirmed the presence or the absence of C9 methylation in NRRL3 09545^{OE} and NRRL3 09545^{OE}09546^{KO}, respectively.

The NRRL3_09545^{OE}09546^{KO} strain was subsequently used to test the methyltransferase library. The deletion of the gene encoding NRRL3_09546 and the resulting production of the non-methylated TAN/BMS product did not influence the strong yellow pigmentation observed in NRRL3_09545^{OE}, which was previously reported as being correlated to the expression of the genes encoding the polyketide synthase (*ada*A) and thioesterase (*ada*B) (Li *et al.,* 2011). A simple and cost-effective screening method using spectrophotometry was also developed to check for positive transformants, in which the TAN/BMS absorbance value reported in the literature (*i.e.,* 414 nm) (Shu *et al.,* 1995) was compared to TAN/BMS absorbance values of the gene-edited strains. Spectrophotometric assays have previously been used to screen for the presence of secondary metabolites in fungi isolated form the environment (Nischitha1 and Shivanna, 2022), indicating that similar procedures could be used to test the molecular and chemical status of TAN/BMS analogues. Further, it has been documented that *O*-methylation can alter the spectral properties of secondary metabolites (Malaj *et al*., 2013) justifying the use of this screening method in the present study. The wild type strain demonstrated no absorbance peak at 414 nm, while NRRL3_09545^{OE} had an absorption maximum at 421 nm. Considering a full scan was performed in the visible light spectrum (*i.e.,* 350 – 500 nm), the difference noted between the observed and expected wavelength may be due to the calibration of the spectrophotometer. The strain NRRL3_09545^{OE}09546^{KO} showed an absorption maximum at 412 nm, and the *gla*A::09546 mutant strain demonstrated an absorbance maxima at 416 nm. Considering NRRL3_09545^{OE}09546^{KO} produced a TAN/BMS analog that does not have C9 methylation, the absorbance maximum was predictably different compared to NRRL3_09545^{OE}. The absorbance maxima for the *gla*A::09546 mutant also demonstrated a level of absorbance different from NRRL3_09545^{OE}, indicating that spectrophotometry could potentially be used to compare the methylation status of the methyltransferase gene library mutants. Furthermore the level of absorbance corresponding to the *gla*A::09546 mutant was intermediate between NRRL3_09545^{OE} and NRRL3_09545^{OE}09546^{KO}, suggesting that a different methylation pattern was produced. However, to better understand the methylation status, more robust screening methods were pursued.

The secondary metabolite profiles were then analysed by mass spectrometry to confirm the *m/z* of different TAN/BMS gene-edited strains. Mass chromatograms and spectra were generated to confirm the retention times and *m/z* corresponding to TAN/BMS and TAN/BMS analogs. First, the wild type strain, parental strains NRRL3_09545^{OE} and NRRL3_09545^{OE}09546^{KO}, as well as the gene-edited strain NRRL3_09545^{OE}09546^{KO}*gla*A::09546 were compared. The wild type strain produced no TAN/BMS or TAN/BMS analogs, while NRRL3_09545^{OE} predominantly produced the methylated form of TAN/BMS (*m/z* 415.1009, 10.30 min) and NRRL3_09545^{OE}09546^{KO} only produced the non-methylated form of TAN/BMS (*m/z* 401.0852, 9.06 min). The *gla*A::09546

mutant however, displayed an expression profile corresponding to both TAN/BMS products (*m/z* 415.1009, 10.29 min and *m/z* 401.0852, 9.05 min, respectively). Interestingly, NRRL3_09545^{OE} also appeared to produce both TAN/BMS products, in which a small level of the non-methylated product was present (*m/z* 401.0852, 9.05 min). While this phenomenon has not been previously described for other genes, it suggests the possibility that gene replacement at the *gla*A locus affects the synthesis of TAN/BMS. In addition, the production of the non-methylated form of TAN/BMS in the overexpression strain possibly indicates that overexpression of the TAN/BMS cluster does not achieve complete methylation of the product.

To examine the extent of methylated versus non-methylated forms of TAN/BMS in each the NRRL3_09545^{OE} and the *gla*A::09546 mutant strains, growth assays were repeated and additional TAN/BMS expression profiles were generated and compared. In NRRL3_09545^{OE}, the estimated ratio between the extent of the methylated and non-methylated TAN/BMS product was 1/10, respectively, while in the *gla*A::09546 mutant strain, the estimated ratio between the extent of the methylated and non-methylated TAN/BMS product was 1/3, respectively. These outcomes correlate to the spectrophotometry results in which shifts were observed in the absorbance peak maximum between NRRL3_09545^{OE} (421 nm) and *gla*A::09546 (416 nm). Despite the commonly accepted architecture of the biosynthetic gene clusters described in the introduction, the regulatory mechanisms involved in the production of secondary metabolites are complex and remain largely unknown. The partial restauration of the TAN/BMS methylation profile when the native methyltransferase gene is overexpressed out-of-cluster potentially suggests that the location of the gene is important for the methylation process.

The production of TAN/BMS was then examined in the gene-edited strains in which ANEp8-Cas9-sgRNA(*gla*A) was used to transform methyltransferase genes using the strain NRRL3_09545^{OE}09546^{KO}. The replacement of the methyltransferase library genes in *gla*A was validated by PCR for each mutant strain, and the *m/z* expression profiles of the mutants were compared to the NRRL3_09545^{OE}09546^{KO} strain to examine the TAN/BMS methylation status and to discern the extent of TAN/BMS production. It was hypothesized that peaks corresponding to methylated TAN/BMS would be present with a *m/z* of 415.1009, however methylation patterns would be later tested by NMR to confirm the molecular configuration. NMR is particularly valuable when examining the molecular optimization of secondary metabolites, in which combinatorial biosynthesis can lead to changes in the configuration and conformation of a secondary metabolite (Broadhurst *et al*., 2003). Overall, no peaks corresponding to the *m/z* of TAN/BMS were identified, but peaks corresponding to the non-methylated product were observably larger compared to NRRL3_09545^{OE}09546^{KO}.

To better understand TAN/BMS production, the relative abundance of protein-encoding genes was examined to gain insights on expression and production levels in each mutant strain. As such, the overexpression of the TAN/BMS biosynthetic gene cluster and overproduction of the methyltransferases in *gla*A was assessed by an intracellular proteomic analysis. As previously mentioned, the *gla*A promoter has been successfully used to overexpress genes (Andersen *et al*., 2011, Kwon *et al*., 2012, Evdokias *et al*., 2021), validating the use of out-of-cluster gene replacement in this analysis. Overall, the relative protein abundance appeared to be the lowest in

the wild type strain, and was the highest in the NRRL3_09545^{OE} strain. More specifically, statistical differences were noted when comparing the wild type to NRRL3_09545^{OE} regarding the relative protein abundance of the *O*-methyltransferase gene NRRL3_09546 ($p < 0.0005$), and the polyketide synthase gene NRRL 09549 ($p < 0.005$). Combined with the metabolomic analysis, these results are the first to substantiate that the biosynthesis of TAN/BMS is present in NRRL3_09545^{OE}, as high intracellular protein levels of the *O*-methyltransferase and polyketide synthase genes were observed, which ultimately corresponded to high TAN/BMS production extracellularly. In contrast, the absence of *O*-methyltransferase production in the *gla*A gene-edited strain may correspond to the high levels of the non-methylated TAN/BMS analog that was observed in metabolomic profiles. The low extent of methylation observed in the *gla*A::09546 is yet to be investigated, and further studies will be pursued. Furthermore, at this stage it is not possible to conclude on the specificity of the methyltransferases regarding TAN/BMS methylation. However, the localization of methyltransferases in the genome, in- or out-of-cluster, seems to play an important role in the methylation patterning of TAN/BMS. A more complex regulatory mechanism could be involved in addition to the transcription factor, and is yet to be investigated.

5. Conclusions and Future Directions

The proof-of-concept described herein posits a mechanism for the *in vivo* modification of secondary metabolites using *A. niger* as a cell factory. By applying different synthetic biology approaches, the overexpression of the TAN/BMS biosynthetic gene cluster and the subsequent knockout of the *O*-methyltransferase gene NRRL3_09546 was performed. An intraspecies methyltransferase library consisting of 13 genes was designed and individual genes were inserted in NRRL3_09545^{OE}09546^{KO} at the *gla*A locus. Screening methods were used to examine methyltransferase gene insertion and changes in methylation patterning. The insertion of methyltransferase genes at *gla*A was validated for all mutants, and TAN/BMS expression was confirmed in NRRL3_09545^{OE} based on *m/z* metabolomic profiles and intracellular protein abundance. However, methyltransferase mutants did not recapitulate methylation nor modify TAN/BMS. To continue studies examining the modification and optimization of secondary metabolites, two experimental approaches will be pursued. First, the methyltransferase library will be inserted at the TAN/BMS locus to identify whether the co-localization of genes affects the selectivity of methyltransferases, and if the abundance of TAN/BMS products will be elevated as observed in NRRL3_09545^{OE}. In addition, to further increase the chemical diversity of secondary metabolites, other biosynthetic gene clusters containing methyltransferases (*e.g.,* Pyranonigrin) will be probed.

6. References

Abotaleb M, Liskova A, Kubatka P, Büsselberg D. 2020. Therapeutic Potential of Plant Phenolic Acids in the Treatment of Cancer. Biomolecules. 10(2):221. doi:10.3390/biom10020221.

Adessi C and Soto C. 2002. Converting a Peptide into a Drug: Strategies to Improve Stability and Bioavailability. Current Medicinal Chemistry. 9(9):963–978. doi:10.2174/0929867024606731.

Aguilar-Pontes MV, Brandl J, McDonnell E, Strasser K, Nguyen TTM, Riley R, Mondo S, Salamov A, Nybo JL, Vesth TC, *et al*. 2018. The gold-standard genome of *Aspergillus niger* NRRL 3 enables a detailed view of the diversity of sugar catabolism in fungi. Studies in Mycology. 91(1):61–78. doi:10.1016/j.simyco.2018.10.001.

Albright JC, Henke MT, Soukup AA, McClure RA, Thomson RJ, Keller NP, Kelleher NL. 2015. Large-Scale Metabolomics Reveals a Complex Response of *Aspergillus* nidulans to Epigenetic Perturbation. ACS Chemical Biology. 10(6):1535–1541. doi:10.1021/acschembio.5b00025.

Almassi F, Ghisalberti EL, Rowland CY. 1994. Alkylcitrate-Derived Metabolites from *Aspergillus niger*. Journal of Natural Products. 57(6):833–836. doi:10.1021/np50108a024.

Amare MG and Keller NP. 2014. Molecular mechanisms of *Aspergillus* flavus secondary metabolism and development. Fungal Genetics and Biology. 66:11–18. doi:10.1016/j.fgb.2014.02.008.

Anderegg RJ, Biemann K, Buechi G, Cushman M. 1976. Malformin C, a new metabolite of *Aspergillus niger*. Journal of the American Chemical Society. 98(11):3365–3370. doi:10.1021/ja00427a051.

Andersen MR, Salazar MP, Schaap PJ, van de Vondervoort PJI, Culley D, Thykaer J, Frisvad JC, Nielsen KF, Albang R, Albermann K, *et al*. 2011. Comparative genomics of citric-acid-producing *Aspergillus niger* ATCC 1015 versus enzyme-producing CBS 513.88. Genome Research. 21(6):885–897. doi:10.1101/gr.112169.110.

Anderson JG and Smith JE. 1971. The Production of Conidiophores and Conidia by Newly Germinated Conidia of *Aspergillus niger* (Microcycle Conidiation). Journal of General Microbiology. 69(2):185–197. doi:10.1099/00221287-69-2-185.

Ang EL, Sun H, Liu Z, Zhao H. 2015 Feb. Recent advances in combinatorial biosynthesis for drug discovery. Drug Design, Development and Therapy. 9:823-833. doi:10.2147/DDDT.S63023.

Arentshorst M, Lagendijk EL, Ram AF. 2015. A new vector for efficient gene targeting to the pyrG locus in *Aspergillus niger*. Fungal Biology and Biotechnology. 2(1):2. doi:10.1186/s40694-015-0012-4

Aslanidis C, and de Jong PJ. 1990. Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Research. 18(20):6069–6074. doi:10.1093/nar/18.20.6069.

Aynetdinova D, Callens MC, Hicks HB, Poh CYX, Shennan BDA, Boyd AM, Lim ZH, Leitch JA,

Dixon DJ. 2021. Installing the “magic methyl” – C–H methylation in synthesis. Chemical Society Reviews. 50(9):5517–5563. doi:10.1039/D0CS00973C.

Ayté J, Leis JF, DeCaprio JA. 1997. The fission yeast protein p73res2 is an essential component of the mitotic MBF complex and a master regulator of meiosis. Molecular and Cellular Biology. 17(11):6246–6254. doi:10.1128/MCB.17.11.6246.

Bajorath J. 2002. Integration of virtual and high-throughput screening. Nature Reviews Drug Discovery. 1(11):882–894. doi:10.1038/nrd941.

Baker SE. 2006. *Aspergillus niger* genomics: Past, present and into the future. Medical Mycology. 44(s1):17–21. doi:10.1080/13693780600921037.

Ball B, Bermas A, Carruthers-Lay D, Geddes-McAlister J. 2019. Mass Spectrometry-Based Proteomics of Fungal Pathogenesis, Host–Fungal Interactions, and Antifungal Development. Journal of Fungi. 5(2):52. doi:10.3390/jof5020052.

Ball B and Geddes-McAlister J. 2019. Quantitative Proteomic Profiling of Cryptococcus neoformans. Current Protocols in Microbiology. 55(1). doi:10.1002/cpmc.94.

Bauerle MR, Schwalm EL, Booker SJ. 2015. Mechanistic Diversity of Radical S-Adenosylmethionine (SAM)-dependent Methylation. Journal of Biological Chemistry. 290(7):3995–4002. doi:10.1074/jbc.R114.607044.

Bayram Ö and Braus GH. 2012. Coordination of secondarymetabolism and development in fungi: the velvet familyof regulatory proteins. FEMS Microbiology Reviews. 36(1):1–24. doi:10.1111/j.1574-6976.2011.00285.x.

Bibb MJ. 2005. Regulation of secondary metabolism in streptomycetes. Current Opinion in Microbiology. 8(2):208–215. doi:10.1016/j.mib.2005.02.016..

Blachowicz A, Raffa N, Bok JW, Choera T, Knox B, Lim FY, Huttenlocher A, Wang CCC, Venkateswaran K, Keller NP. 2020. Contributions of Spore Secondary Metabolites to UV-C Protection and Virulence Vary in Different *Aspergillus* fumigatus Strains. Andrew Alspaugh J, editor. mBio. 11(1). doi:10.1128/mBio.03415-19.

Bos CJ, Debets AJM, Swart K, Huybers A, Kobus G, Slakhorst SM. 1988. Genetic analysis and the construction of master strains for assignment of genes to six linkage groups in *Aspergillus niger*. Current Genetics. 14(5):437–443. doi:10.1007/BF00521266.

Bourgaud F, Gravot A, Milesi S, Gontier E. 2001. Production of plant secondary metabolites: a historical perspective. Plant Science. 161(5):839–851. doi:10.1016/S0168-9452(01)00490-3.

Broadhurst RW, Nietlispach D, Wheatcroft MP, Leadlay PF, Weissman KJ. 2003. The structure of docking domains in modular polyketide synthases. Chem. Biol. 10, 723–731. doi: 10.1016/S1074-5521(03)00156-X.

Buchanan R and, Ayres JC. 1975. Effect of initial pH on aflatoxin production. Applied Microbiology. 30(6):1050–1051. doi:10.1128/am.30.6.1050-1051.1975.

Bugni TS, Abbanat D, Bernan VS, Maiese WM, Greenstein M, Van Wagoner RM, Ireland CM. 2000. Yanuthones: Novel Metabolites from a Marine Isolate of *Aspergillus* ni ger. The Journal of Organic Chemistry. 65(21):7195–7200. doi:10.1021/jo0006831.

Cairns TC, Nai C, Meyer V. 2018 How a fungus shapes biotechnology: 100 years of Aspergillus niger research. Fungal Biol Biotechnol. 5(13). doi: 10.1186/s40694-018-0054-5.

Carvalho NDSP, Arentshorst M, Jin Kwon M, Meyer V, Ram AFJ. 2010. Expanding the ku70 toolbox for filamentous fungi: establishment of complementation vectors and recipient strains for advanced gene analyses. Applied Microbiology and Biotechnology. 87(4):1463–1473.

Castilho LR, Medronho RA, Alves TL. 2000. Production and extraction of pectinases obtained by solid state fermentation of agroindustrial residues with *Aspergillus niger*. Bioresource Technology. 71(1):45–50. doi:10.1016/S0960-8524(99)00058-9.

Challenger F and Klein L. 1929. The formation of l-malic acid from fumaric acid by *Aspergillus niger*. J Chem Soc.1644–1647. doi:10.1039/JR9290001644.

Challis GL and Hopwood DA. 2003. Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by Streptomyces species. Proceedings of the National Academy of Sciences. 100(suppl_2):14555–14561. doi:10.1073/pnas.1934677100.

Chang HHY, Pannunzio NR, Adachi N, Lieber MR. 2017. Non-homologous DNA end joining and alternative pathways to double-strand break repair. Nature Reviews Molecular Cell Biology. 18(8):495–506. doi:10.1038/nrm.2017.48.

Chang PK, Ehrlich KC, Yu J, Bhatnagar D, Cleveland TE. 1995. Increased expression of *Aspergillus* parasiticus aflR, encoding a sequence-specific DNA-binding protein, relieves nitrate inhibition of aflatoxin biosynthesis. Applied and Environmental Microbiology. 61(6):2372–2377. doi:10.1128/aem.61.6.2372-2377.1995.

Chiang Y-M, Meyer KM, Praseuth M, Baker SE, Bruno KS, Wang CCC. 2011. Characterization of a polyketide synthase in *Aspergillus niger* whose product is a precursor for both dihydroxynaphthalene (DHN) melanin and naphtho-γ-pyrone. Fungal Genetics and Biology. 48(4):430–437. doi:10.1016/j.fgb.2010.12.001.

Clikeman JA, Khalsa GJ, Barton SL, Nickoloff JA. 2001. Homologous Recombinational Repair of Double-Strand Breaks in Yeast Is Enhanced by MAT Heterozygosity Through yKU-Dependent and -Independent Mechanisms. Genetics. 157(2):579–589. doi:10.1093/genetics/157.2.579.

Coleman JJ, Ghosh S, Okoli I, Mylonakis E. 2011. Antifungal Activity of Microbial Secondary Metabolites. Chaturvedi V, editor. PLoS ONE. 6(9):e25321. doi:10.1371/journal.pone.0025321.

Cramer J, Sager CP, Ernst B. 2019. Hydroxyl Groups in Synthetic and Natural-Product-Derived Therapeutics: A Perspective on a Common Functional Group. Journal of Medicinal Chemistry. 62(20):8915–8930. doi:10.1021/acs.jmedchem.9b00179.

Cramer RA, Stajich JE, Yamanaka Y, Dietrich FS, Steinbach WJ, Perfect JR. 2006. Phylogenomic analysis of non-ribosomal peptide synthetases in the genus *Aspergillus*. Gene. 383:24–32. doi:10.1016/j.gene.2006.07.008.

Critchlow SE and Jackson SP. 1998. DNA end-joining: from yeast to man. Trends in Biochemical Sciences. 23(10):394–398. doi:10.1016/S0968-0004(98)01284-5.

CSFG genomes. Accessed 2022 August 16. https://www.fungalgenomics.ca/.

Cui JJ, Tran-Dubé M, Shen H, Nambu M, Kung P-P, Pairish M, Jia L, Meng J, Funk L, Botrous I, *et al*. 2011. Structure Based Drug Design of Crizotinib (PF-02341066), a Potent and Selective Dual Inhibitor of Mesenchymal–Epithelial Transition Factor (c-MET) Kinase and Anaplastic Lymphoma Kinase (ALK). Journal of Medicinal Chemistry. 54(18):6342–6363. doi:10.1021/jm2007613.

Currie JN. 1917. The citric acid fermentation of *Aspergillus niger*. Journal of Biological Chemistry. 31(1):15–37. doi:10.1016/S0021-9258(18)86708-4.

Cutler HG, Crumley FG, Cox RH, Hernandez O, Cole RJ, Dorner JW. 1979. Orlandin: a nontoxic fungal metabolite with plant growth inhibiting properties. Journal of Agricultural and Food Chemistry. 27(3):592–595. doi:10.1021/jf60223a043.

Dai J and Mumper RJ. 2010. Plant Phenolics: Extraction, Analysis and Their Antioxidant and Anticancer Properties. Molecules. 15(10):7313–7352. doi:10.3390/molecules15107313.

David H, Åkesson M, Nielsen J. 2003. Reconstruction of the central carbon metabolism of *Aspergillus niger*. European Journal of Biochemistry. 270(21):4243–4253. doi:10.1046/j.1432-1033.2003.03798.x.

Debets AJM and Bos CJ. 1986. Isolation of small protoplasts from *Aspergillus niger*. Fungal Genetics Reports. 33(1):24. doi:10.4148/1941-4765.1575.

Debets F, Swart K, Hoekstra RF, Bos CJ. 1993. Genetic maps of eight linkage groups of *Aspergillus niger* based on mitotic mapping. Current Genetics. 23(1):47–53. doi:10.1007/BF00336749.

Demirci E, Arentshorst M, Yilmaz B, Swinkels A, Reid ID, Visser J, Tsang A, Ram AFJ. 2021. Genetic Characterization of Mutations Related to Conidiophore Stalk Length Development in *Aspergillus niger* Laboratory Strain N402. Frontiers in Genetics. 12. doi:10.3389/fgene.2021.666684.

Derkx P, Madrid S. 2001. The *Aspergillus niger* cypA gene encodes a cyclophilin that mediates sensitivity to the immunosuppressant cyclosporin A. Molecular Genetics and Genomics. 266(4):527–536. doi:10.1007/s004380100586.

Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, Smith I, Tothova Z, Wilen C, Orchard R, *et al*. 2016. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nature Biotechnology. 34(2):184–191. doi:10.1038/nbt.3437.

Dolan SK, O'Keeffe G, Jones GW, Doyle S. 2015. Resistance is not futile: gliotoxin biosynthesis, functionality and utility. Trends in Microbiology. 23(7):419–428. doi:10.1016/j.tim.2015.02.005.

Dong H, Zheng J, Yu D, Wang B, Pan L. 2019. Efficient genome editing in *Aspergillus niger* with an improved recyclable CRISPR-HDR toolbox and its application in introducing multiple copies of heterologous genes. Journal of Microbiological Methods. 163:105655. doi:10.1016/j.mimet.2019.105655.

Dong L, Yu D, Lin X, Wang B, Pan L. 2020. Improving expression of thermostable trehalase from Myceliophthora sepedonium in *Aspergillus niger* mediated by the CRISPR/Cas9 tool and its purification, characterization. Protein Expression and Purification. 165:105482. doi:10.1016/j.pep.2019.105482.

Drewes AM, Jensen RD, Nielsen LM, Droney J, Christrup LL, Arendt-Nielsen L, Riley J, Dahan A. 2013. Differences between opioids: pharmacological, experimental, clinical and economical perspectives. British Journal of Clinical Pharmacology. 75(1):60–78. doi:10.1111/j.1365-2125.2012.04317.x.

Duran R, Cary JW, Calvo AM. 2010. Role of the Osmotic Stress Regulatory Pathway in Morphogenesis and Secondary Metabolism in Filamentous Fungi. Toxins. 2(4):367–381. doi:10.3390/toxins2040367.

Edgar, RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32(5):1792-1797. doi: 10.1093/nar/gkh340

Eisenman HC and Casadevall A. 2012. Synthesis and assembly of fungal melanin. Applied microbiology and biotechnology. 93(3):931–40. doi:10.1007/s00253-011-3777-2.

Espeso EA, Tilburn J, Arst HN, Peñalva MA. 1993. pH regulation is a major determinant in expression of a fungal penicillin biosynthetic gene. The EMBO Journal. 12(10):3947–3956. doi:10.1002/j.1460-2075.1993.tb06072.x.

Evdokias G, Semper C, Mora-Ochomogo M, Di Falco M, Minh Nguyen TT, Savchenko A, Tsang A Benoit-Gelber I. 2021. Identification of a Novel Biosynthetic Gene Cluster in Aspergillus niger Using Comparative Genomics. J. Fungi. 7(5), 374. doi: 10.3390/jof7050374.

Feng K, Quevedo RE, Kohrt JT, Oderinde MS, Reilly U, White MC. 2020. Late-stage oxidative C(sp3)–H methylation. Nature. 580(7805):621–627. doi:10.1038/s41586-020-2137-8.

Ferracin LM, Fier CB, Vieira MLC, Monteiro-Vitorello CB, Varani A de M, Rossi MM, Müller-Santos M, Taniwaki MH, Thie Iamanaka B, Fungaro MHP. 2012. Strain-specific polyketide synthase genes of *Aspergillus niger*. International Journal of Food Microbiology. 155(3):137–145. doi:10.1016/j.ijfoodmicro.2012.01.020.

Fiedler MRM, Barthel L, Kubisch C, Nai C, Meyer V. 2018. Construction of an improved *Aspergillus niger* platform for enhanced glucoamylase secretion. Microbial Cell Factories. 17(1):95. doi:10.1186/s12934-018-0941-8.

Firn RD and Jones CG. 2000. The evolution of secondary metabolism - a unifying model. Molecular Microbiology. 37(5):989–994. doi:10.1046/j.1365-2958.2000.02098.x.

Fleming A. 1932. On the specific antibacterial properties of penicillin and potassium tellurite. Incorporating a method of demonstrating some bacterial antagonisms. The Journal of Pathology and Bacteriology. 35(6):831–842. doi:10.1002/path.1700350603.

Fleming A. 2001. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. 1929. Bulletin of the World Health Organization. 79(8):780–90.

Foster JW and Karow EO. 1945. Microbiological Aspects of Penicillin. Journal of Bacteriology. 49(1):19–29. doi:10.1128/jb.49.1.19-29.1945.

Fountain JC, Bajaj P, Pandey M, Nayak SN, Yang L, Kumar V, Jayale AS, Chitikineni A, Zhuang W, Scully BT, *et al*. 2016. Oxidative stress and carbon metabolism influence *Aspergillus* flavus transcriptome composition and secondary metabolite production. Scientific Reports. 6(1):38747. doi:10.1038/srep38747.

Fowler T, Berka RM, Ward M. 1990. Regulation of the glaA gene of *Aspergillus niger*. Current Genetics. 18(6):537–545. doi:10.1007/BF00327025.

Franken ACW, Lechner BE, Werner ER, Haas H, Lokman BC, Ram AFJ, van den Hondel CAMJJ, de Weert S, Punt PJ. 2014. Genome mining and functional genomics for siderophore production in *Aspergillus niger*. Briefings in Functional Genomics. 13(6):482–492. doi:10.1093/bfgp/elu026.

Frisvad J.C., Larsen TO, de Vries R, Meijer M, Houbraken J, Cabañes FJ, Ehrlich K, Samson RA. 2007. Secondary metabolite profiling, growth profiles and other tools for species recognition and important *Aspergillus* mycotoxins. Studies in Mycology. 59:31–37. doi:10.3114/sim.2007.59.04.

Frisvad JC, Møller LLH, Larsen TO, Kumar R, Arnau J. 2018. Safety of the fungal workhorses of industrial biotechnology: update on the mycotoxin and secondary metabolite potential of *Aspergillus niger*, *Aspergillus* oryzae, and Trichoderma reesei. Applied Microbiology and Biotechnology. 102(22):9481–9515. doi:10.1007/s00253-018-9354-1.

Frisvad Jens C., Smedsgaard J, Samson RA, Larsen TO, Thrane U. 2007. Fumonisin B 2 Production by *Aspergillus niger*. Journal of Agricultural and Food Chemistry. 55(23):9727–9732. doi:10.1021/jf0718906.

Fukuda T, Hasegawa Y, Hagimori K, Yamaguchi Y, Masuma R, Tomoda H, Ōmura S. 2006. Tensidols, New Potentiators of Antifungal Miconazole Activity, Produced by *Aspergillus niger* FKI-2342. The Journal of Antibiotics. 59(8):480–485. doi:10.1038/ja.2006.67.

Gailey FB, Stefaniak JJ, Olson BH, Johnson MJ. 1946. A Comparison of Penicillin-producing Strains of Penicillium notatum-chrysogenum. Journal of bacteriology. 52(1):129–40.

Gaisser S, Lill R, Wirtz G, Grolle F, Staunton J, Leadlay PF. 2008. New erythromycin derivatives from Saccharopolyspora erythraea using sugar O-methyltransferases from the spinosyn biosynthetic gene cluster. Molecular Microbiology. 41(5):1223–1231. doi:10.1046/j.1365-2958.2001.02594.x.

Garcia-Conesa M-T, Kroon PA, Ralph J, Mellon FA, Colquhoun IJ, Saulnier L, Thibault J-F, Williamson G. 1999. A cinnamoyl esterase from *Aspergillus niger* can break plant cell wall cross-links without release of free diferulic acids. European Journal of Biochemistry. 266(2):644–652. doi:10.1046/j.1432-1327.1999.00910.x.

Geneious Prime 2021.2.2.(https://www.geneious.com)

Goldman GH, McGuire SL, Harris SD. 2002. The DNA Damage Response in Filamentous Fungi. Fungal Genetics and Biology. 35(3):183–195. doi:10.1006/fgbi.2002.1344.

Goosen T, Bloemheuvel G, Gysler C, de Bie DA, van den Broek HWJ, Swart K. 1987. Transformation of *Aspergillus niger* using the homologous orotidine-5′-phosphate-decarboxylase gene. Current Genetics. 11(6–7):499–503. doi:10.1007/BF00384612.

Gottlieb TM and Jackson SP. 1993. The DNA-dependent protein kinase: Requirement for DNA ends and association with Ku antigen. Cell. 72(1):131–142. doi:10.1016/0092-8674(93)90057-W.

GraphPad Prism v7.4 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com.

Guzman-Chavez F, Salo O, Samol M, Ries M, Kuipers J, Bovenberg RAL, Vreeken RJ, Driessen AJM. 2018. Deregulation of secondary metabolism in a histone deacetylase mutant of Penicillium chrysogenum. MicrobiologyOpen. 7(5):e00598. doi:10.1002/mbo3.598.

van Hartingsveldt W, Mattern IE, van Zeijl CMJ, Pouwels PH, van den Hondel CAMJJ. 1987. Development of a homologous transformation system for *Aspergillus niger* based on the pyrG gene. Molecular and General Genetics MGG. 206(1):71–75. doi:10.1007/BF00326538.

Hasegawa Y, Fukuda T, Hagimori K, Tomoda H, Ōmura S. 2007. Tensyuic Acids, New Antibiotics Produced by *Aspergillus niger* FKI-2342. Chemical and Pharmaceutical Bulletin. 55(9):1338–1341. doi:10.1248/cpb.55.1338.

Hiort J, Maksimenka K, Reichert M, Perović-Ottstadt S, Lin WH, Wray V, Steube K, Schaumann K, Weber H, Proksch P, *et al*. 2004. New Natural Products from the Sponge-Derived Fungus *Aspergillus niger*. Journal of Natural Products. 67(9):1532–1543. doi:10.1021/np030551d.

Holm DK, Petersen LM, Klitgaard A, Knudsen PB, Jarczynska ZD, Nielsen KF, Gotfredsen CH, Larsen TO, Mortensen UH. 2014. Molecular and Chemical Characterization of the Biosynthesis of the 6-MSA-Derived Meroterpenoid Yanuthone D in *Aspergillus niger*. Chemistry & Biology. 21(4):519–529. doi:10.1016/j.chembiol.2014.01.013.

Hoppert M, Gentzsch C, Schörgendorfer K. 2001. Structure and localization of cyclosporin synthetase, the key enzyme of cyclosporin biosynthesis in Tolypocladium inflatum. Archives of Microbiology. 176(4):285–293. doi:10.1007/s002030100324.

Hornbogen T, Pieper R, Hoffmann K, Kleinkauf H, Zocher R. 1992. Two new cyclophilins from Fusarium sambucinum and *Aspergillus niger*: Resistance of cyclophilin/cyclosporin a complexes against proteolysis. Biochemical and Biophysical Research Communications. 187(2):791–796. doi:10.1016/0006-291X(92)91265-R.

Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, *et al*. 2013. DNA targeting specificity of RNA-guided Cas9 nucleases. Nature Biotechnology. 31(9):827–832. doi:10.1038/nbt.2647.

Hurst PL, Nielsen J, Sullivan PA, Shepherd MG. 1977. Purification and properties of a cellulase from *Aspergillus niger*. Biochemical Journal. 165(1):33–41. doi:10.1042/bj1650033.

Inglis DO, Binkley J, Skrzypek MS, Arnaud MB, Cerqueira GC, Shah P, Wymore F, Wortman JR, Sherlock G. 2013. Comprehensive annotation of secondary metabolite biosynthetic genes and gene clusters of *Aspergillus* nidulans, A. fumigatus, A. *niger* and A. oryzae. BMC Microbiology. 13(1):91. doi:10.1186/1471-2180-13-91.

Isogai A, Washizu M, Kondo K, Murakoshi S, Suzuki A. 1984. Isolation and identification of (+)-hexylitaconic acid as a plant growth regulator. Agricultural and Biological Chemistry. 48(10):2607–2609. doi:10.1271/bbb1961.48.2607.

Iwamoto T, Shima S, Hirota A, Akira I, Sakai H. 1983. *Niger*azine B, a New Metabolite from *Aspergillus niger* Screening, Isolation, and Chemical and Biological Properties. Agricultural and Biological Chemistry. 47(4):739–743. doi:10.1080/00021369.1983.10865731.

Jacobson ES. 2000. Pathogenic roles for fungal melanins. Clinical microbiology reviews. 13(4):708–17. doi:10.1128/CMR.13.4.708.

Jefferson WE. 1967. The Isolation and Characterization of Asperenone, a New Phenylpolyene from *Aspergillus niger*. Biochemistry. 6(11):3479–3484. doi:10.1021/bi00863a019.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science. 337(6096):816–821. doi:10.1126/science.1225829.

Jungmann J, Reins HA, Lee J, Romeo A, Hassett R, Kosman D, Jentsch S. 1993. MAC1, a nuclear regulatory protein related to Cu-dependent transcription factors is involved in Cu/Fe utilization and stress resistance in yeast. The EMBO Journal. 12(13):5051–5056. doi:10.1002/j.1460-2075.1993.tb06198.x.

Keller G, Bird A, Winge DR. 2005. Independent Metalloregulation of Ace1 and Mac1 in Saccharomyces cerevisiae. Eukaryotic Cell. 4(11):1863–1871. doi:10.1128/EC.4.11.1863-1871.2005.

Keller G, Ray E, Brown PO, Winge DR. 2001. Haa1, a Protein Homologous to the Copper-regulated Transcription Factor Ace1, Is a Novel Transcriptional Activator. Journal of Biological Chemistry. 276(42):38697–38702. doi:10.1074/jbc.M107131200.

Keller NP. 2019. Fungal secondary metabolism: regulation, function and drug discovery. Nature Reviews Microbiology. 17(3):167–180. doi:10.1038/s41579-018-0121-1.

Keller NP, Dischinger HC, Bhatnagar D, Cleveland TE, Ullah AH. 1993. Purification of a 40-kilodalton methyltransferase active in the aflatoxin biosynthetic pathway. Applied and Environmental Microbiology. 59(2):479–484. doi:10.1128/aem.59.2.479-484.1993.

Kelly JM, Hynes MJ. 1985. Transformation of *Aspergillus niger* by the amdS gene of *Aspergillus* nidulans. The EMBO Journal. 4(2):475–479. doi:10.1002/j.1460-2075.1985.tb03653.x.

Kennedy J, Auclair K, Kendrew SG, Park C, Vederas JC, Richard Hutchinson C. 1999. Modulation of Polyketide Synthase Activity by Accessory Proteins During Lovastatin Biosynthesis. Science. 284(5418):1368–1372. doi:10.1126/science.284.5418.1368.

Kim K-W, Sugawara F, Yoshida S, Murofushi N, Takahashi N, Curtis RW. 1993. Structure of Malformin A, a Phytotoxic Metabolite Produced by *Aspergillus niger*. Bioscience, Biotechnology, and Biochemistry. 57(2):240–243. doi:10.1271/bbb.57.240.

Kirimura K, Kobayashi K, Ueda Y, Hattori T. 2016. Phenotypes of gene disruptants in relation to a putative mitochondrial malate–citrate shuttle protein in citric acid-producing *Aspergillus niger*. Bioscience, Biotechnology, and Biochemistry. 80(9):1737–1746. doi:10.1080/09168451.2016.1164583.

Kobbe B, Cushman M, Wogan GN, Demain AL. 1977. Production and antibacterial activity of malforming C, a toxic metabolite of *Aspergillus niger*. Applied and Environmental Microbiology. 33(4):996–997. doi:10.1128/aem.33.4.996-997.1977.

Kodaki T, Yamashita S. 1987. Yeast phosphatidylethanolamine methylation pathway. Cloning and characterization of two distinct methyltransferase genes. The Journal of biological chemistry. 262(32):15428–35.

Kodukula K, Arcuri, Maredith, Cutrone JQ, Hugill RM, Lowe SE, Pirnik DM, Shu Y-Z, Fernandes PB, Seethala R. 1995. BMS-192548, a Tetracyclic Binding Inhibitor of Neuropeptide Y Receptors, from *Aspergillus niger* WB2346. I. Taxonomy, Fermentation, Isolation and Biological Activity. The Journal of Antibiotics. 48(10):1055–1059. doi:10.7164/antibiotics.48.1055.

Kooistra R, Hooykaas PJJ, Steensma HY. 2004. Efficient gene targeting inKluyveromyces lactis. Yeast. 21(9):781–792. doi:10.1002/yea.1131.

Kubicek CP, Röhr M, Rehm HJ. 1985. Citric Acid Fermentation. Critical Reviews in Biotechnology. 3(4):331–373. doi:10.3109/07388558509150788.

Kwon MJ, Jørgensen TR, Nitsche BM, Arentshorst M, Park J, Ram AFJ, Meyer V. 2012. The transcriptomic fingerprint of glucoamylase over-expression in Aspergillus niger. BMC Genomics, 13:701. doi: 10.1186/1471-2164-13-701.

van Lanen JM SM. 1968. Process of producing glucamylase and an alcohol product. US patent 3,418,211. Hiram Walker & Sons, Inc., Peoria, IL

Ledford H. 2015. CRISPR, the disruptor. Nature. 522(7554):20–24. doi:10.1038/522020a.

Li Y, Chooi Y-H, Sheng Y, Valentine JS, Tang Y. 2011. Comparative Characterization of Fungal Anthracenone and Naphthacenedione Biosynthetic Pathways Reveals an α-Hydroxylation-Dependent Claisen-like Cyclization Catalyzed by a Dimanganese Thioesterase. Journal of the American Chemical Society. 133(39):15773–15785. doi:10.1021/ja206906d.

Machado AK, Morgan BA, Merrill GF. 1997. Thioredoxin Reductase-dependent Inhibition of MCB Cell Cycle Box Activity in Saccharomyces cerevisiae. Journal of Biological Chemistry. 272(27):17045–17054. doi:10.1074/jbc.272.27.17045.

Macheleidt J, Mattern DJ, Fischer J, Netzker T, Weber J, Schroeckh V, Valiante V, Brakhage AA. 2016. Regulation and Role of Fungal Secondary Metabolites. Annual Review of Genetics. 50(1):371–392. doi:10.1146/annurev-genet-120215-035203.

Macheleidt J, Scherlach K, Neuwirth T, Schmidt-Heck W, Straßburger M, Spraker J, Baccile JA, Schroeder FC, Keller NP, Hertweck C, *et al*. 2015. Transcriptome analysis of cyclic AMP-dependent protein kinase A-regulated genes reveals the production of the novel natural compound fumipyrrole by A spergillus fumigatus. Molecular Microbiology. 96(1):148–162. doi:10.1111/mmi.12926.

MacPherson S, Larochelle M, Turcotte B. 2006. A Fungal Family of Transcriptional Regulators: the Zinc Cluster Proteins. Microbiology and Molecular Biology Reviews. 70(3):583–604. doi:10.1128/MMBR.00015-06.

Malaj N, De Simone BC, Quartarolo AD, Russo N. 2013. Spectrophotometric study of the copigmentation of malvidin 3-O-glucoside with p-coumaric, vanillic and syringic acids. 4(15): 3614–3620. doi: 10.1016/j.foodchem.2013.06.017.

Mandal S, Moudgil M, Mandal SK. 2009. Rational drug design. European Journal of Pharmacology. 625(1–3):90–100. doi:10.1016/j.ejphar.2009.06.065.

Manzanares P, de Graaff LH, Visser J. 1998. Characterization of Galactosidases from *Aspergillus niger*: Purification of a Novel α-Galactosidase Activity. Enzyme and Microbial Technology. 22(5):383–390. doi:10.1016/S0141-0229(97)00207-X.

Mas-Moruno C, Rechenmacher F, Kessler H. 2010. Cilengitide: The First Anti-Angiogenic Small Molecule Drug Candidate. Design, Synthesis and Clinical Evaluation. Anti-Cancer Agents in Medicinal Chemistry. 10(10):753–768. doi:10.2174/187152010794728639.

Mattern IE, Unkles S, Kinghorn JR, Pouwels PH, van den Hondel CAMJJ. 1987. Transformation of *Aspergillus* oryzae using the A. *niger* pyrG gene. Molecular and General Genetics MGG. 210(3):460–461. doi:10.1007/BF00327197.

Mayr LM and Bojanic D. 2009. Novel trends in high-throughput screening. Current Opinion in Pharmacology. 9(5):580–588. doi:10.1016/j.coph.2009.08.004.

McDonnell E, Strasser K, Tsang A. 2018. Manual Gene Curation and Functional Annotation. p. 185–208.

Meyer V, Arentshorst M, El-Ghezal A, Drews A-C, Kooistra R, van den Hondel CAMJJ, Ram AFJ. 2006. Highly efficient gene targeting in the *Aspergillus niger* kusA mutant. Journal of Biotechnology. 128(4):770–775. doi:10.1016/j.jbiotec.2006.12.021.

Meyer V, Wanka F, van Gent J, Arentshorst M, van den Hondel CAMJJ, Ram AFJ. 2011. Fungal Gene Expression on Demand: an Inducible, Tunable, and Metabolism-Independent Expression System for *Aspergillus niger*. Applied and Environmental Microbiology. 77(9):2975–2983. doi:10.1128/AEM.02740-10.

Microorganisms & Microbial-Derived Ingredients Used in Food (Partial List) [internet]. 2018. US FDA; cited 2022 August 16, Available from https://www.fda.gov/food/generally-recognized-safe-gras/microorganisms-microbial-derived-ingredients-used-food-partial-list

Minetoki T, Nunokawa Y, Gomi K, Kitamoto K, Kumagai C, Tamura G. 1996. Deletion analysis of promoter elements of the *Aspergillus* oryzae agdA gene encoding α -glucosidase. Current Genetics. 30(5):432–438. doi:10.1007/s002940050153.

Morton LW, Caccetta RA-A, Puddey IB, Croft KD. 2000. Chemistry And Biological Effects Of Dietary Phenolic Compounds: Relevance To Cardiovascular Disease. Clinical and Experimental Pharmacology and Physiology. 27(3):152–159. doi:10.1046/j.1440-1681.2000.03214.x.

Mouafi FE, Ibrahim GS, Abo Elsoud MM. 2016. Optimization of lovastatin production from *Aspergillus* fumigatus. Journal of Genetic Engineering and Biotechnology. 14(2):253–259. doi:10.1016/j.jgeb.2016.10.006.

Mulder HJ, Saloheimo M, Penttilä M, Madrid SM. 2004. The transcription factor HACA mediates the unfolded protein response in *Aspergillus niger*, and up-regulates its own transcription. Molecular Genetics and Genomics. 271(2):130–140. doi:10.1007/s00438-003-0965-5.

Nguyen KT, Ritz D, Gu J-Q, Alexander D, Chu M, Miao V, Brian P, Baltz RH. 2006. Combinatorial biosynthesis of novel antibiotics related to daptomycin. Proceedings of the National Academy of Sciences. 103(46):17462–17467. doi:10.1073/pnas.0608589103.

Nie X, Li B, Wang S. 2018. Epigenetic and Posttranslational Modifications in Regulating the Biology of *Aspergillus* Species. 105:191-226. doi: 10.1016/bs.aambs.2018.05.004.

Nielsen KF, Mogensen JM, Johansen M, Larsen TO, Frisvad JC. 2009. Review of secondary metabolites and mycotoxins from the *Aspergillus niger* group. Analytical and Bioanalytical Chemistry. 395(5):1225–1242. doi:10.1007/s00216-009-3081-5.

Ninomiya Y, Suzuki K, Ishii C, Inoue H. 2004. Highly efficient gene replacements in Neurospora strains deficient for nonhomologous end-joining. Proceedings of the National Academy of Sciences. 101(33):12248–12253. doi:10.1073/pnas.0402780101.

Nischitha1 R and Shivanna MB. 2022. Screening of secondary metabolites and antioxidant potential of endophytic fungus Penicillium citrinum and host Digitaria bicornis by spectrophotometric and electrochemical methods. Archives of Microbiology. 204:206. doi: 10.1007/s00203-022-02795-z.

Noonim P, Mahakarnchanakul W, Nielsen KF, Frisvad JC, Samson RA. 2009. Fumonisin B 2 production by *Aspergillus niger* in Thai coffee beans. Food Additives & Contaminants: Part A. 26(1):94–100. doi:10.1080/02652030802366090.

Oakley CE, Ahuja M, Sun W-W, Entwistle R, Akashi T, Yaegashi J, Guo C-J, Cerqueira GC, Russo Wortman J, Wang CCC, *et al*. 2017. Discovery of McrA, a master regulator of *Aspergillus* secondary metabolism. Molecular Microbiology. 103(2):347–365. doi:10.1111/mmi.13562.

Ōmura S, Miyadera H, Ui H, Shiomi K, Yamaguchi Y, Masuma R, Nagamitsu T, Takano D, Sunazuka T, Harder A, *et al*. 2001. An anthelmintic compound, nafuredin, shows selective inhibition of complex I in helminth mitochondria. Proceedings of the National Academy of Sciences. 98(1):60–62. doi:10.1073/pnas.98.1.60.

Osman Y, Gebreil A, Mowafy AM, Anan TI, Hamed SM. 2019. Characterization of *Aspergillus niger* siderophore that mediates bioleaching of rare earth elements from phosphorites. World Journal of Microbiology and Biotechnology. 35(6):93. doi:10.1007/s11274-019-2666-1.

Padhi S, Masi M, Panda SK, Luyten W, Cimmino A, Tayung K, Evidente A. 2020. Antimicrobial secondary metabolites of an endolichenic *Aspergillus niger* isolated from lichen thallus of Parmotrema ravum. Natural Product Research. 34(18):2573–2580. doi:10.1080/14786419.2018.1544982.

Palys S. 2017. A bioinformatics characterization of secondary metabolism and alkyl citric acid pathway reconstruction in *Aspergillus niger* NRRL3. Masters Thesis, Concordia University, Spectrum ID code 982550.

Papagianni M. 2007. Advances in citric acid fermentation by *Aspergillus niger*: Biochemical aspects, membrane transport and modeling. Biotechnology Advances. 25(3):244–263. doi:10.1016/j.biotechadv.2007.01.002.

Pel HJ, de Winde JH, Archer DB, Dyer PS, Hofmann G, Schaap PJ, Turner G, de Vries RP, Albang R, Albermann K, *et al*. 2007. Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. Nature Biotechnology. 25(2):221–231. doi:10.1038/nbt1282.

Perlman D, Kita DA, Peterson WH. 1946. Production of citric acid from cane molasses. Archives of biochemistry. 11:123–9.

Perrone G, Susca A, Cozzi G, Ehrlich K, Varga J, Frisvad JC, Meijer M, Noonim P, Mahakarnchanakul W, Samson RA. 2007. Biodiversity of *Aspergillus* species in some important agricultural products. Studies in Mycology. 59:53–66. doi:10.3114/sim.2007.59.07.

Petersen LM, Holm DK, Knudsen PB, Nielsen KF, Gotfredsen CH, Mortensen UH, Larsen TO. 2015. Characterization of four new antifungal yanuthones from *Aspergillus niger*. The Journal of Antibiotics. 68(3):201–205. doi:10.1038/ja.2014.130.

Pöggeler S, Kück U. 2006. Highly efficient generation of signal transduction knockout mutants using a fungal strain deficient in the mammalian ku70 ortholog. Gene. 378:1–10. doi:10.1016/j.gene.2006.03.020.

Pfam: The protein families database in 2021: J. Mistry, S. Chuguransky, L. Williams, M. Qureshi, G.A. Salazar, E.L.L. Sonnhammer, S.C.E. Tosatto, L. Paladin, S. Raj, L.J. Richardson, R.D. Finn, A. Bateman. Punt PJ, Oliver RP, Dingemanse MA, Pouwels PH, van den Hondel CA. 1987. Transformation of *Aspergillus* based on the hygromycin B resistance marker from Escherichia coli. Gene. 56(1):117–124. doi:10.1016/0378-1119(87)90164-8.

Price, MN, Dehal, PS, Arkin, AP. 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PloS ONE. 5(3):e9490. doi:10.1371/journal.pone.0009490.

Price, MN, Dehal, PS, Arkin, AP. 2009. FastTree: Computing Large Minimum-Evolution Trees with Profiles instead of a Distance Matrix. Molecular Biology and Evolution. 26:1641-1650. doi:10.1093/molbev/msp077.

Raistrick H and Clark AB. 1919. On the Mechanism of Oxalic Acid Formation by *Aspergillus niger*. Biochemical Journal. 13(4):329–344. doi:10.1042/bj0130329.

Rao KCS, Karanth NG, Sattur AP. 2005. Production of *niger*loxin, an enzyme inhibitor and a free radical scavenger, by *Aspergillus niger* using solid state fermentation. Process Biochemistry. 40(7):2517–2522. doi:10.1016/j.procbio.2004.10.008.

Rao P, Shukla A, Parmar P, Rawal RM, Patel B V., Saraf M, Goswami D. 2022. Proposing a fungal metabolite-flaviolin as a potential inhibitor of 3CL pro of novel coronavirus SARS-CoV-2 identified using docking and molecular dynamics. Journal of Biomolecular Structure and Dynamics. 40(1):348–360. doi:10.1080/07391102.2020.1813202.

Rambaut, A. 2010. FigTree v1.4.1. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh. Cited 2020 August 16. http://tree.bio.ed.ac.uk/software/figtree/.

Rateb ME, Hallyburton I, Houssen WE, Bull AT, Goodfellow M, Santhanam R, Jaspars M, Ebel R. 2013. Induction of diverse secondary metabolites in *Aspergillus* fumigatus by microbial co-culture. RSC Advances. 3(34):14444. doi:10.1039/c3ra42378f.

Rehman S, Aslam H, Ahmad A, Khan SA, Sohail M. 2014. Production of plant cell wall degrading enzymes by monoculture and co-culture of *Aspergillus niger* and *Aspergillus* terreus under SSF of banana peels. Brazilian Journal of Microbiology. 45(4):1485–1492. doi:10.1590/S1517-83822014000400045.

Riko R, Nakamura H, Shindo K. 2014. Studies on pyranonigrins–isolation of pyranonigrin E and biosynthetic studies on pyranonigrin A. The Journal of Antibiotics. 67(2):179–181. doi:10.1038/ja.2013.91.

Ruiz B, Chávez A, Forero A, García-Huante Y, Romero A, Sánchez M, Rocha D, Sánchez B, Rodríguez-Sanoja R, Sánchez S, *et al*. 2010. Production of microbial secondary metabolites: Regulation by the carbon source. Critical Reviews in Microbiology. 36(2):146–167. doi:10.3109/10408410903489576.

Sakurai M, Kohno J, Yamamoto K, Okuda T, Nishio M, Kawano K, Ohnuki T. 2002. TMC-256A1 and C1, New Inhibitors of IL-4 Signal Transduction Produced by *Aspergillus niger* var *niger* TC 1629. The Journal of Antibiotics. 55(8):685–692. doi:10.7164/antibiotics.55.685.

Sander JD and Joung JK. 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. Nature Biotechnology. 32(4):347–355. doi:10.1038/nbt.2842.

Schuster, E, Dunn-Coleman, D, Frisvad, JC, Van Dijck P. 2002. On the safety of *Aspergillus niger* - a review. Applied Microbiology and Biotechnology. 59(4–5):426–435. doi:10.1007/s00253-002-1032-6.

Segal BH. 2009. Aspergillosis. New England Journal of Medicine. 360(18):1870–1884. doi:10.1056/NEJMra0808853.

Semova N, Storms R, John T, Gaudet P, Ulycznyj P, Min XJ, Sun J, Butler G, Tsang A. 2006. Generation, annotation, and analysis of an extensive *Aspergillus niger* EST collection. BMC Microbiology. 6(1):7. doi:10.1186/1471-2180-6-7.

Shelest E. 2008. Transcription factors in fungi. FEMS Microbiology Letters. 286(2):145–151. doi:10.1111/j.1574-6968.2008.01293.x.

Shen L, Ye Y-H, Wang X-T, Zhu H-L, Xu C, Song Y-C, Li H, Tan R-X. 2006. Structure and Total Synthesis of Asper*niger*in: A Novel Cytotoxic Endophyte Metabolite. Chemistry - A European Journal. 12(16):4393–4396. doi:10.1002/chem.200501423.

Shu Y-Z, Cutrone JQ, Klohr SE, Huang S. 1995. BMS-192548, a Tetracyclic Binding Inhibitor of Neuropeptide Y Receptors, from *Aspergillus niger* WB2346. II. Physico-chemical Properties and Structural Characterization. The Journal of Antibiotics. 48(10):1060–1065. doi:10.7164/antibiotics.48.1060.

Šimonovičová A, Vojtková H, Nosalj S, Piecková E, Švehláková H, Kraková L, Drahovská H, Stalmachová B, Kučová K, Pangallo D. 2021. *Aspergillus niger* Environmental Isolates and Their Specific Diversity Through Metabolite Profiling. Frontiers in Microbiology. 12. doi:10.3389/fmicb.2021.658010.

Skellam E. 2019. Strategies for Engineering Natural Product Biosynthesis in Fungi. Trends in Biotechnology. 37(4):416–427. doi:10.1016/j.tibtech.2018.09.003.

Song L, Ouedraogo J-P, Kolbusz M, Nguyen TTM, Tsang A. 2018. Efficient genome editing using tRNA promoter-driven CRISPR/Cas9 gRNA in *Aspergillus niger*. Han K-H, editor. PLOS ONE. 13(8):e0202868. doi:10.1371/journal.pone.0202868.

Song L, Wu S, Tsang A. 2018. Phylogenetic Analysis of Protein Family. 1775:267-275. doi: 10.1007/978-1-4939-7804-5_21.

Stajich JE, Block D, Boulez K, Brenner SE, Chervitz SA, Dagdigian C, Fuellen G, Gilbert JG, Korf I, Lapp H, *et al*. 2002. The Bioperl toolkit: Perl modules for the life sciences. Genome Res. 12(10):1611-8. doi:http://dx.doi.org/10.1101/gr.361602.

Stevenson PC, Nicolson SW, Wright GA. 2017. Plant secondary metabolites in nectar: impacts on pollinators and ecological functions. Manson J, editor. Functional Ecology. 31(1):65–75. doi:10.1111/1365-2435.12761.

Storms R, Zheng Y, Li H, Sillaots S, Martinez-Perez A, Tsang A. 2005. Plasmid vectors for protein production, gene expression and molecular manipulations in *Aspergillus niger*. Plasmid. 53(3):191–204. doi:10.1016/j.plasmid.2004.10.001.

Sugawara F, Kim K-W, Uzawa J, Yoshida S, Takahashi N, Curtis RW. 1990. Structure of malformin A2, reinvestigation of phytotoxic metabolites produced by *Aspergillus niger*. Tetrahedron Letters. 31(30):4337–4340. doi:10.1016/S0040-4039(00)97615-9.

Sun S and Fu J. 2018. Methyl-containing pharmaceuticals: Methylation in drug design. Bioorganic & Medicinal Chemistry Letters. 28(20):3283–3289. doi:10.1016/j.bmcl.2018.09.016.

Sun X, Zhu J, Bao L, Hu C, Jin C, Harris SD, Liu H, Li S. 2013. pyrG is required for maintaining stable cellular uracil level and normal sporulation pattern under excess uracil stress in *Aspergillus* nidulans. Science China Life Sciences. 56(5):467–475. doi:10.1007/s11427-013-4480-6.

Svensson B and Larsen K, G. A. 1986. Characterization of a glucoamylase G2 from *Aspergillus niger*. European Journal of Biochemistry. 154(3):497–502. doi:10.1111/j.1432-1033.1986.tb09425.x.

Svensson B, Svendsen TG, Svendsen I, Sakai T, Ottesen M. 1982. Characterization of two forms of glucoamylase from *Aspergillus niger*. Carlsberg Research Communications. 47(1):55–69. doi:10.1007/BF02907797.

Tamayo-Ramos JA, Barends S, de Lange D, de Jel A, Verhaert R, de Graaff L. 2013. Enhanced production of *Aspergillus niger* laccase-like multicopper oxidases through mRNA optimization of the glucoamylase expression system. Biotechnology and Bioengineering. 110(2):543–551. doi:10.1002/bit.24723.

Tang M, Zou Y, Yee D, Tang Y. 2018. Identification of the pyranonigrin A biosynthetic gene cluster by genome mining in Penicillium thymicola IBT 5891. AIChE Journal. 64(12):4182–4186. doi:10.1002/aic.16324.

Thorne N, Auld DS, Inglese J. 2010. Apparent activity in high-throughput screening: origins of compound-dependent assay interference. Current Opinion in Chemical Biology. 14(3):315–324. doi:10.1016/j.cbpa.2010.03.020.

Tobert JA. 2003. Lovastatin and beyond: the history of the HMG-CoA reductase inhibitors. Nature Reviews Drug Discovery. 2(7):517–526. doi:10.1038/nrd1112.

Tsang A, Butler G, Powlowski J, Panisko EA, Baker SE. 2009. Analytical and computational approaches to define the *Aspergillus niger* secretome. Fungal Genetics and Biology. 46(1):S153–S160. doi:10.1016/j.fgb.2008.07.014.

Tsukamoto S, Yoshida T, Hosono H, Ohta T, Yokosawa H. 2006. Hexylitaconic acid: A new inhibitor of p53–HDM2 interaction isolated from a marine-derived fungus, Arthrinium sp. Bioorganic & Medicinal Chemistry Letters. 16(1):69–71. doi:10.1016/j.bmcl.2005.09.052.

Turnbull IF, Smith DR, Sharp PJ, Cobon GS, Hynes MJ. 1990. Expression and secretion in *Aspergillus* nidulans and *Aspergillus niger* of a cell surface glycoprotein from the cattle tick, Boophilus microplus, by using the fungal amdS promoter system. Applied and Environmental Microbiology. 56(9):2847–2852. doi:10.1128/aem.56.9.2847-2852.1990.

Vardanyan RS and Hruby VJ. 2006. Analgesics. In: Synthesis of Essential Drugs. Elsevier. p. 19–55.

Vesth TC, Nybo JL, Theobald S, Frisvad JC, Larsen TO, Nielsen KF, Hoof JB, Brandl J, Salamov A, Riley R, *et al*. 2018. Investigation of inter- and intraspecies variation through genome sequencing of *Aspergillus* section Nigri. Nature Genetics. 50(12):1688–1695. doi:10.1038/s41588-018-0246-1.

de Vries R. 2003. Regulation of *Aspergillus* genes encoding plant cell wall polysaccharide-degrading enzymes; relevance for industrial production. Applied Microbiology and Biotechnology. 61(1):10–20. doi:10.1007/s00253-002-1171-9.

de Vries RP and Visser J. 1999. Regulation of the Feruloyl Esterase (faeA) Gene from *Aspergillus niger*. Applied and Environmental Microbiology. 65(12):5500–5503. doi:10.1128/AEM.65.12.5500-5503.1999.

Waksman SA and Reilly HC. 1944. Strain Specificity and Production of Antibiotic Substances. Proceedings of the National Academy of Sciences. 30(5):99–105. doi:10.1073/pnas.30.5.99.

Walisko R, Moench-Tegeder J, Blotenberg J, Wucherpfennig T, Krull R. 2015. The Taming of the Shrew - Controlling the Morphology of Filamentous Eukaryotic and Prokaryotic Microorganisms. p. 1–27.

Wang B, Li X, Tabudravu J, Wang S, Deng H, Pan L. 2021. The chemical profile of activated secondary metabolites by overexpressing LaeA in *Aspergillus niger*. Microbiological Research. 248:126735. doi:10.1016/j.micres.2021.126735.

Wanka F, Cairns T, Boecker S, Berens C, Happel A, Zheng X, Sun J, Krappmann S, Meyer V. 2016. Tet-on, or Tet-off, that is the question: Advanced conditional gene expression in *Aspergillus*. Fungal Genetics and Biology. 89:72–83. doi:10.1016/j.fgb.2015.11.003.

Waterhouse, AM, Procter, JB, Martin, DMA, Clamp, M, Barton, GJ (2009). Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. Bioinformatics. 25 (9) 1189-1191 doi: 10.1093/bioinformatics/btp033.

Weissman KJ and Leadlay PF. 2005. Combinatorial biosynthesis of reduced polyketides. Nature Reviews Microbiology. 3(12):925–936. doi:10.1038/nrmicro1287.

Wells PA and Herrick HT. 1938. Citric Acid Industry. Industrial & Engineering Chemistry. 30(3):255–262. doi:10.1021/ie50339a004.

Wernars K, Goosen T, Wennekes BMJ, Swart K, van den Hondel CAMJJ, van den Broek HWJ. 1987. Cotransformation of *Aspergillus* nidulans: a tool for replacing fungal genes. Molecular and General Genetics MGG. 209(1):71–77. doi:10.1007/BF00329838.

Williams RB, Henrikson JC, Hoover AR, Lee AE, Cichewicz RH. 2008. Epigenetic remodeling of the fungal secondary metabolome. Organic & Biomolecular Chemistry. 6(11):1895. doi:10.1039/b804701d.

Wu X, Wu J, Luo Y, Bragg J, Anderson O, Vogel J, Gu YQ. 2013. Phylogenetic, Molecular, and Biochemical Characterization of Caffeic Acid o-Methyltransferase Gene Family in Brachypodium distachyon. Int J Plant Genomics. 2013:423189. doi: 10.1155/2013/423189.

Xu Y, Shan L, Zhou Y, Xie Z, Ball AS, Cao W, Liu H. 2019. Development of a Cre-loxP-based genetic system in *Aspergillus niger* ATCC1015 and its application to construction of efficient organic acid-producing cell factories. Applied Microbiology and Biotechnology. 103(19):8105–8114. doi:10.1007/s00253-019-10054-3.

Yang L, Henriksen MM, Hansen RS, Lübeck M, Vang J, Andersen JE, Bille S, Lübeck PS. 2020. Metabolic engineering of *Aspergillus niger* via ribonucleoprotein-based CRISPR–Cas9 system for succinic acid production from renewable biomass. Biotechnology for Biofuels. 13(1):206. doi:10.1186/s13068-020-01850-5.

Yoshizawa T, Tsuchiya Y, Morooka N, Sawada Y. 1975. Malformin A 1 as a Mammalian Toxicant from *Aspergillus niger*. Agricultural and Biological Chemistry. 39(6):1325–1326. doi:10.1080/00021369.1975.10861776.

Zabala AO, Xu W, Chooi Y-H, Tang Y. 2012. Characterization of a Silent Azaphilone Gene Cluster from *Aspergillus niger* ATCC 1015 Reveals a Hydroxylation-Mediated Pyran-Ring Formation. Chemistry & Biology. 19(8):1049–1059. doi:10.1016/j.chembiol.2012.07.004.

Zhang W and Tang Y. 2008. Combinatorial Biosynthesis of Natural Products. Journal of Medicinal Chemistry. 51(9):2629–2633. doi:10.1021/jm701269v.

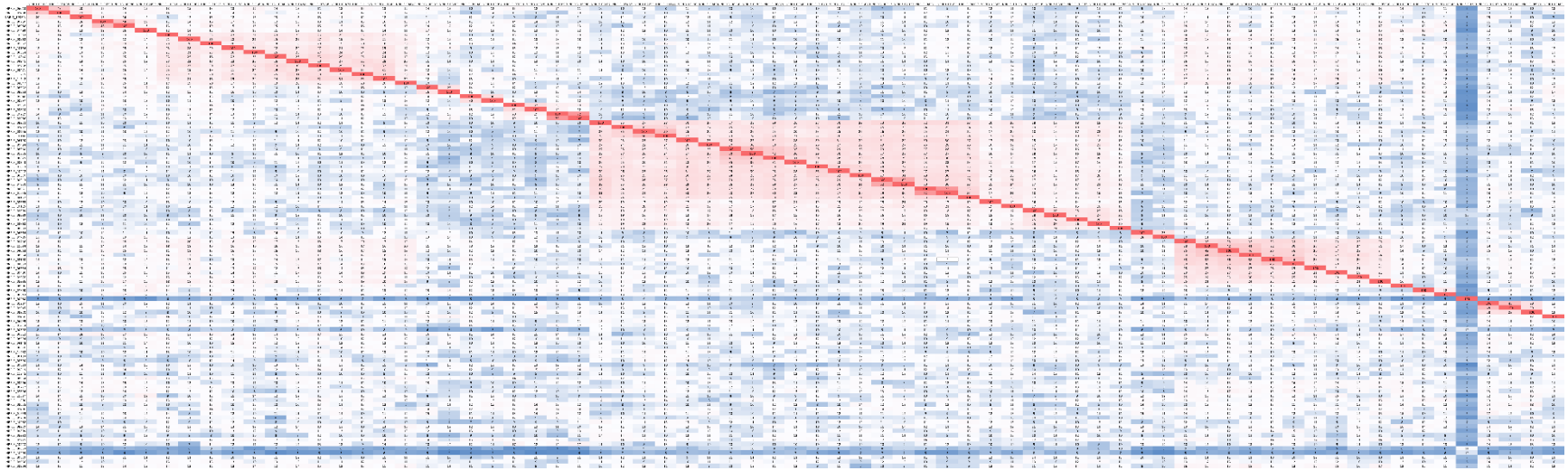
Zheng X, Zheng P, Sun J, Kun Z, Ma Y. 2018. Heterologous and endogenous U6 snRNA promoters enable CRISPR/Cas9 mediated genome editing in *Aspergillus niger*. Fungal Biology and Biotechnology. 5(1):2. doi:10.1186/s40694-018-0047-4.

Zhong J-J. 2002. Plant cell culture for production of paclitaxel and other taxanes. Journal of Bioscience and Bioengineering. 94(6):591–599. doi:10.1016/S1389-1723(02)80200-6.

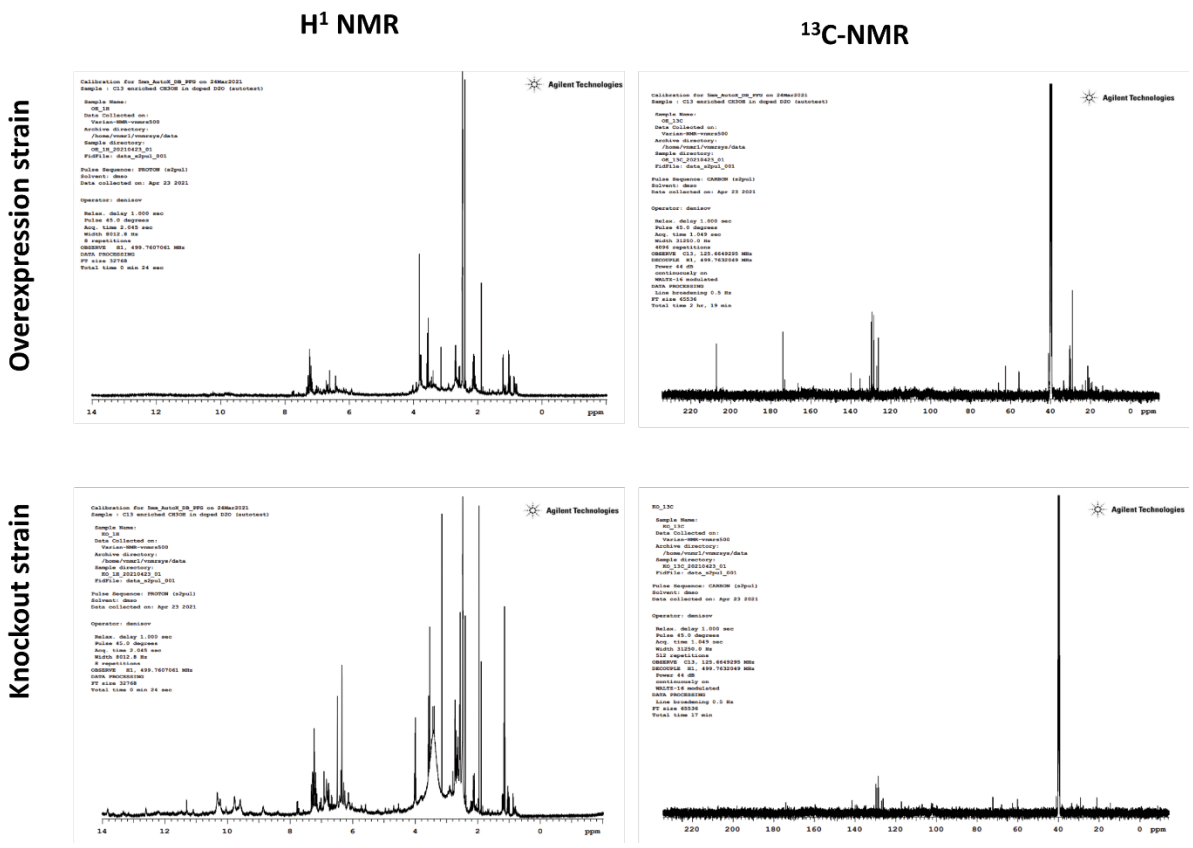
Ziegler J, Facchini PJ, Geißler R, Schmidt J, Ammer C, Kramell R, Voigtländer S, Gesell A, Pienkny S, Brandt W. 2009. Evolution of morphine biosynthesis in opium poppy. Phytochemistry. 70(15–16):1696–1707. doi:10.1016/j.phytochem.2009.07.006.

7. Appendices

Appendix 1: Intraspecies NRRL3 methyltransferase gene pairwise comparison matrix



Appendix 2: ^{1}H and ^{13}C NMR analysis of TAN/BMS strains.



Appendix 3: TAN/BMS extracellular protein production and statistical analysis

Methylated versus non-methylated TAN/BMS production in NRRL3_09545^{OE} (t-test (α = 0.05))

Two-tailed t-test	Mean Diff.	95.00% CI of diff.	Summary	Adjusted P Value
NRRL3_09545^{OE} Non-methylated vs. NRRL3_09545^{OE} Methylated	-11123000 ± 1490960	-14283695 to -7962305	****	p < 0.0001

*Methylated versus non-methylated TAN/BMS production in NRRL3_09545^{OE}09546^{KO}*glaA*::09546 (t-test (α = 0.05))*

Two-tailed t-test	Mean Diff.	95.00% CI of diff.	Summary	Adjusted P Value
*gla*A::09645 Non-methylated 1 vs *gla*A::09645 Methylated 1	12360000 ± 1393760	8490302 to 16229698	***	p < 0.005
*gla*A::09645 Non-methylated 2 vs *gla*A::09645 Methylated 2	20493333 ± 1991326	14964527 to 26022140	****	p < 0.005
*gla*A::09645 Non-methylated 3 vs *gla*A::09645 Methylated 3	28390000 ± 1048009	25480261 to 31299739	****	p < 0.0001
*gla*A::09645 Non-methylated 4 vs *gla*A::09645 Methylated 4	19500000 ± 1306395	15872867 to 23127133	****	p < 0.0001
*gla*A::09645 Non-methylated 5 vs *gla*A::09645 Methylated 5	26680000 ± 1504645	22502437 to 30857563	****	p < 0.0001
*gla*A::09645 Non-methylated 6 vs *gla*A::09645 Methylated 6	13696667 ± 4004208	2579204 to 24814129	*	p < 0.05

Appendix 4: TAN/BMS intracellular protein production and statistical analysis

1-way ANOVA (α = 0.05)

Gene	Dunnett's multiple comparisons test	Mean Diff.	95.00% CI of diff.	Summary	Adjusted P Value	Comparison
NRRL3_09545	Wild type vs. Overexpression	-2.34E+08	-1022627512 to 555538915	ns	0.9305	Overexpression
	Wild type vs. Knockout	-7.56E+07	-864689754 to 713476673	ns	0.9996	Knockout
	Wild type vs. *gla*A::09546	-2.14E+08	-1003519116 to 574647312	ns	0.9529	*gla*A::09546
	Wild type vs. *gla*A::00997	-1.87E+08	-976008466 to 602157962	ns	0.9761	*gla*A::00997
	Wild type vs. *gla*A::07487	-1.55E+08	-944310405 to 633856022	ns	0.9916	*gla*A::07487
	Wild type vs. *gla*A::10125	-9.80E+07	-887040031 to 691126397	ns	0.9995	*gla*A::10125
	Wild type vs. *gla*A::10369.1	-3.50E+07	-824085317 to 754081111	ns	0.9998	*gla*A::10369.1
NRRL3_09546	Wild type vs. Overexpression	-1.72E+10	-26363941011 to -8042411067	***	0.0002	Overexpression
	Wild type vs. Knockout	-1.04E+08	-9264764972 to 9056764972	ns	>0.9999	Knockout
	Wild type vs. *gla*A::09546	-3.04E+08	-9464289576 to 8857240369	ns	0.9999	*gla*A::09546
	Wild type vs. *gla*A::00997	1.65E+05	-9160599764 to 9160930181	ns	>0.9999	*gla*A::00997
	Wild type vs. *gla*A::07487	7.08E+07	-9089995605 to 9231534340	ns	>0.9999	*gla*A::07487
	Wild type vs. *gla*A::10125	6.39E+07	-9096872593 to 9224657351	ns	>0.9999	*gla*A::10125
	Wild type vs. *gla*A::10369.1	5.06E+07	-9844164510 to 9945349618	ns	>0.9999	*gla*A::10369.1
NRRL3_09547	Wild type vs. Overexpression	-1.15E+10	-23774081234 to 712973831	ns	0.0708	Overexpression
	Wild type vs. Knockout	-6.31E+09	-18553527533 to 5933527533	ns	0.5559	Knockout
	Wild type vs. *gla*A::09546	-7.00E+09	-19243550234 to 5243504831	ns	0.4514	*gla*A*::09546*
	Wild type vs. *gla*A::00997	-5.60E+09	-17847298591 to 6639756475	ns	0.6687	*gla*A::00997
	Wild type vs. *gla*A::07487	-3.61E+09	-15858316451 to 8628738615	ns	0.9312	*gla*A::07487
	Wild type vs. *gla*A::10125	-7.65E+09	-19892891641 to 4594163424	ns	0.3631	*gla*A::10125
	Wild type vs. *gla*A::10369.1	-4.73E+09	-16970623433 to 7516431632	ns	0.8028	*gla*A::10369.1

Gene	Dunnett's multiple comparisons test	Mean Diff.	95.00% CI of diff.	Summary	Adjusted P Value	Comparison
NRRL3_09548	Wild type vs. Overexpression	-1.48E+09	-5415695743 to 2447957474	ns	0.8177	Overexpression
	Wild type vs. Knockout	-1.01E+09	-4943195743 to 2920457474	ns	0.9639	Knockout
	Wild type vs. *gla*A::09546	1.94E+08	-3737734920 to 4125918296	ns	0.9998	*gla*A::09546
	Wild type vs. *gla*A::00997	-2.00E+09	-5934553289 to 1929099927	ns	0.5675	*gla*A::00997
	Wild type vs. *gla*A::07487	-1.28E+09	-5207408730 to 2656244487	ns	0.8973	*gla*A::07487
	Wild type vs. *gla*A::10125	-3.34E+08	-4266231867 to 3597421350	ns	0.9997	*gla*A::10125
	Wild type vs. *gla*A::10369.1	3.71E+08	-3560907022 to 4302746195	ns	0.9996	*gla*A::10369.1
NRRL3_09549	Wild type vs. Overexpression	-5.41E+10	-93357358489 to -14756141511	**	0.0043	Overexpression
	Wild type vs. Knockout	-3.70E+10	-76304858489 to 2296358489	ns	0.0709	Knockout
	Wild type vs. *gla*A::09546	-3.14E+10	-70704061386 to 7897155591	ns	0.1571	*gla*A::09546
	Wild type vs. *gla*A::00997	-3.21E+10	-71437922488 to 7163294489	ns	0.1422	*gla*A::00997
	Wild type vs. *gla*A::07487	-2.70E+10	-66301572088 to 12299644889	ns	0.2745	*gla*A::07487
	Wild type vs. *gla*A::10125	-2.85E+10	-67802418086 to 10798798891	ns	0.2286	*gla*A::10125
	Wild type vs. *gla*A::10369.1	-2.27E+10	-61994441837 to 16606775140	ns	0.4414	*gla*A::10369.1
NRRL3_09550	Wild type vs. Overexpression	-6.56E+08	-1516759967 to 204632380	ns	0.19	Overexpression
	Wild type vs. Knockout	-6.83E+08	-1543382712 to 178009635	ns	0.1619	Knockout
	Wild type vs. *gla*A::09546	-5.14E+08	-1374809752 to 346582595	ns	0.4075	*gla*A::09546
	Wild type vs. *gla*A::00997	-3.80E+08	-1240833323 to 480559024	ns	0.6999	*gla*A::00997
	Wild type vs. *gla*A::07487	-3.65E+08	-1225418779 to 495973568	ns	0.7342	*gla*A::07487
	Wild type vs. *gla*A::10125	-5.94E+08	-1454878110 to 266514236	ns	0.2703	*gla*A::10125
	Wild type vs. *gla*A::10369.1	-1.97E+08	-1057762103 to 663630244	ns	0.98	*gla*A::10369.1