TECHNISCHE UNIVERSITÄT WIEN

D I P L O M A R B E I T

Transcription factor engineering for the activation of biosynthetic gene clusters

---

Ausgeführt am Institut für

Verfahrenstechnik, Umweltechnik und technische Biowissenschaften E166

der Technischen Universität Wien

unter der Anleitung von

Assistant Prof. Dr.techn. Dipl.-Ing. Astrid Mach-Aigner

Senior Scientist Dr.rer.nat. Mag.rer.nat. Christian Derntl

durch

Fourtis Lukas, BSc.

19.11.2021

Datum

Unterschrift

Abstract

Fungal secondary metabolites are small bioactive compounds that are encoded in biosynthetic gene clusters in the fungal genome. These gene clusters are a source of undiscovered treasures, considering the bioactivities of the resulting metabolite. The metabolites from these clusters are mainly used as treatment in medical applications but also for food- and textile coloring in our modern world. Many of these clusters often remain silent under standard laboratory conditions and the organism only activates them when it needs to perform a certain task. Hence the metabolites often remain undiscovered. A major task in modern biotechnology is therefore to activate these silent gene clusters, identify the resulting secondary metabolite or derivate and utilize them.

In this project, bioinformatic analyses on *Trichoderma reesei* QM6a's genome was performed and 91 biosynthetic gene cluster were discovered. Four were chosen for activation according to their organization and potential metabolite. Five synthetic transcription factor cassettes (syn) were designed, bearing a truncated DNA binding domain fused to the transactivating domain of another gene. Classic cloning approaches paved the way to transform these cassettes into *T. reesei* and additionally *T. reesei* strains were generated with overexpression cassettes (OE), which inherited the native transcription factor. After cultivation on glucose as sole carbon source, transcript analyses were performed via qPCR by targeting key core biosynthetic genes of each cluster. We were able to activate three out of four BGCs: One cluster was activated in both types of transformants (OE and syn), one cluster only in syn-transformants and the third cluster only in OE-transformants. Furthermore, we determined co-expression behavior of predicted genes in these three activated clusters. All analyzed core biosynthetic genes were highly co-expressed in the activated clusters.

Future steps to come consist of identifying the produced secondary metabolites, isolate them and test them for their biochemical properties.

Zusammenfassung

Sekundärmetabolite aus Pilzen sind bioaktive Verbindungen, die in biosynthetischen Genclustern im Pilzgenom kodiert sind. Diese Gencluster stellen eine Quelle versteckter Schätze dar, wenn man die Vielfalt der Bioaktivitäten der entstehenden Metaboliten bedenkt. Die aus diesen Clustern stammenden Metabolite werden in unserer modernen Welt hauptsächlich in der Medizin, aber auch zur Färbung von Lebensmitteln und Textilien verwendet. Viele dieser Cluster bleiben unter Standardlaborbedingungen oft stumm und werden nur aktiviert, wenn der Pilz sie braucht. Der Organismus aktiviert sie nur, wenn er eine bestimmte Aufgabe zu erfüllen hat, daher bleiben die Metaboliten oft unentdeckt. Eine wichtige Aufgabe der modernen Biotechnologie besteht daher darin, diese stillen Gencluster zu aktivieren und die daraus entstehenden Sekundärmetaboliten oder Derivate zu identifizieren und sie zu nutzen.

In diesem Projekt wurden bioinformatische Analysen am Genom von *Trichoderma reesei* QM6a durchgeführt und 91 biosynthetische Gencluster entdeckt. Vier davon wurden aufgrund ihres Aufbaus und ihres potenziellen Metaboliten zur Aktivierung ausgewählt. Es wurden fünf synthetische Transkriptionsfaktorkassetten (syn) entworfen, die eine verkürzte DNA-Bindungsdomäne enthalten, welche mit der transaktivierenden Domäne eines anderen Gens fusioniert ist. Klassische Klonierungsansätze ebneten den Weg zur Transformation dieser Kassetten in *T. reesei* und zusätzlich wurden *T. reesei* Stämme mit Überexpressionskassetten (OE) transformiert, die den nativen Transkriptionsfaktor enthielten. Nach der Kultivierung mit Glukose als einziger Kohlenstoffquelle wurden Transkriptionsanalysen mittels qPCR durchgeführt und die wichtigsten biosynthetischen Schlüsselgene jedes Clusters analysiert. Es gelang uns drei von vier BGCs zu aktivieren: Ein Cluster wurde in beiden Arten von Transformanten (OE und syn) aktiviert, ein Cluster nur in syn-Transformanten und der dritte Cluster nur in OE-Transformanten. Darüber hinaus haben wir das Co-expressionsverhalten von bestimmten Genen in diesen drei aktivierten Clustern analysiert. Alle, als wichtig für die Cluster vorhergesagten Gene, die analysiert wurden waren stark co-exprimiert.

Zukünftige Aufgaben bestehen darin, die produzierten Sekundärmetaboliten zu identifizieren, zu isolieren und auf ihre biochemischen Eigenschaften zu testen.

1. Introduction

1.1. Secondary metabolites – pioneers of a modern world

Secondary metabolism is coarsely defined as metabolism that is not necessary for growth, maintaining or reproduction of the organism[1]. Its counterpart is primary metabolism which fulfills the formerly mentioned vital tasks in each living organism. Both have in common that they use the same substrates (e.g. acetyl-CoA or amino acids) but different enzymes[1]. While primary metabolism enzymes are ubiquitous and highly conserved across nearly all kingdoms of life[2–4], secondary metabolism enzymes remain specialized in terms for task and host[5]. Furthermore, secondary metabolism has some unique characteristics: The resulting metabolites are normally only built under specific circumstances or to perform a certain task like defense[6,7], predatorious[8] or to gain the edge when competing for space and resources[9]. Secondary metabolites (SMs) are often small active molecules that are used by mankind for different purposes[10]. The most important application of SMs is their use in medicine as antibiotics. Prominent examples are penicillin[11], cephalosporin[12] or erythromycin[13]. Although there are countless substances and different classes of antibiotics, antibiotic resistance[14–16] is on the rise and poses a great threat to humanity. This is one of the main reasons to intensify research on novel substance classes and until this day hidden SMs. Other medical applications of SMs include chemotherapeutic drugs[17], immune suppressants[18] and anti-cancer drugs[18]. Industrially used SMs include food coloring pigments[22] and coloring of textiles[23]. Fungi are further used as biocontrol agents, thereby their SMs exhibit antifungal[19], antimicrobial[20] and cytotoxic[21] properties. Agriculturally, fungi are also applied to enhance the yield and growth of crops and to fend off parasites[18].

The focus of this project is centered around fungal SMs because fungi pose an enormous source of useful yet undiscovered SMs[24]. Fungal SMs can be grouped into three major groups, although being chemically unbelievably diverse and heterogenous: polyketides, non-ribosomal peptides and terpenes[24]. There are also hybrids of the former mentioned ones that display characteristics from either of these[24]. An overview of the three main classes of fungal SMs is given in Figure 1.

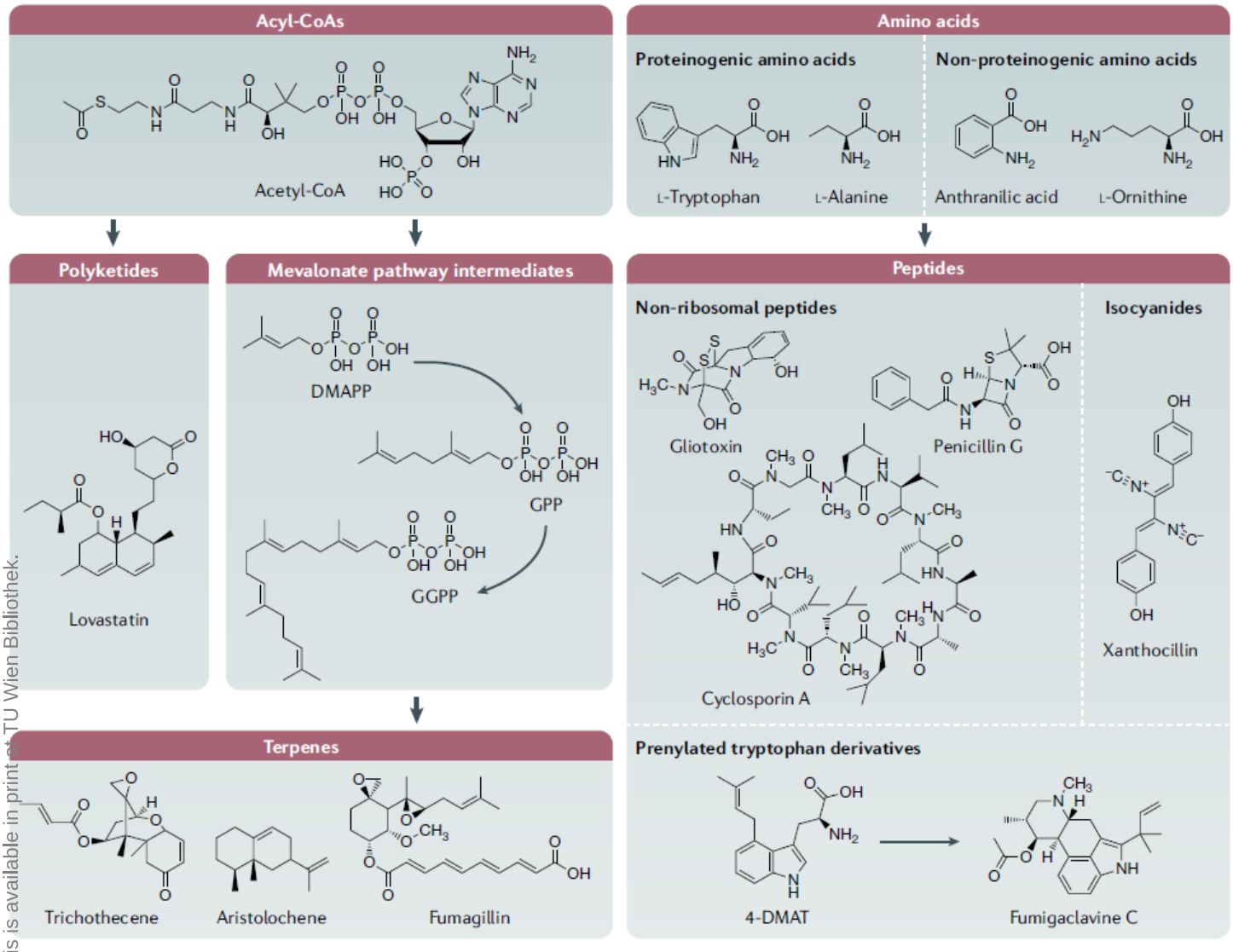

Figure 1. Overview of fungal secondary metabolites.
Polyketides as well as terpenes are derived from acyl-CoA building blocks whereas non-ribosomal peptides are derived from amino acids. DMAPP … dimethylallyl diphosphate; 4-DMAT … dimethylallyl tryptophan; GPP … geranyl diphosphate (C10) ; GGPP … geranylgeranyl diphosphate (C15); Adapted with permission from REF[24], Springer Nature Limited.

Polyketides got their name from their basal chemical structure consisting of multiple condensed ketones and are built from acetyl-CoA or malonyl-CoA as substrates by polyketide-synthases (PKSs) via condensation[24]. Generally, PKSs can be grouped in three main groups: type I, type II or type III[25]. Type I comprise PKSs where functional domains are covalently linked on one large poly-peptide chain[25]. Type I PKSs are typically organized in several modules of similar domains[25]. Type II refers to PKSs where some functional domains are not covalently linked but extra-standing in modules[25] and are mainly found in bacteria[26]. Type III are distinguishable from type I and II by their structural simplicity, acting generally as homodimers of identical keto-synthases[27] and being able to use a wider variety of different starter substrates[27,28]. The domains that most type I PKSs are minimally comprised of are a

keto-(acyl)synthase (K(A)S), a (malonyl)-acyltransferase ((M)AT) and an acyl-carrier protein (ACP)[25]. Furthermore the following domains can be part of PKSs: a ketoreductase (KR), a dehydratase (DH), a enoyl reductase (ER), a C-methyltransferase (CMeT) and a thioesterase (TE)[25]. Each domain asserts a specific function and is crucial for the building of the basal chemical scaffold. Thereby, the task of each domain is similar to their homologs in the fatty acid synthases (FASs)[29]. The KS domain catalyzes the condensation of two acyl substrates[25]. The (M)AT domain transfers either acetyl or malonyl-CoA to ACP[25]. KR reduces keto groups[25]. DH removes water thereby creating a double or triple bond[25]. ER reduces the former created double or triple bond and TE splits the product from the PKSs[25]. Some domains appear more than once in type I PKSs[30]. Type I PKSs can either act modular like in *Aspergillus terreus*'s biosynthesis of lovastatin[31] or iterative like in *Penicillium patulum*'s biosynthesis of 6-methylsaliclyic acid[32,33]. Thereby modular means that each module consists of several domains and acts only once per resulting product. Contrary to that, iterative means that each module acts more than once per resulting product, except for the AT domain[34]. Modular acting type 1 PKSs product backbones are mainly dependent on number and order of the modules, whereas it's tricky to unravel the backbone of products derived from iterative acting type 1 PKSs only from number and order of the modules[34].

Type II PKSs are mainly composed of the "minimal core subunits" which are the ACP and the KS, lacking an AT domain[27]. Other domains are either substrate or reaction-specific like in the case of actinorhodin biosynthesized by *Streptomyces coelicolor*[32,35]. This specific type II PKS inherits three reaction-specific subunits, a KR, a cyclase (CYC) and an aromatase (ARO)[32,35]. Type III PKSs were first found in plants[36] and believed to be exclusively expressed in them but recently discovered in fungi[37], bacteria[38] and actinobacteria[39]. Type III PKSs core structure are mainly homodimers with catalytic condensation domain KSs[28]. The various type III PKSs differ significantly from each other by their starting substrate and reaction mechanisms, leading to a wide variety of different products[28].

The second group of fungal SMs are synthesized by non-ribosomal peptide synthetases (NRPSs) via the condensation of amino acids to peptides. Generally, NRPSs can be grouped into three types[40]:

Type I NRPSs are organized on a single polypeptide chain in modules and are mainly found in fungi[41]. Thereby each module acts once per product and consist of several domains[40]. A prominent example for type I NRP is penicillin G biosynthesized by *Penicillium chrysogenium*[42]. Type II NRPSs are organized in several stand-alone proteins with several domains that interact more than once (iterative) in the building of the product[40]. Type

II NRPSs are mainly found in bacteria[43]. A prominent example for type II NRP is vancomycin biosynthesized by *Amycolatopsis orientalis*[44]. Vancomycin was applied in medicine as antibiotic, especially since the rise of methicillin-resistant *Staphylococcus aureus* (MRSA) and penicillin-resistant *Streptococcus pneumoniae*[45]. Type III NRPSs present themselves as mix of type I and type II NRPSs and inherit modules organized in one polypeptides as well as stand-alone units of other modules that act either modular or iterative[40]. An example for a type III NRP is vibriobactin biosynthesized by the bacterium *Vibrio cholerae*[46].

The basic units each NRPS modules inherits are a adenylation domain (A), a peptidyl carrier protein domain (PCP) and a condensation domain (C) with exception from the first module in each NRPS which lacks a C domain[40]. Generally, the PCP domain fixates the peptides that will be condensed by the C domain while the A domain recognize and activates the peptides priory by formation of a adenylate intermediate[40]. Furthermore, modules can contain an epimerase (E), a thioesterase (TE), a methyltransferase (MT) or oxidation (Ox) and cyclisation (Cyc) domain[40].

The third group of fungal SMs are synthesized by terpene synthases (TSs)[47]. Terpenes are natural products that were isolated from plants and other microbial organisms[47]. Fungal terpenoids are all derived from isopentenyl diphosphate (IPP) or its isomer dimethylallyl diphosphate (DMAPP)[47]. These building blocks are resulting from the mevalonate pathway, which requires acetyl-CoA as starter substrate and leads to the production of cholesterol[48]. Thereby, two coarse classes of terpenes are distinguished: Terpenes that are derived solely from isoprenyl units (class I) and those terpenes that are derived from mixed building blocks (class II), namely isoprenyl units and polyketide or indol precursor[47]. An example for class I terpenoids are mono-, sesqui-, di- or triterpenoids, namely trichothecenes biosynthesized by *Fusarium* and other fungal species[47]. This terpenoid is a mycotoxin which poses a health threat to mankind by poising cereal crops[47]. An example for class II terpenoids are meroterpenoids, namely pyripyropene A biosynthesized by *Aspergillus fumigatus*[49]. Thereby, an iterative type I PKS utilizes nicotinyl-CoA and condenses it with two malonyl-CoA units[49]. The following steps in the biosynthesis include the prenyltransfer by a prenyltransferase and epoxidation as well as cyclization by a terpene cyclase leading to pyripyropene A[50]. Pyripyropene A is a potential candidate for application as acyl-coenzyme A:cholesterol acyltransferase (ACAT) inhibitor, resulting in lowered cholesterol levels[51].

Although many metabolites in all three major classes of SMs are already discovered there is still many more to unravel. Especially when it comes to the organizational structure these metabolites originate from: biosynthetic gene clusters.

1.1.1. Biosynthetic gene clusters – immense dormant potential for novel substance classes

The term biosynthetic gene cluster (BGC) comprises several genes that are in a close distance to each other which play a part in the formation of a SM. Most of these clusters consist of at least one core biosynthetic gene that encodes either a PKS, a NRPS, a hybrid enzyme of former mentioned (PKS-NRPS) or a TS[24]. These enzymes execute the primary formation of the basal chemical scaffold of the corresponding SM[24]. They are supported by additional biosynthetic enzymes, often referred as auxiliary or tailoring enzymes, that modify the chemical scaffold[24]. Examples for additional biosynthetic enzymes are oxidoreductases, transferases, hydrolases, or epimerases and more[24]. Sometimes transport related genes are also of part these clusters[24].

BGCs are found in all kingdom of lives and were extensively studied in bacteria[52–54], fungi[55,56] and plants[57–59]. Thereby, the vast potential of silent BGCs was just recently unraveled when genomic data became widely available for public with the advent of the genomics era in the early 2000s. Prior to the prediction of BGCs based on genomic data, High-throughput screening of synthetic compound libraries[60] or fragment based design[61] was conducted. Drawbacks of these methods included the inefficiency of developed assays in vivo, the limited number of heterocyclic aromatic scaffolds in former mentioned libraries and the approaches being generally time-consuming, labor-intensive and expensive in costs[60-62]. These days the journey for the discovery of new SMs starts with a thorough *in silico* analysis of sequencing data of the organism of interest, with sophisticated bioinformatic tools. antiSMASH[63], PRISM[64], SMURF[65] and BAGEL3[66] are just few examples of today's publicly available tools.

Two different strategical approaches are applied for the activation of BGC: The first strategy can be summarized as pleiotropic approaches and the second as pathway-specific approaches. In pleiotropic approaches the main idea is that by modifying global influences like variation of culturing conditions[67], tampering with the transcriptional and translational machinery[54], overexpressing or deleting global known regulators[52] or epigenetic disturbances[68] the induction of a former silent BGCs can be achieved. Pathway-specific approaches comprise manipulating pathway-specific regulators[69], reporter-guided mutant selection[70], refactoring[71] or heterologous expression[72]. All the mentioned strategies have been applied successfully in different organisms to induce the expression of secondary metabolites.

In this project the focus is directed to *Trichoderma reesei*'s silent BGCs. In *Trichoderma* species generally, the potential of BGCs is varying widely: While genome mining of *T. virens*, *T. atroviride* and *T. reesei* revealed that *T. virens* and *T. atroviride* have more putative PKSs

(18 vs. 11) and NRPSs (28, 16 vs. 10)[73] which are located in putative BGCs than *T. reesei*, there are still interesting and valuable insights to be gained when studying *T. reesei's* BGCs. An example is the elucidated and well understood sorbicillinoid BGC. Sorbicillin was originally isolated and identified in *P. notatum* by Cram et al[74,75]. It was identified as disturbing yellow pigment in the production of penicillin from the genera *Penicillium*[75]. Nearly 50 years later, Abe et al. described the isolation of demethylsorbicillin and oxosorbicillinol after cultivation of *Trichoderma* sp. USF-2690, laying open that at least one genera of *Trichoderma* is able to produce sorbicillinoid compounds[76]. Derntl et. al, identified the main regulator responsible for the yellow pigment production in *T. reesei*[77]. They showed that the BGC, which is responsible for sorbicillinoid production, consists of two PKS adjacent to each other and that in the near distance (< 30 kilo bases) the main regulator, the yellow pigment regulator (*ypr1*), of this cluster was present[77]. A knockout of *ypr1* lead to the loss of yellow pigment formation in *T. reesei* QM6a and was regained after complementation of this gene[77]. This study demonstrated how first light can be shed into BGCs and their regulation with techniques like Reverse Transcriptase quantitative PCR (qPCR) by also showing that seven of the nine genes of this cluster, including the two PKSs, were upregulated while two were not in the knockout strain (Δ*ypr1*)[77]. It was also demonstrated that the second transcription factor (*ypr2*) of this gene cluster was repressing most other genes of the cluster and proposed that this occurs via repression of *ypr1*[77]. In a second study regarding the sorbicillinoid cluster the same authors elucidated the sorbicillinoid cluster even further showing that the main product formed by this cluster is sorbicillinol and that sorbicillinol was the main building block for the formation of other sorbicillinoids[78]. These two studies demonstrated the pathway from identifying a BGCs up to characterization of the products and even proposing a reaction mechanism for the formation of these SMs.

These discoveries are motivating for imitators which follow the trails of elucidating putative BGCs. While the sorbicillinoid cluster does not inherit a SM or derivate that is actively utilized, this example shows why the elucidation of BGCs is important: The so gained knowledge about the sorbicillinoid gene cluster could be used to optimize downstream processes by directly targeting and deleting *ypr1*, which might lead to a more cost-efficient production of enzymes because the need for removal of the sorbicillinoids in the downstream processing would be diminished.

1.2. Transcription factors and their role in gene regulation

Transcription factors (TFs) are proteins that regulate gene expression by either binding to the DNA and recruiting other enzymes to induce gene expression[80] or by binding to the DNA and blocking other enzymes from transcribing DNA[81]. If gene expression is induced the TF is labelled as activator and if gene expression is suppressed the TF is labeled as repressor. Thereby, TFs are in general structured in a modular fashion with several domains[82]. A key domain thereby is the DNA binding domain (DBD)[82]. This domain identifies and binds regulatory sequences of the gene to be transcribed[82]. In eukaryotes these sequences can be adjacent to the to-be-transcribed-gene or in distance to it[82]. The second important domain is the transactivating domain (TAD) of the TF[82]. The TAD initiates transcription via interaction with the RNA-polymerase or associated proteins after the TF has bound the DNA[82]. Another domain that TF can be comprised of are regulatory domains, that hinders the TF of binding DNA under certain circumstances[82].

TFs can be classified via their structure, resulting in different DNA binding approaches[83]. Well studied examples for common eukaryotic DBD structures are homeo-, basic-leucine zipper- or zinc-finger domains, although several more are known[83]. Homeodomain-DBD proteins form either heterodimeric structures which identify asymmetric DNA sequences and one helix of each the helix-turn-helix motifs interacts with the major groove of the DNA[83]. Basic-leucine zipper DBD-proteins consist of two helices in which opposite parts of the helices are composed of the hydrophobe amino acid leucine[83]. These leucine residues on opposite side interact with each other and form a “zipper” structure[83] similar to a zipper of a jacket or jean. The other parts of the helices are composed of basic amino acids that interact with the DNA[83]. The last example that will be elaborated are zinc finger TF[83]. They are categorized according to their zinc-binding motifs. A famous and often occurring motif thereby is the Cys_2His_2 (C_2H_2) motif, where two cysteines and two histidine residues interact with one zinc ion (Zn^{2+})[84], thereby forming a “finger-like structure” that consist of one α-helix and two antiparallel β-strands[85]. Several of these micro units can be linear arranged and interact with the DNA’s major groove[83].

In fungi, the Gal4 TF which regulates the galactose metabolism in *S. cerevisiae*, was intensively studied since its discovery in 1974[86]. The Gal4 TF of *S. cerevisiae* thereby inherits six cysteines (C_6) that bind two zinc ions, thereby being labelled as zinc binuclear cluster- or Zn_2C_6 proteins. Generally, Zn_2C_6 are able to interact with DNA as monomers[87], homodimers[88]

or heterodimers and are strictly related to the fungal kingdom[84]. Zn_2C_6 TFs can be divided in three sections, the DBD, the TAD and the regulatory regions as mentioned above[82,84]. The DBD is furthermore dissected into the zinc finger domain, the linker region and the dimerization domain[84]. The cysteine rich zinc finger domain (Zn_2C_6) is located at the N-terminus of the TF[89]. Zn_2C_6 recognizes two CGG triplets as either inverted or directed repeats and bind these triplets as homodimers while interacting with the major groove of the DNA[84]. The linker region between DBD and TAD is located C-terminal to the Zn_2C_6 motif and is widely diverse and non-conserved across fungal species[84]. The Gal4 TF linker region interacts with the phosphate backbone of the DNA and paves the way for the interaction of the dimerized TF with the major groove[90]. The dimerization region is positioned at the C-terminus of the linker and is highly conserved across Zn_2C_6 TFs[84]. This region forms coiled-coils with heptad repeats which is dimerizing and furthermore suggested to play a role in protein-protein interaction[89]. The regulatory domain is not conserved across fungal Zn_2C_6 TF and separates the DBD from the TAD[84]. The role of this region is not finally resolved but studies suggest that this regulatory domain exerts inhibition influence on TF[91,92]. Finally, the third section of zinc binuclear cluster TFs is the TAD, also referred to as acidic region and located at the C-terminus of the TF[84,89]. Its sequence is often not conserved and function as well as structure not well understood in zinc binuclear cluster TFs[84]. However, the observed modularity of the Gal4 TF lead to the development of the yeast-two-hybrid system where protein-protein interactions were observable when fusing the DBD to a protein of interest (DBD-A) and the TAD to another protein (TAD-B) which was expected to interact with protein A[93]. If this was the case then the Gal4 TF was reconstituted by bringing both domains “into the necessary distance” to activate gene expression and thus being observable[93]. The modularity and their manipulation of Gal4 TFs was again demonstrated when the TAD of the original Gal4 protein was fused to the LexA-DBD and shown to induce a reporter gene (*lacZ*) if the *lexA* operator is present near the transcription start site in *E. coli*[94].

There are several other examples where the Gal4 TF modularity was exploited[95–97], and the mentioned modularity of TFs is a promising lead to open up new regulation approaches, especially in silent BGCs.

1.2.1. Synthetic transcription factors – pathway specific modification that leads to unprecedented possibilities

Synthetic transcription factors (synTFs) make us of the modularity of native TFs. SynTFs are constructs, where the TF is modified in either its DBD or TAD or one of these two is swapped with the corresponding domain of another gene. Furthermore, the promotor of the construct can be swapped with the corresponding element from another gene for either permanent expression (constitutive promotor) or inducible expression (inducible promotor). The terminator can also be swapped from the initial gene for strong terminator to prevent leakiness.

The mentioned example of the sorbicillinoid cluster (see 1.1.1) described a BGCs that is already activated in *T. reesei* without any further necessary effort. However, most of the other BGCs remain silent under laboratory conditions and must be activated. Some strategies for the activation have already been shortly addressed in the previous abstract.

One project demonstrating the vast opportunities of synthetic TFs as novel regulation approach in activating silent BGCs, was published by Grau et al in 2018[98]. The group fused the DBD of the putative TF of the targeted cluster *alnR*, to the TAD of *afoA* (AN1029), a transcription factor shown to induce the expression of the asperfuranone cluster[69]. The authors exchanged the native promoter of *alnR* with the constitutive promotor of *alcA*, an alcohol dehydrogenase, and transformed this construct into an *Aspergillus nidulans* strain (LO9577) which lacked key genes necessary for the induction of other highly expressed BGCs (see [99]; e.g. asperfuranone cluster)[98]. In this genetic background the identification of the product of the targeted BGC was presumed to be easier[98]. Lastly, they demonstrated that with this approach and a following RNAseq analysis, they were not only able to induce the expression of the targeted BGC, resulting in the expression of (+)-Asperlin, but also defined the genes (*alnI, alnH, alnR, alnG, alnF, alnE, alnD, alnC, alnB*, and the PKS *alnA*) that were an actual part of this cluster[98] and co-expressed. Another interesting outcome of this study was that the mere replacement of the native promotor of *alnR* with the promoter of the alcohol dehydrogenase *alcA* did not result in an activation of the targeted BGC[98]. To reach this goal it was utterly necessary to replace the native TAD (*alnR*)[98] with a more potent one (*afoA*[69]).

Another modified pathway-specific approach to activate silent BGCs in which the promotor of a TF was swapped, was the study of Chiang et al. in which an *Aspergillus* strain was created that produced asperfuranone[69]. The approach included the direct insertion of the constitutive

alcohol dehydrogenase promotor (*alcA)*, replacing the original promotor of the TF, via fusion PCR[100] into an non homologous end-joining (NHEJ) deficient *Aspergillus* strain (*ΔnkuA*)[69]. The authors reported that this cluster remained silent under laboratory conditions but could be activated after the promotor-swap approach[69]. Although this activation wasn't carried out with a direct modification of the open reading frame (ORF) of the TF, it still showed that targeting TFs located in such clusters can be the key when aiming for activating silent BGCs.

Another modified pathway-specific approach was conducted when Laureti et al. were overexpressing a synTF in *Streptomyces ambofaciens* which led to the production of stambomycins A-D[101]. In this study the promotor of the putative TF (*samR0484*) was also exchanged for a constitutive promotor (*ermE*) but was not inserted at the original locus of the gene[101]. Laureti et al. designed an overexpression- cassette (OE) and integrated it into the *att*B site of the *S. ambofaciens* chromosome, making this approach slightly different than the one for the activation of (+)-Asperlin cluster, again showing that modulation of regulators in those clusters can be a successful strategy[101].

Another example for a hybrid/synthetic transcription factor and its application in *T. reesei* was demonstrated by Derntl et. al when they fused the DBD of the main cellulase regulator Xyr1 to the TAD of Ypr1 to improve cellulase and xylanase induction[102]. This was achieved nearly carbon source independent (exception lactose) which unraveled the immense potential of such constructs[102]. The development of this synTF expression system was prolonged when Derntl et al. added the ligand binding domain of the human estrogen receptor α to the former mentioned construct[102,103]. This new synTF was able to induce the expression of xylanases after addition of the inducer 17β-estradiol without the simultaneously noteworthy induction of cellulases[103]. Furthermore, the induction of xylanases was once again demonstrated to be nearly carbon source independent and in this case was also fine-tunable by addition of different concentrations of the inducer[103].

The next example where a synTF was applied in the context of a heterologous expression system in *T. reesei*, was the light-induced synthetic expression system from Zhang et al, where the author constructed chimeric TFs which dimerized light-induced, bind 5' upstream of the promotor in the second synthetic cassette with the gene of interest and induce the expression of the inserted gene of this cassette[104].

All these examples showed that there are various ways of making knowledge about regulatory mechanism count, in the relatively new scientific field of synthetic biology. The following project will be another step towards activating BGCs.

2. Research aim

The motivation to carry out this project was to activate, under standard condition silent or weakly expressed, BGCs and pave the road to the discovery of a new substance class. *T. reesei* was therefore the perfect model organism, regarding its potential of silent BGCs. The other advantages of this fungus were its GRAS-status, vast experience in handling, easy maintenance and accessibility.

To achieve the research, aim a thorough *in silico* analysis was performed. It was decided to design synthetic transcription factors which were presumed to active the *in silico* analyzed clusters after incubation on glucose as sole carbon source. Furthermore, the focus laid on determine if and which of the associated predicted genes of these clusters were also co-expressed with the core biosynthetic genes.

With this knowledge future experiments will include to screen the metabolome of the most promising transformants with the goal of identifying the newly expressed SM(s) in the end.

3. Results

3.1. *In silico* analysis yields 91 putative biosynthetic gene clusters – 4 selected

The *in silico* analysis, performed with antiSMASH[63], with the genomic data[105] of *T. reesei's* QM6a yielded 91 putative biosynthetic gene clusters. Some of these clusters were assigned one or more core biosynthetic genes. These were either encoding a polyketide synthase (PKS), non-ribosomal peptide synthetase (NRPS), hybrids of the former mentioned enzymes (PKS-NRPS) or other similar biosynthetic core enzyme, like terpene synthases (TS). Some clusters were even assigned two or more of the former mentioned core biosynthetic enzymes. On the other hand, there were clusters that were identified as such but did not inherit any of the mentioned enzymes (see Table 1/Cluster 85). Few of these cluster's core biosynthetic genes inherited one conserved domain that is also part of PKSs, NRPSs or TSs but were missing several other crucial domains and were therefore not identified as PKS, NRPSs or TS (see Table 1/Cluster 3) by the algorithm. Table 1 summarizes the chosen BGCs, the core biosynthetic enzymes of these clusters and illustrates three examples of non-chosen BGCs.

Table 1. Selection of BGCs after antiSMASH analysis. Some clusters contained PKSs, NRPSs, TSs or none of the mentioned ones; put … putative

Cluster	Biosynthetic core enzyme(s)	Chosen
21	PKS-NRPS: 58285	YES
22	PKS: 105804 & PKS-NRPS: 59315	YES
52	PKS: 65172	YES
78	PKS: 81964	YES
3	put. condensation domain containing protein: 38640	NO
23	put. TS: 59597	NO
85	put. cytochrome P450: 70339 & put. dehydrogenase/reductase: 70334	NO

It was decided to focus on clusters 21, 22, 52 and 78 because two of them inherited a type I-PKS (52 and 78), one a hybrid type I-PKS-NRPS (21) and one both (22). It was assumed that the chances of actually expressing and finding a new secondary metabolite or derivate after activation of these clusters was higher compared to other clusters that had not been clearly assigned a PKS, PKS-NRPS or NRPS. Other criteria for these four clusters were that they all inherited at least one putative transcription factor and that the predicted products by the antiSMASH analysis were all labelled putatively non-hazardous.

Three examples of clusters that were not chosen were nonetheless displayed in Table 1 to demonstrate the reasoning behind our choices:

The first example of a non-chosen cluster was cluster 3. Cluster 3 was assigned a core biosynthetic gene (38640) with a condensation domain but this gene lacked any other conserved domain necessary for being identified as PKS, NRPS or TS. The second example of the non-chosen BGC was cluster 23 which was not chosen for activation because although inheriting a putative TS (59597), there was no putative TF in this cluster. The last exemplary displayed cluster, which was not chosen, was cluster 85 because it did not inherit a biosynthetic core gene which is known to form a secondary metabolite but the only two core biosynthetic genes in it were a putative cytochrome P450 (70339) and a putative dehydrogenase/reductase (70334). The organization of the chosen clusters can be seen in Figure 2.

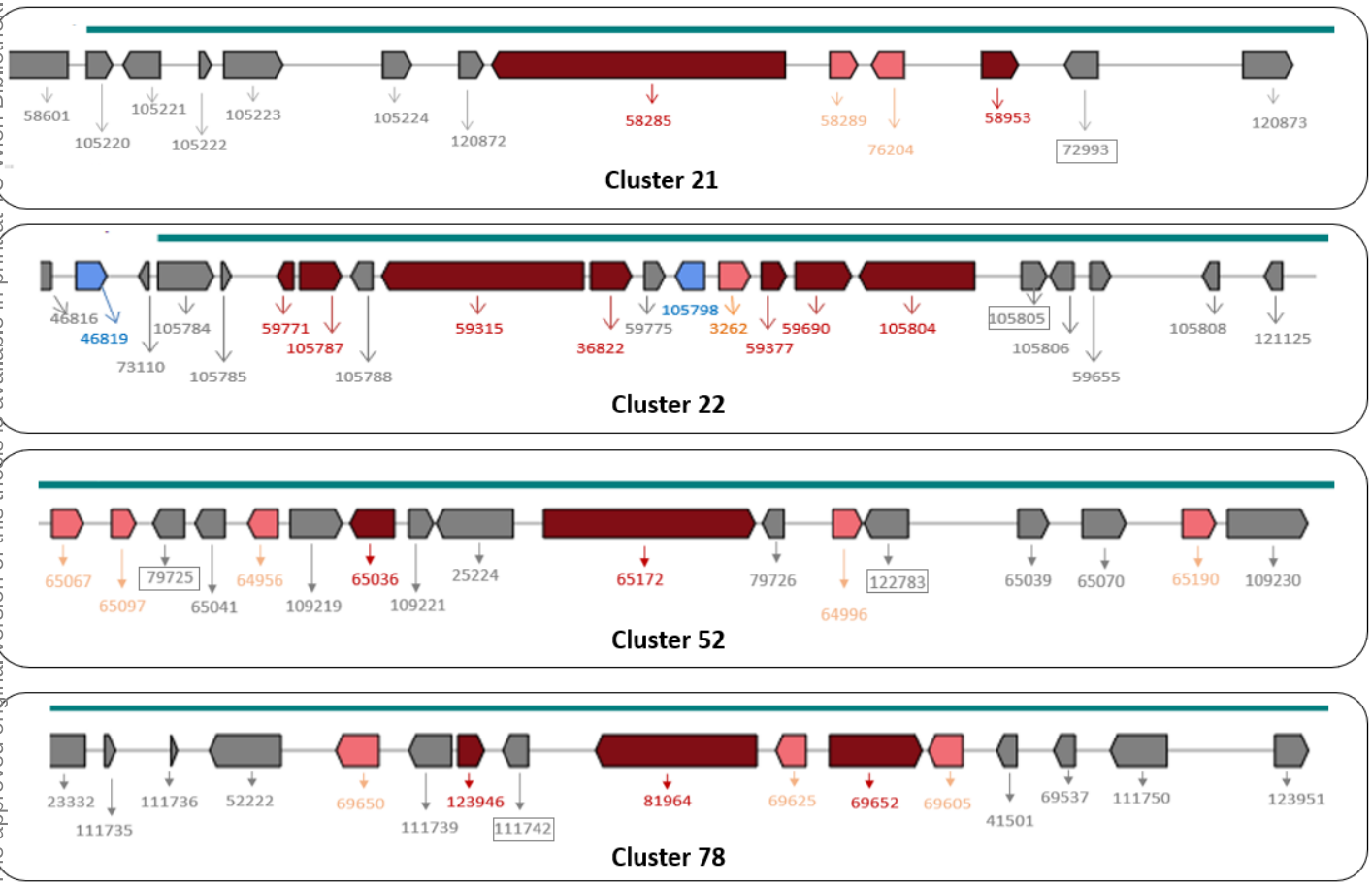

Figure 2. Cluster organization of chosen BGCs.
Numbers indicate the gene numbers/protein IDs assigned to these sequences after gene prediction by the Joint Genome Institute[106]; Grey arrow… other genes; Blue arrows … transport related genes; Red arrows … core biosynthetic genes; Pink arrows … additional biosynthetic genes; Grey boxes … putative transcription factors of corresponding cluster; Blue/greenish bar … ClusterFinder[107] association of genes that are predicted to be an actual part of the cluster

The genes that were associated with each cluster were either already annotated, assigned a putative function or had no assigned functions. Each cluster showed more than one core (red arrow - Figure 2) and additional biosynthetic gene (pink arrow - Figure 2), which indicated that the core chemical structure of the secondary metabolite will be modified[24]. Blue marked genes (blue arrows - Figure 2) indicated that this gene was most likely to perform a transport function and grey genes (grey arrows - Figure 2) were genes that have been identified as genes but couldn't be assigned a specific function by the antiSMASH algorithm. The blue/greenish bars (blue/greenish bar - Figure 2) over certain genes of each cluster indicated that the associated genes are an actual part of the cluster.

Cluster 21 was assigned two core (58285 and 58953) and two additional biosynthetic genes (58289 and 76204) and several other genes with no annotated function. Gene 58601 was excluded by the ClusterFinder algorithm to be an actual part of this BGC. Cluster 22 was assigned seven core (59771, 105787, 36882, 59337, 59690 and 105804) and one additional biosynthetic gene (3262). In this cluster three genes were excluded by ClusterFinder algorithm from being an actual part of this BGCs (46818, 46819 and 73110). Cluster 52 was assigned two core (65036 and 65172) and five additional biosynthetic genes (65067, 65097, 64956, 64996 and 65190). Cluster 78 was assigned three core (123946, 81964 and 69652) and three additional biosynthetic genes (69650, 69625 and 69605). ClusterFinder did include all genes to be part of the cluster 52 and 78.

PANNZER2[108] (Protein ANNotation with Z-scorE) analysis identified the putative TFs of each cluster that were suspected to regulate the expression of the corresponding PKS or PKS-NRPS. This bioinformatic tool, generally predicts functions and/or biological processes in which putative genes might be involved[108]. The PANNZER2 algorithm thereby executes homology sequence comparisons with mass-annotation based on sequence similarity using SANSparallel[109]. An overview of the identified TFs is prepared in Table 2.

Table 2. PANNZER2 identification of transcription factors in the chosen clusters.
Est. … Estimated; CR … cluster regulator; TR … transcription regulator; TF … transcription factor; *apdR* … aspyridone cluster regulator; *nscR* … neosartoricin regulator

Cluster	Transcription factor	Est. Description	Est. Biological process	Est. Molecular function
21	72993	Aspyridone CR *apdR* (0.76)	Transcription (0.65)	Zinc ion binding (0.63) DNA binding (0.56)
22	105805	Zn_2Cys_6 TR (0.45)		
52	79725	C_6 domain TF *nscR* (0.40)	Transcription (0.64)	Zinc ion binding (0.63) DNA binding (0.55) DNA- binding transcription factor activity (0.70)
	122783	Zn_2Cys_6 TR (0.41)	Transcription (0.69)	Zinc ion binding (0.63) DNA- binding transcription factor activity (0.72)
78	111742	Zn_2Cys_6 TR (0.44)	Transcription (0.51)	Zinc ion binding (0.63) DNA binding (0.56) DNA- binding transcription factor activity (0.72)

One TF was identified in cluster 21, 22 and 78 and two TFs were identified in cluster 52. All of them inherited a GAL4-like Zn_2Cys_6 zinc finger motif which made it likely that these were in fact regulatory proteins. The estimated biological processes (transcription) and molecular function (zinc ion binding, DNA binding and DNA- binding transcription factor activity) did also fit well, having predicted probability numbers of at least 51 %. These numbers after each prediction are an indicator for how likely it is that the prediction was likely and is ranging between zero and one. The closer to zero the unlikelier the prediction and vice versa.

The final length of the truncated TFs was chosen according to a BLASTX[110] analysis for four of the five TFs (72993, 105805, 79725, 111742). The final length of truncated TF 122783 was chosen after analysis with BLASTP[110] because analysis with BLASTX didn't reveal the zinc finger motif. The results are displayed in Figure 3.

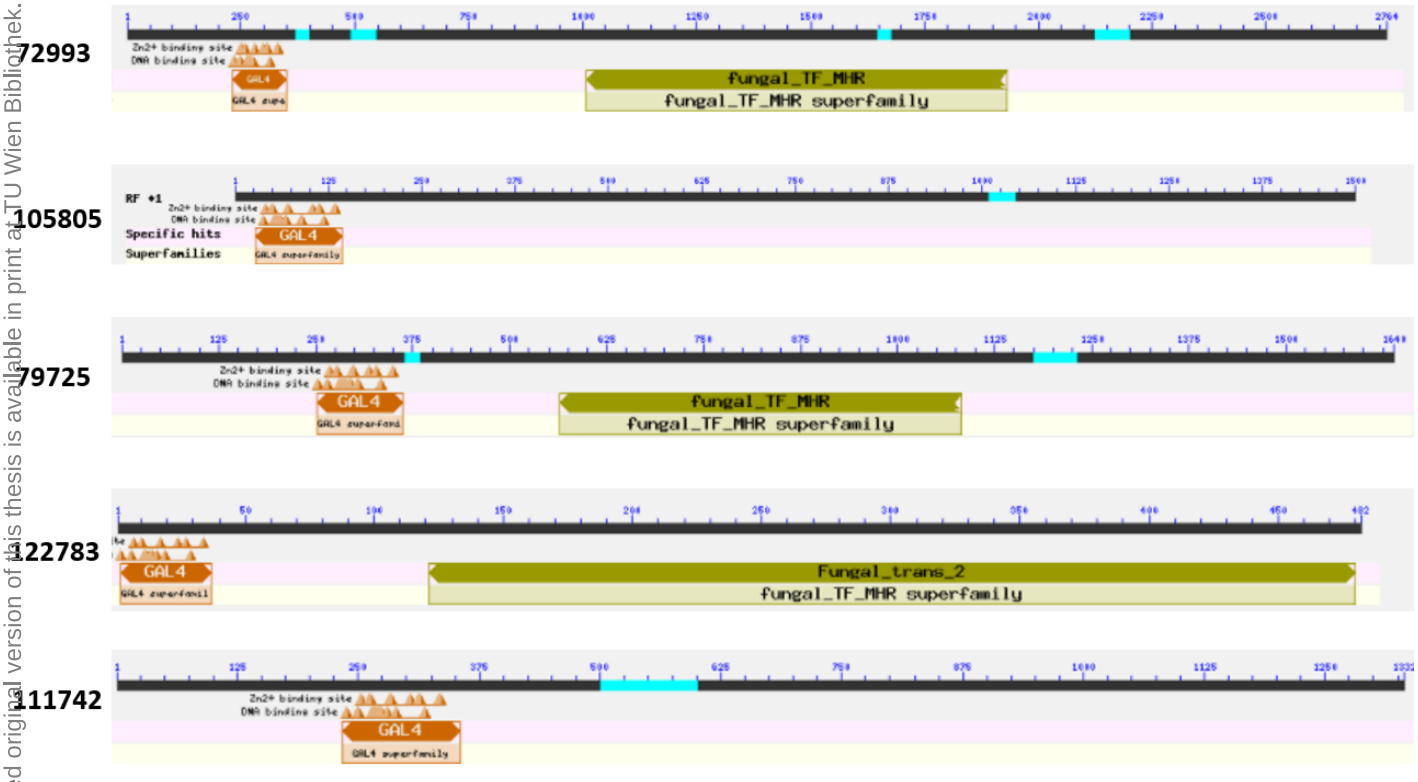

Figure 3. BLASTX and BLASTP analysis of TFs.
All five TFs showed a GAL4-like zinc finger domain at the N-terminus of their sequences.

The lengths of the truncated TFs were chosen for each TF in order to comprise the conserved Gal4-like domain and the conserved areas up to the fungal transcription factor middle homology region (FTFMHR) where applicable. An extensive explanation on how these areas have been chosen is provided in the original publication where this strategy was applied firstly by Derntl et al[102]. Three of the four TFs (72993, 79725 and 122783) did have the conserved FTFMHR

domain which is suggested to play a regulatory role[84] and was therefore strengthening the correctness of the gene prediction. An overview of the lengths of the DBD' is given in Table 3.

Table 3. Chosen length of the TFs after *in silico* analysis

TF	DBD' - length (bp)
72993	807
105805	553
79724	377
122783	454
111742	401

3.2. Generation of synthetic transcription factor strains and characterization

The generation of the strains bearing the synthetic transcription factor cassettes (synTF) ranged from cloning over transforming up to the characterization of the obtained strains via genotyping.

3.2.1. Cloning

The cloning procedure comprised three steps:

Firstly, the truncated TFs (DBD') were amplified via PCR and the resulting PCR products were cloned into the vector pJET1.2. The strategy behind this step was on the one hand to send this plasmid to our sequencing collaborator and on the other hand to excise the DBD' in the next step with a higher accuracy.

In the following step, the excised DBD' was cloned into vector pCD_pXY1 to assemble the DBD' with the truncated transactivation domain (TAD') from Ypr1. This TAD' was already in pCD_pXY1 in its original design only with a different DBD', namely a truncated version of the DBD of Xyr1. The sequence of pCD_pXY1 is provided in section 7.1.

In the last step the excised synthetic transcription factor (DBD' + TAD') was cloned into vector pRP4_TX with the purpose to reintegrate these vectors into the *pyr4* locus.

A schematic overview of the cloning strategy that was applied is given in Figure 4. This figure was created with Biorender.com

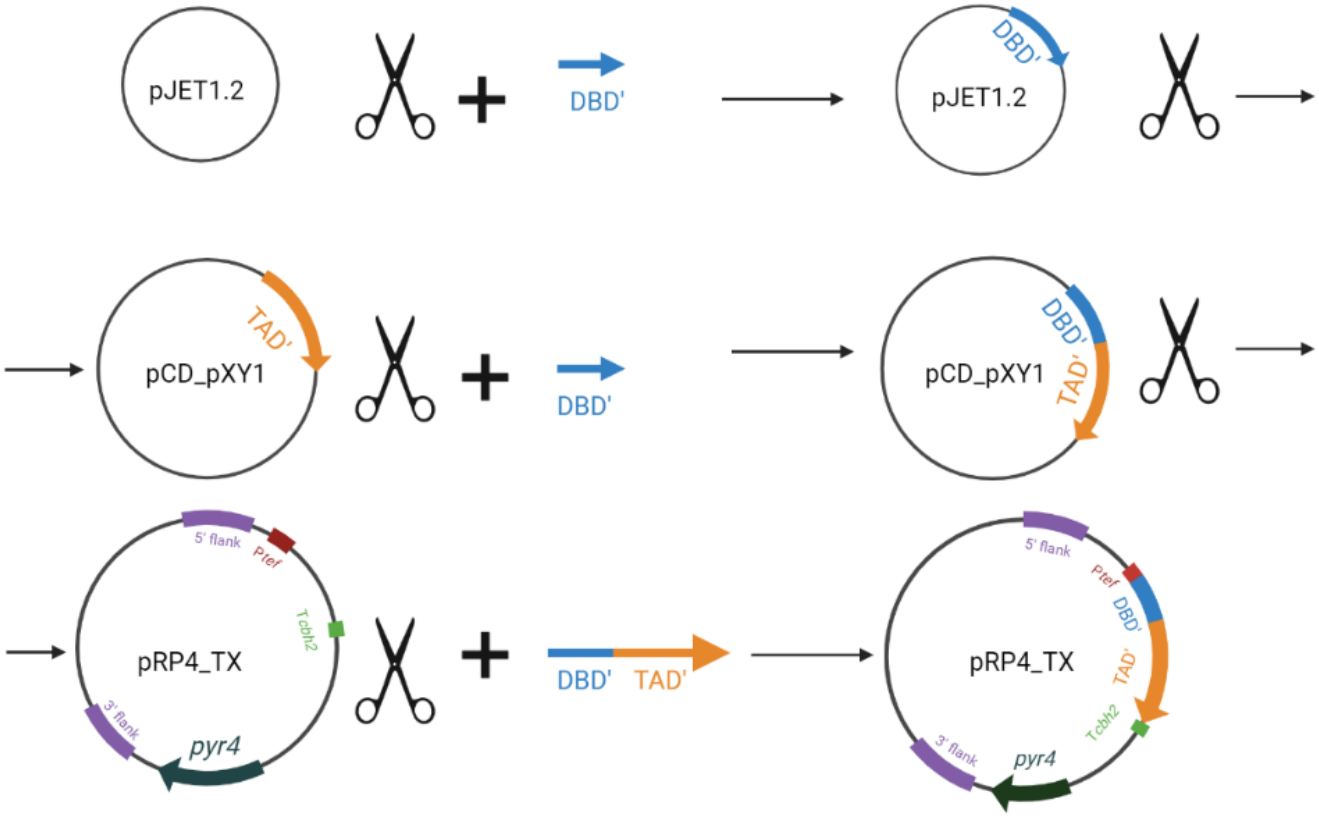

Figure 4. Schematic representation of cloning strategy.
5'/3' flank – 5'/3' homologous flank of *pyr4* locus; P*tef1* – promotor of *tef1;* DBD' – DNA-binding domain of truncated TF; TAD' – truncated transactivation domain of Ypr1 (Yellow pigment repressor 1); T*cbh2* – terminator *cbh2*; *pyr4; pyr4 – pyr4* gene; Figure created with biorender.com

The sequence result of the PCR products of the DBD' of each TF was compared to the genomic sequence of the TFs with MAFFT[111] algorithm and visualized in benchling[112]. Figure 5 shows the comparison.

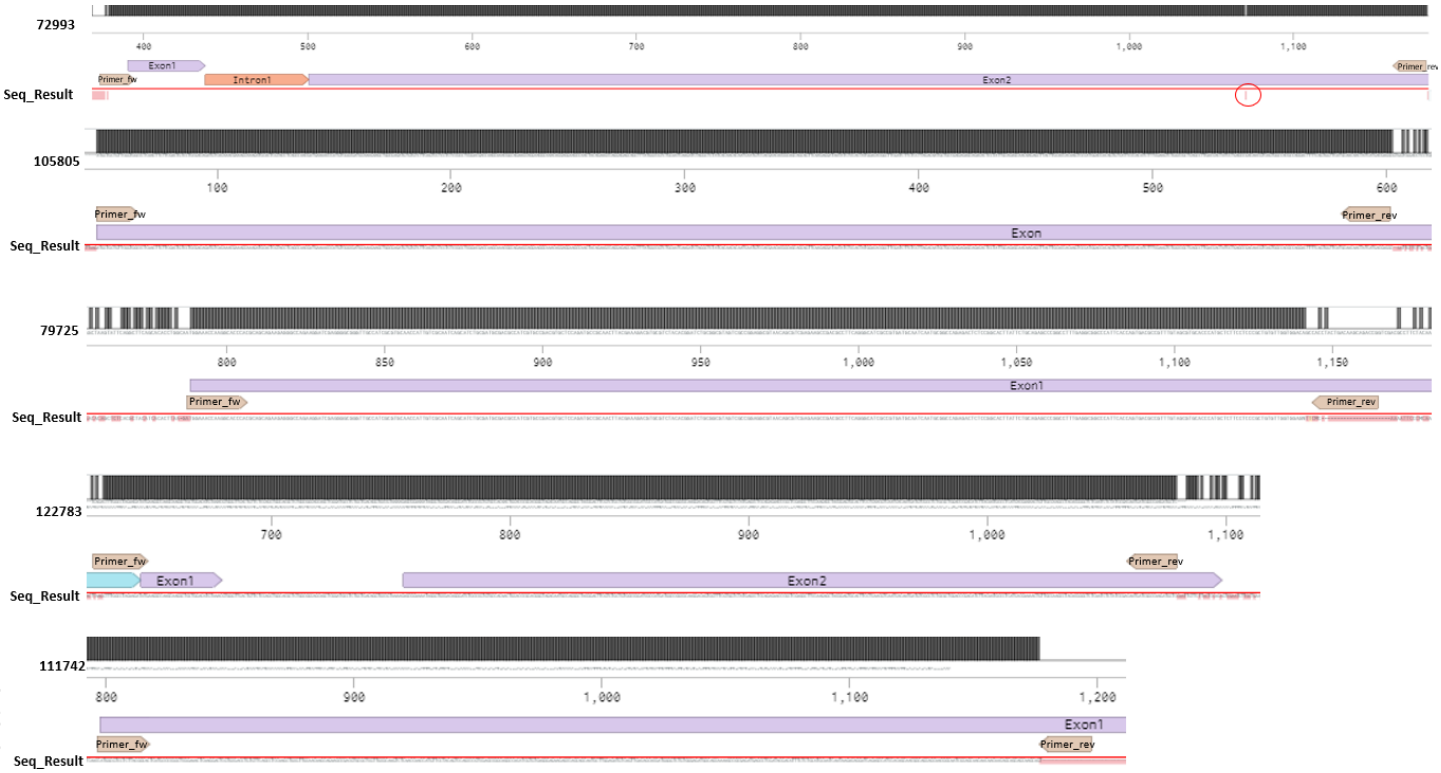

Figure 5. Sequence comparison of genomic sequences of the TFs with DBD' sequences of the TFs.
Seq_Result … sequencing result; Dark-grey vertical bars: consensus sequences, brown arrows: amplification primers, orange arrows: intron, purple arrows: exons, red vertical bars (circle): mismatch

The dark-grey vertical bars on top of each transcription factor sequence indicated consensus sequences. The absence of these bars (white space between grey bars) indicated different bases and/or failed measurement. **Four** truncated TFs (105805, 79725, 122783, 111742) were amplified completely correct and the fifth (72993) showed one single mismatch at position 690 in exon 2 (red circle). The mismatch was incorporated in the codon CGG where the third position was cytosine instead of guanine. This mismatch resulted in a silent mutation and furthermore was not located in the conserved sequence Gal4-like motif of the DBD'. We therefore decided to presume with this mismatch.

Verification of successful cloning was carried out for all five TFs after cloning the DBD' into pCD_pXY1 and for three synTFs (105805, 79725, 111742) after cloning the synTFs into the final construct pRP4_TX. Figure 6 displays the Colony PCR results. Figure 7 displays the test restriction results. The lengths of the final constructs are shown in Table 4.

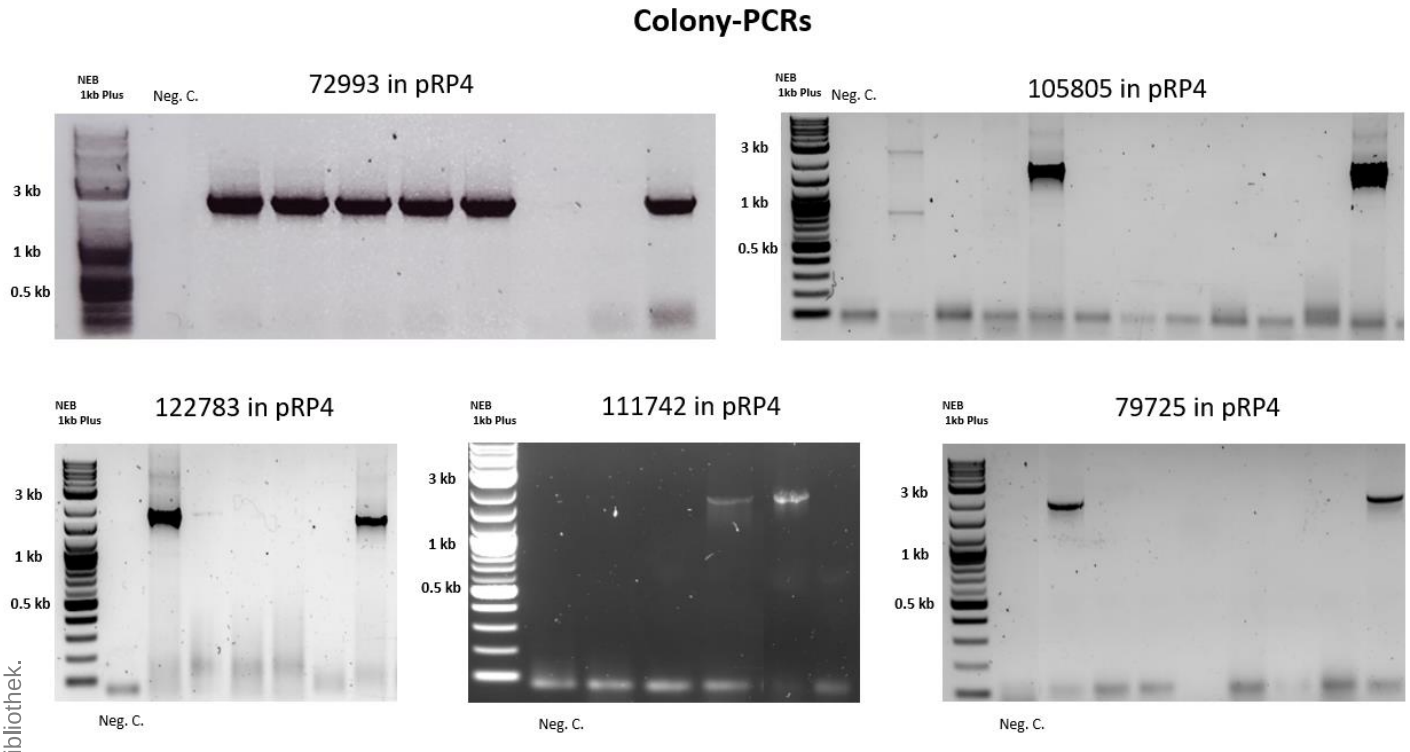

Figure 6. Colony-PCRs of synTFs-cassettes in *E. coli* single colonies after transformation of the final vector pRP4_TX. Neg. C. … negative control of PCR (= water); Primers from Table 14 have been used for amplification: Primer Ptef was used to bind inside the promotor of *tef1* and reverse primers corresponding to the DBD' were used to bind 3' end of the DBD'. At least two colonies from each synTF show the expected bands (inserts).

At least two *E. coli* colonies for each synTF showed the expected signals in the Colony-PCR (see Figure 6) and the vectors thereof were subsequently used for transformation into *T. reesei*.

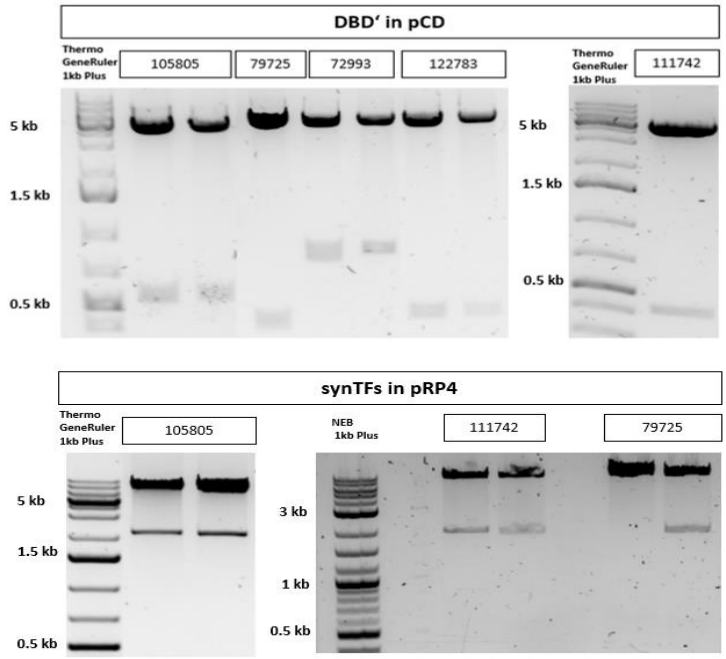

Figure 7. Test restriction digests of vectors pCD and pRP4 bearing either the truncated DBD (DBD' in pCD_pXY1) or the complete synTF-cassette; DBD' … truncated DNA-binding domain; pCD_pXY1 … intermediate vector 2; pRP4_TX … final vector;

Test restriction digests (Figure 7) showed that all the DBD' had the correct length (DBD' in pCD). Three of the assembled synTFs (105805, 111742 and 79725) were also demonstrated to have the expected length in the final vector (synTF in pRP4 - Figure 7 - see Table 4). synTFs 122783 and 72993 weren't analyzed via test restriction due to lack of time.

Table 4. Overview - synTFs and vectors

synTF (DBD'+TAD')	length [bp]	vectors	vector lengths [bp]
72993	2197	pCD pRP4	~ 4100 ~ 11000
105805	2027		
79724	1847		
122783	1827		
111742	1867		

3.2.2. Transformation

The synTF-expression cassettes were reintegrated into the *pyr4* locus of *T. reesei* QM6a *Δtmus53 Δpyr4* via transformation. This strain was chosen because it was shown that the knockout of the *tmus53* gene led to a lack of efficient Non-homologous end joining (NHEJ) mechanism which massively elevates the chances of integrating an insert into the favored locus by homologous recombination[113]. The *pyr4* gene encodes the orotidine 5-monophosphate decarboxylase[114]. This enzyme is vital in the pyrimidine synthesis of the organism and therefore a *pyr4* knockout-strain can be used in selection for prototrophy[102]. The parental strain is not able to grow on media without the external addition of uridine because *pyr4* is missing. By retransforming the syn-cassettes into the *pyr4* locus the transformants will gain back a copy of the *pyr4* gene and would therefore again able to grow on minimal medium without the addition of uridine.

The transformation process consisting of transforming, colony picking, homokaryon streaking and single colony picking is exemplary shown for one transformant, with the synTF-cassette of 72993 inserted, in Figure 8 and was carried out for all transformants listed in Table 5.

Table 5.

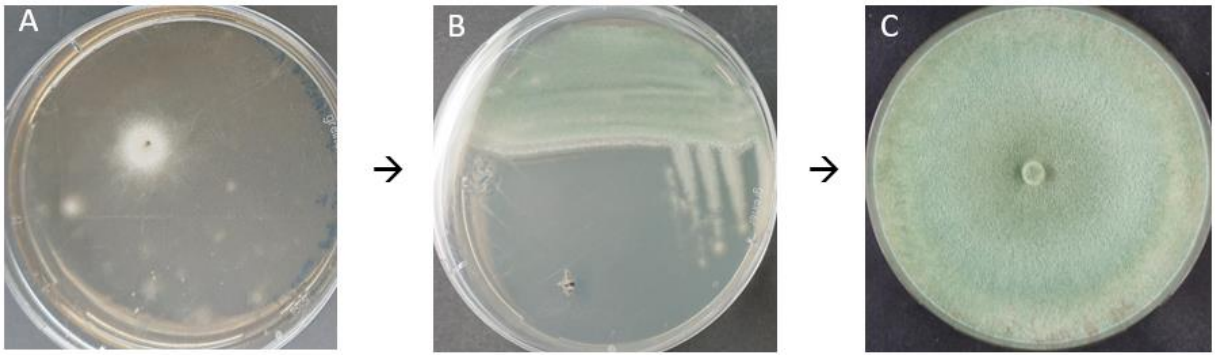

Figure 8. Transformation process by example of *T. reesei* QM6a *Δtmus53* synTF-72993 (*pyr4*).
A …Transformed *T. reesei* colonies after 4-5 days; B … homokaryon streaking and single colony picking; C … Transformant after single colony picking after 2-3 days; Selection pressure was kept up throughout all stages by using minimal medium (MA) without the addition of a uridine source.

3.2.3. Genotyping

The promising syn-TF transformed candidates were subject to cultivation on liquid MEX medium at 30 °C. DNA was extracted after 2 days and subsequently analyzed via PCR. At least three transformants per synTF-cassette were screened for the correct genotype. Additionally, at least three transformants with overexpression-cassettes (OE) were screened. The OE-cassettes

were constructed and transformed by colleagues in this research group but analyzed subsequently (see 6.2). The aim of reinsertion into the *pyr4* locus was so confirmed in both types of transformants (OE and syn). The primer pairs were designed to confirm *pyr4* locus specific integration with the following strategy:

The first pair was designed so that the forward primer would bind outside the 5' transformed flank and the reverse primer was designed to bind in the *tef1* promotor. The second pair was designed so that the forward primer would bind inside *pyr4* gene and the reverse primer was designed to bind outside the 3' transformed flank. With this strategy it was possible to screen all generated strains rapidly and simultaneously. Figure 9 shows the genotyping results. The used primers are shown in Table 15.

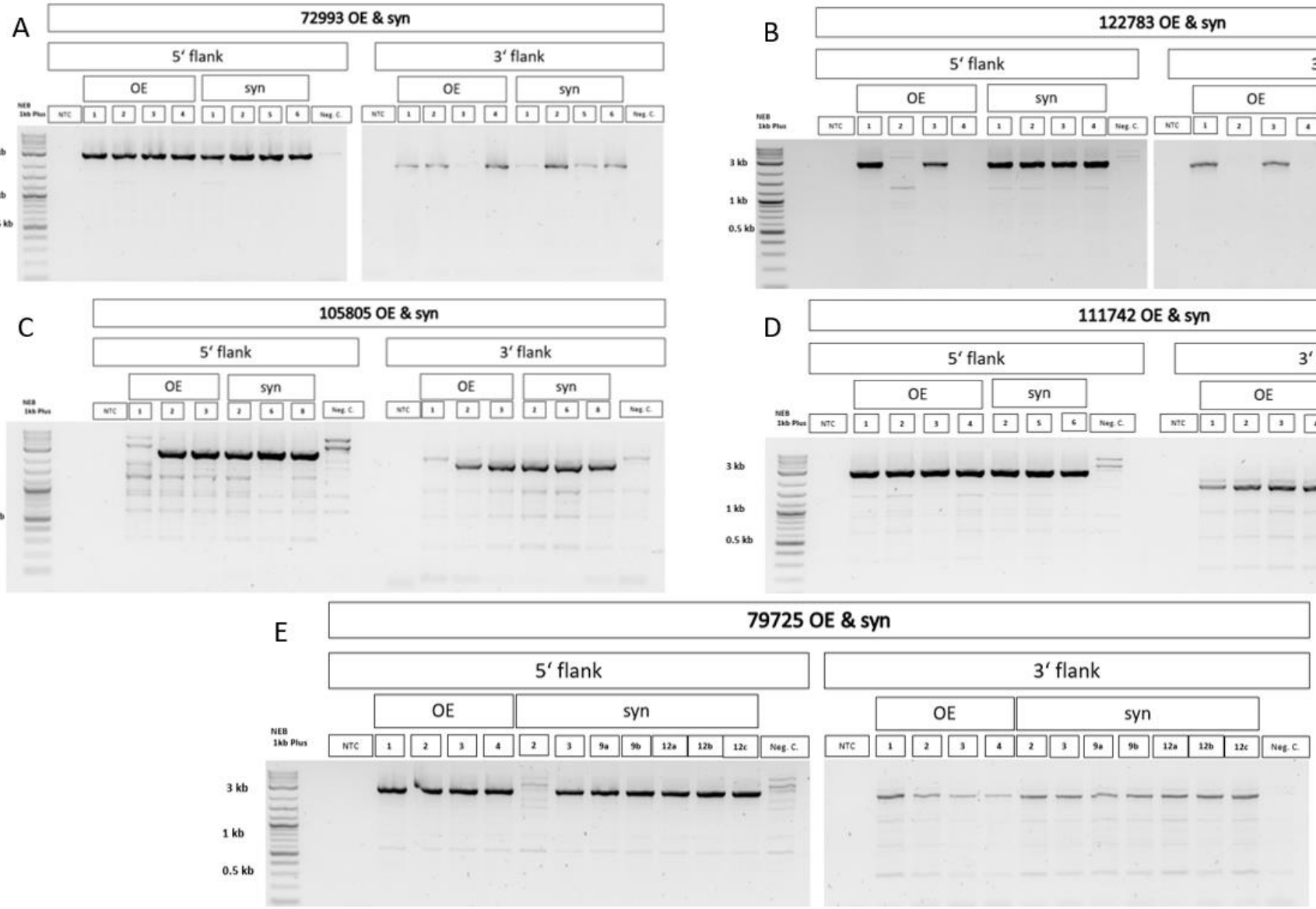

Figure 9. Genotyping of generated OE and synTFs strains via PCR.
5'/3' flank … amplification outside transformed flank and inside respective construct; NTC … No Template Control (water); Neg. C. … Negative Control – *T. reesei* QM6a *Δtmus53 Δpyr4;* OE … overexpression transformants; syn … synTF transformants; kB … kilo bases; Signals that were strong and different compared to the negative control for the 5' and 3' flank, were judged as successfully transformed.

In Figure 9/A/B/C/D/E all no template controls (NTC - water) showed no amplicons, which was expected. The negative controls (Neg. C – *T. reesei* QM6a *Δtmus53 Δpyr4*) yielded none or not the correct signal at the 5' and 3' flank in all PCRs. This was also expected because the reverse primer of the 5' flank and the forward primer of the 3' flank were designed to not bind in the parental strains *pyr4* locus.

Some transformants did not show the expected band at either the 3' or 5' flank. The signal was either missing completely (e.g. missing 3' flank - 72993-OE-3; missing 5' flank – 122783-OE-3 or 79725-OE-3) or did show a weak signal (e.g. 72993-syn-2). Interestingly, in some of these transformants one flank did show the expected signal whereas the other flank did not. All the transformants where such results were obtained did not integrate the cassette correctly and were therefore discarded.

For all other shown transformants the expected signals at the 5' and the 3' flank were obtained correctly. At least two transformants per OE and syn-cassette were processed further and analyzed with Reverse-Transcriptase quantitative PCR (qPCR). An overview of these strains is given in Table 5.

Table 5. Summary synTF and OE strains that were analyzed via qPCR for cluster activation.

Strains	Transformants	Abbreviation	Cluster
T. reesei Δtmus53 synTF-72993 (*pyr4*)	2	72993-syn-2	21
		72993-syn-6	
T. reesei Δtmus53 OE-72993 (*pyr4*)	2	72993-OE-1	
		72993-OE-2	
T. reesei Δtmus53 synTF-105805 (*pyr4*)	2	105805-syn-2	22
		105805-syn-2	
T. reesei Δtmus53 OE-105805 (*pyr4*)	2	105805-OE-2	
		105805-OE-3	
T. reesei Δtmus53 synTF-79725 (*pyr4*)	4	79725-syn-3	52
		79725-syn-9a	
		79725-syn-9b	
		79725-syn-12a	
T. reesei Δtmus53 OE-79725 (*pyr4*)	2	79725-OE-1	
		79725-OE-3	
T. reesei Δtmus53 synTF-122783 (*pyr4*)	2	122783-syn-2	
		122783-syn-9	
T. reesei Δtmus53 OE-122783 (*pyr4*)	2	122783-OE-1	
		122783-OE-3	
T. reesei Δtmus53 synTF-111742 (*pyr4*)	3	111742-syn-2	78
		111742-syn-5	
		111742-syn-6	
T. reesei Δtmus53 OE-111742 (*pyr4*)	4	111742-OE-1	
		111742-OE-2	
		111742-OE-3	
		111742-OE-4	

3.3. Three out of four clusters successfully activated

The generated *T. reesei* transformants from Table 5 and the WT (*T. reesei* QM6a *Δtmus53*) have been cultivated on minimal medium (MA without citric acid or peptone – see Table 17) with glucose as sole carbon source for 72 hours to estimate the transcript levels of the targeted PKS and PKS-NRPS genes. This medium did not contain any other carbon source except glucose. This was achieved by modifying the recipe of the MA buffer with adjusting the pH to 5 with HCl instead of citric acid. With this setup it was made sure that the only carbons that would be metabolized by the fungus would be the ones deriving from D-glucose and in the following project, the search for the putative newly expressed SM(s) will be easier (see tracer derived approach – section 7.5). RNA was extracted, reverse transcribed to cDNA and subsequently analyzed with qPCR. The results of this transcript level estimation are displayed Figure 10.

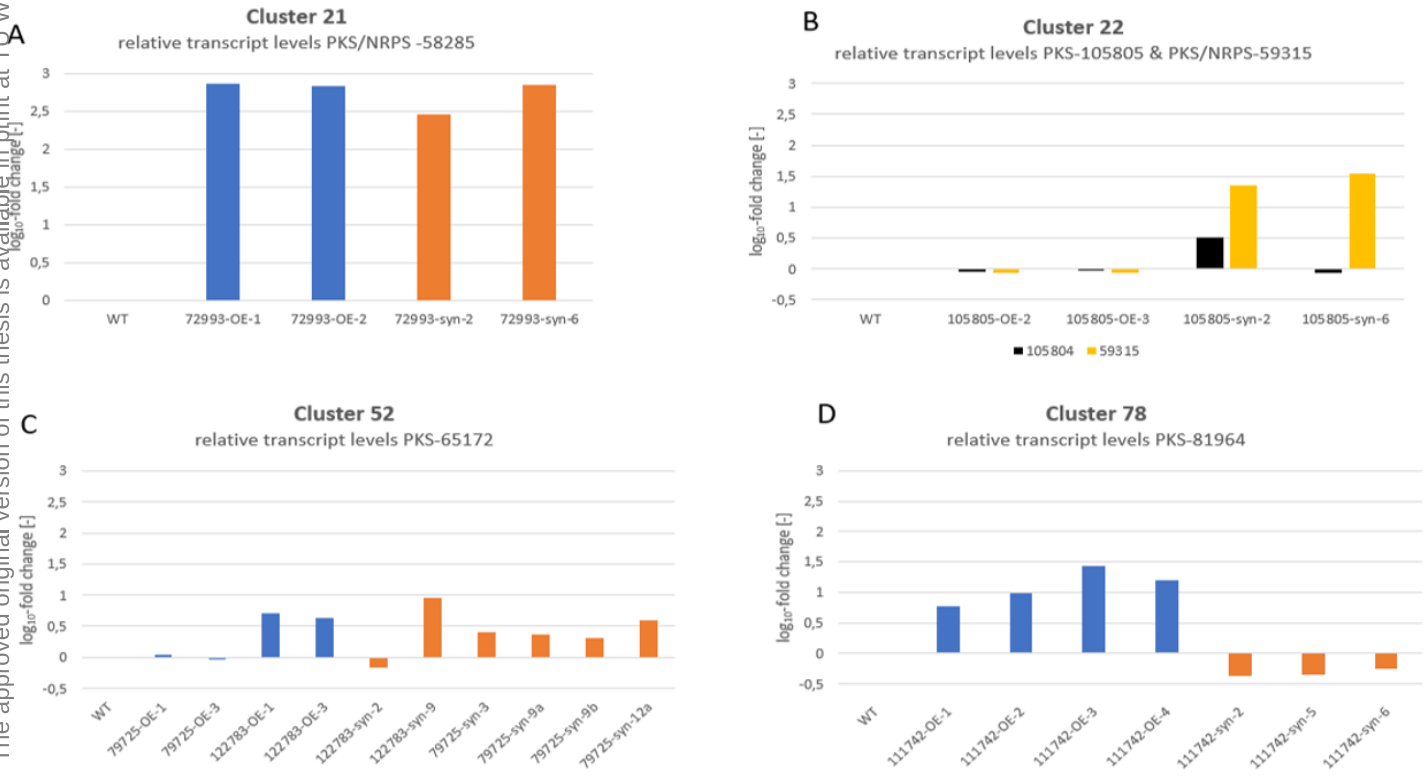

Figure 10.Transcript level estimation of PKS, NRPS/PKS of each cluster.
WT … *T. reesei* QM6a *Δtmus53;* OE … overexpression transformants; syn … synTFs transformants; Transformants and WT have been cultivated on glucose as sole carbon source. Samples have been taken after 72 hours. Relative transcript levels of the indicated genes have been normalized to the WT samples using the reference genes *act1* & *sar1*. Mean values of two technical replicates are given.

The transcript levels of PKS-NRPS hybrid enzyme 58285 of cluster 21 were between 2.45 and 2.85 log_{10}-fold higher in both tested OE and syn strains (72993-OE-1/2 and 72993-syn-2/6) compared to its parental strain (WT - see Figure 10/A). We therefore considered this cluster activated.

The transcript levels of PKS 105804 of cluster 22 were slightly lower in both OE transformants (105805-OE-2/3) and showed mixed regulation signals in the syn transformants (105805-syn-2/6) compared to the parental strain (see Figure 10/B). In 105805-syn-2, PKS 105804 transcript level was 0.5 log_{10}-fold higher, whereas the transcript levels of 105804 were slightly lower (- 0.07 log_{10}-fold) in 105805-syn-6 compared to the parental strain. Cluster 22 would have been considered not activated judging from these results, if this was the only PKS but there is a second core biosynthetic enzyme a few genes besides 105804, namely the hybrid PKS-NRPS 59315. While there were slightly lower transcript levels in both OE transformants (105805-OE-2/3) regarding 59315, higher transcript levels of 59315 were detected in both syn transformants (105805-syn-2/6), ranging between 1.35 and 1.55 log_{10}-fold compared to the parental strain. Cluster 22 was therefore, considered activated in the syn transformants regarding PKS-NRPS 59315.

The transcript levels of PKS 81964 in cluster 78 were between 0.77 and 1.44 log_{10}-fold higher in all four OE transformants (111742-OE-1/2/3/4) but lower in all three tested syn transformants (111742-syn-2/5/6 – see Figure 10/D) compared to the parental strain. Cluster 78 was therefore considered activated in the OE transformants.

The only cluster that that wasn't considered activated was cluster 52 in both OE and syn transformants. The OE transformants of both TFs (79725-OE-1/3 and 122783-OE-1/3) showed mixed or no sufficiently different log-fold change of the PKS 65172 compared to the previous described transformants of the other three clusters (see Figure 10/C). In 79725-OE-1, the transcript level of 65172 was slightly higher (0.03 log_{10}-fold) and in 79725-OE-3 it was slightly lower (- 0.03 log_{10}-fold) compared to the parental strain. In 122783-OE-1/3, the transcript levels of 65172 were both moderately higher (0.62 and 0.70 log_{10}-fold) but not sufficient enough to our judgement (see section 6.6.2 for explanation). In the synTF transformants of 79725 (79725-syn-3/9a/9b/12a), the transcript levels of 65172 were all ranging between 0.25 and 0.55 log_{10}-fold higher compared to the parental strain but were again judged as insufficient to declare the cluster activated in these strains. The synTF transformants of 122783 (122783-syn-2/9) showed mixed regulation signals: While the transcript levels of 65172 were slightly lower (- 0.16 log_{10}-fold) in 122783-syn-2, they were moderately higher (0.98 log_{10}-fold) in 122783-syn-9 compared to the parental strain. Cluster 52 was considered not activated in these strains due to

the mixed regulation signals in 122783-syn transformants and only a small regulation difference in 122783-OE transformants. The same was stated for this cluster in the 79725 transformants of both types (syn and OE).

The explanation why most results were considered sufficient and a few not, for a clear statement regarding the activation of a cluster is prepared in the discussion part of this thesis (see section 4.2 and 6.6.2).

3.3.1. Co-expression analysis within the activated BGCs

Having three clusters activated, we were interested to investigate the prediction of the biosynthetic core enzymes and additional biosynthetic enzymes in those clusters (21, 22, 78) given by antiSMASH analysis. One of the reasons to shed light into the activated clusters was to confirm co-expression of other putatively necessary genes, namely the core and additional biosynthetic genes. If the activation, in form of higher transcript levels of the PKS, PKS-NRPS, was sincere and the predicted clusters are in fact BGCs, then we can assume that the other biosynthetic genes or additional biosynthetic genes are co-expressed too.

We decided to add another layer of bioinformatic analysis before deciding on which genes to focus. This bioinformatic tool is named FunORDER[115] and was designed to identify essential genes, that belong to the same BGC, based on computational molecular co-evolution. It yields heat maps, principal component analyses (PCA) and dendrograms calculated with ward's methods that aims at predicting co-evolution between genes in a putative BGC[115]. The tool thereby declares genes as co-evolved if their pairwise calculated strict and evolutionary distance was below or equal to the value of 0.7 and the combined distance of both was below or equal to the value of 0.6[115]. These calculations were performed by the TreeKO algorithm and the difference between both distances is that the strict distance is penalizing dissimilarities in evolution, such as gene duplication and gene loss events, whereas the evolutionary distance does that not[116]. It is therefore proposed by the authors of the TreeKo algorithm that the strict distance is more suitable in detecting co-evolution[116]. This was taken into account and therefore, this work only displayed the plots and heatmaps in the supplementary section (see 7.3), in which the strict distances were calculated (see Figure 23/Figure 24/Figure 25), although all resulting plots (e.g. evolutionary distance and combined) were analyzed and regarded before the final decision which genes were chosen for qPCR analysis.

3.3.1.1. Cluster 21

The co-expression analysis started for each cluster with the comparison of the prediction results from the antiSMASH and the FunORDER algorithm. Therefore, cluster 21 is schematically displayed in Figure 11 with the predictions of the antiSMASH algorithm. The antiSMASH algorithm predicted the type of gene in this cluster. The types ranged from core biosynthetic genes (red -Figure 11) over additional biosynthetic genes (pink - Figure 11) up to transport related genes (blue – not displayed here) and other genes (grey - Figure 11) with no assigned function.

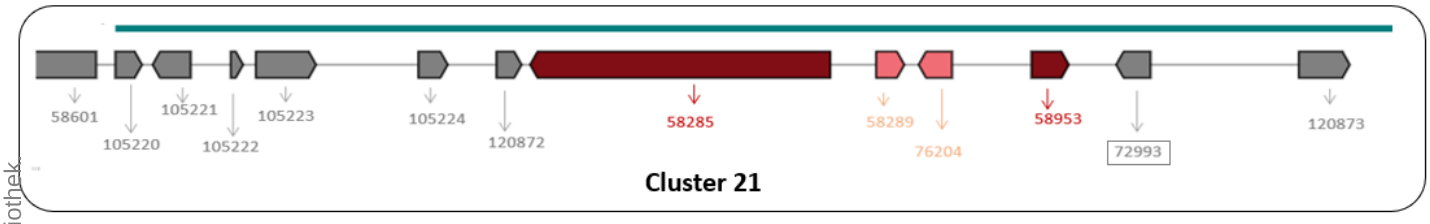

Figure 11. Cluster overview of cluster 21.
Numbers indicate the gene numbers/protein IDs assigned to these sequences after gene prediction by the Joint Genome Institute[106]; Grey arrow… other genes; Red arrows … core biosynthetic genes; Pink arrows … additional biosynthetic genes; Grey box … putative transcription factors of corresponding cluster; Blue/greenish bar … ClusterFinder[107] association of genes that are predicted to be an actual part of the cluster;

Table 6 gives an overview of the predictions both applied algorithms made and the genes that were finally chosen for subsequent qPCR analysis.

Table 6. Overview of gene predictions by antiSMASH, clustering by FunORDER analysis and chosen target genes. Gene-IDs were derived from after gene prediction by the Joint Genome Institute[106]; Functions were assigned after BLASTP and PANNZER2 analysis; put … putative; DUF … domain of unknown function; hyp. … hypothetical protein; PKS … Polyketide synthase; TF … Transcription Factor; Green and Blue clustered groups refer to the analysis results of the FunORDER tool displayed in section 7.3.

Gene-ID	Function	FunORDER – clustered groups	antiSMASH – type	Chosen for qPCR analysis
105220	hyp. protein – DUF3328	-	other	No
105221	hyp. protein – DUF3328	-	other	Yes
105222	hyp. protein	-	other	Yes
105223	put. peptidase S41 fam	Green	other	Yes
105224	hyp. protein	Green	other	No
120872	put. GCN5-acetlytransferase	Green	other	Yes
58285	PKS	Blue	core biosyn.	already analyzed
58289	put. reductase	Blue	additional biosyn.	Yes
76204	put. oxidoreductase	Green	additional biosyn.	Yes
58953	put. P450 monooxygenase	Blue	core biosyn.	Yes
72993	put. TF	-	other	No
120873	glycoside hydrolase	-	other	No

The antiSMASH tool yielded 12 genes as part of the gene cluster and assigned two genes as core biosynthetic genes (58285 and 58953) as well as two genes as additional biosynthetic genes (58289 and 76204).

FunORDER analysis predicted evolutionary background between 10 genes in this BGC. Four genes (105223, 76204, 120872 and 105224) clustered together in the PCA as well as in the wards minimum distance dendrogram and sub summarized as "Green clustered" in Table 6 and Figure 23. Another three genes (58285, 58953 and 58289) were also clustered together in the PCA and wards minimum distance dendrogram and sub summarized as "Blue clustered" (see Table 6 and Figure 23). The three genes where no relationship with the FunORDER tool has been established were 58601, 120873 and 105222. The assignment of colors (Green and Blue) to the clustered genes of the FunORDER analysis were based on the figures presented in the supplementary that are prepared in section 7.3.

Gene 58601was already excluded to be an actual part of cluster 21 by ClusterFinder within the antiSMASH analysis (see Figure 2/Cluster 21). The other two excluded ones were 120873 and 105222. Furthermore, 58601 as well as 105222 were classified by BLASTP analysis as hypothetical proteins with a domain of unknown function 3328 (DUF_3328) and 120873 as

glycoside hydrolase. There are several reasons for the FunORDER tool to not cluster genes that were associated with the BGC by other bioinformatic tools, ranging from identifying these genes as pseudogenes, or these genes having no known homologs in the searched database or not being able to calculate the phylogenetic tree[115]. For a detailed explanation please refer to the original publication[115]. The FunORDER tool predicted two likely co-evolved clades of genes (strict distance ≤ 0.7 – "Green and Blue" clustered). The clustered genes, that are sub summarized as "Blue" in Table 6, consisted of the genes 58285 PKS-NRPS, 58289 and 58953. 58953 was assigned as core biosynthetic gene by antiSMASH and BLASTP as well a PANNZER2 analysis revealed that its most likely a P450 cytochrome like enzyme. 58289 was assigned as additional biosynthetic gene by antiSMASH and PANNZER2/BLASTP analysis suggested that this protein was most likely an enoyl-reductase. These three genes were chosen for transcript analysis because of their possible importance for the formation and modification of the resulting metabolite and their predicted clustering.

The other clustered genes by FunORDER, that are sub summarized as "Green" in Table 6, consisted of genes 105223, 105224, 120872 and 76204. 76204 was identified by BLASTP analysis as putative FMN/NADPH-oxidoreductase and antiSMASH assigned this gene to be an additional biosynthetic gene. The other three (105223, 105224 and 120872) weren't assigned specific tasks by antiSMASH analysis but BLASTP analysis revealed that 105223 might be a peptidase (S41 family), 105224 a putative PTH11-GPCR (membrane integrated protein with G-protein coupled receptor[117]) and 120872 a putative acetyl transferase. 105221, which wasn't clustered with the previous described genes by FunORDER and 105222, a putative hypothetical protein, which also wasn't clustered by FunORDER, were still chosen to be analyzed by qPCR together with 105223, 76204 and 120872. 105221 contains a DUF-3328 domain (domain of unknown function) which was interesting because proteins with these domains appear in the regulation of the long known[118] but only recently standardized[119] classes of RIPPs[120]. 105223, the peptidase and 120872, an acetyltransferase, were speculated to be potentially involved in post formational modification of the derived SM and therefore also selected. 58601 and 120873 were not considered in the following qPCR analysis due to them both being non-clustered by FunORDER analysis.

The chosen genes were measured via qPCR from the same cDNAs samples that were generated for the estimation of the PKS of this cluster (see 3.3) Figure 12 shows the transcript level estimations from these selected genes in both types of transformants (OE and syn).

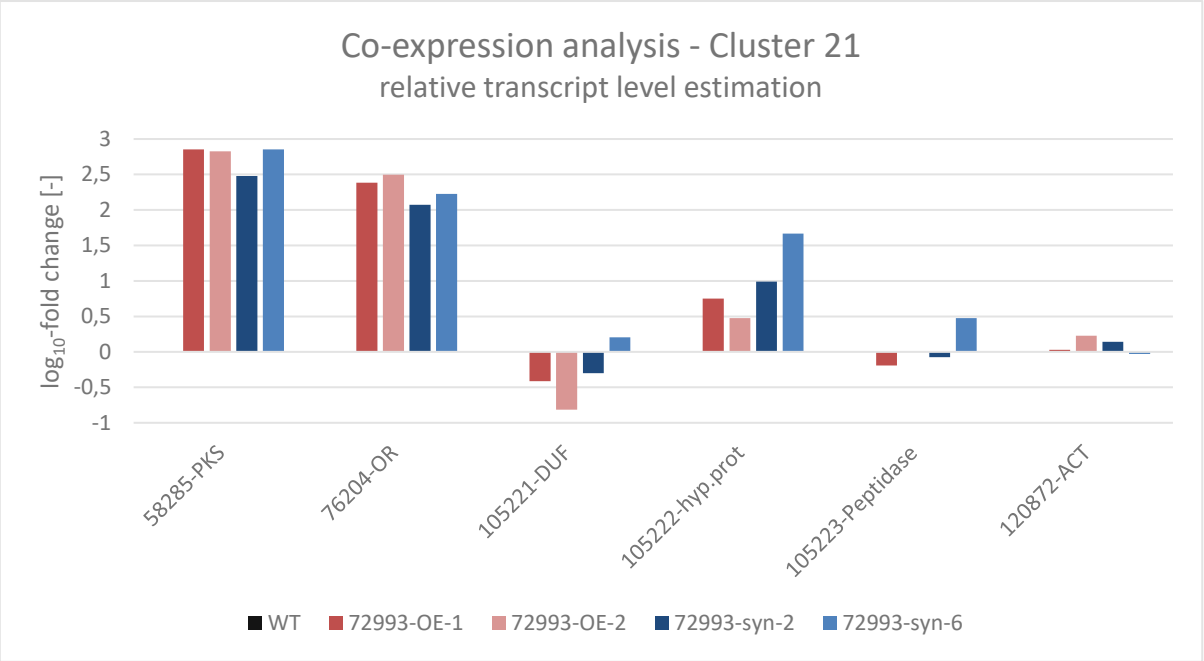

Figure 12. Co-expression analysis - BGC 21
WT … *T. reesei* QM6a *Δtmus53;* OE … overexpression cassette transformants; syn … synTFs-cassette transformants; PKS … Polyketide Synthase; OR … Oxidoreductase; DUF … domain of unknown function; hyp. prot … hypothetical protein; ACT … acetyl transferase;
Transformants and WT have been cultivated on glucose as sole carbon source. Samples have been taken after 72 hours. Relative transcript levels of the indicated genes have been normalized to the WT samples using the reference genes *act1* & *sar1*. Mean values of two replicates are given.

The putative oxidoreductase (76204) transcript levels were as similarly high in both transformant types (OE & syn), as the PKS (58285) of this cluster, ranging from 2.07 to 2.49 log_{10}-fold compared to the parental strain (WT). This makes sense because the predicted functions of these genes suggest that they must play an important role in the formation of the SM(s) of this cluster. The other core - (58953 – putative oxidoreductase) and additional biosynthetic gene (58289 – putative P45 like cytochrome) transcript levels were measured but the relative expression could not be calculated. An explanation why that was the case is following: Before calculating the relative transcript levels, one must inspect the raw data of the transcript measurements of the putative P450 like cytochrome. This data is prepared in the supplementary section (see 7.4) of this work in Table 18.

While on the first look the raw data seems solid there was a seemingly minor, yet big problem with it: The measured C_t-values of the technical replicates of 58289s transcript levels in the transformants (both OE and syn) were close to each other (e.g. 21.5/21.4). This was similarly good for the average C_t-values (21.5/20 and 22.6/22.2), efficiencies (1.71/1.72) and the no template control (NTC - water) did not yield a measurable quantitative fluorescence signal at all, which was expected (see efficiencies Table 18).

On the other hand, when taking a closer look on the WT-sample, one technical replicate with expectable signals for C_t-value and the efficiency (31.5/1.78) was measured and one technical replicate with non-utilizable C_t-value and efficiency (red data in Table 18- 11.1/0). Because there was no third technical replicate, it wouldn't be serious to perform the transcript level calculation with only the one plausible measured WT sample (31.5/1.78). If now though the raw C_t-values from 58289 in both transformants (OE and syn) and only the plausible WT C_t-value were compared to each other (20/22.2 vs. 31.5), then it was clearly visible that the putative oxidoreductases 58289 transcript levels were higher in the transformant and the data suggested that this gene played a part in the biosynthesis of this cluster's metabolite. Furthermore, if theoretically the $\Delta\Delta C_t$-value would have been calculated from these values then the $\log_{10}$-fold change would be ranging between 2.19 and 2.67. These $\log_{10}$-fold changes would fit and made sense in the perspective of the prediction that this gene was assigned as additional biosynthetic genes and should be highly co-expressed too in the activated cluster. Another completely different possibility would be that the WT transcript levels of 58289 were generally undetectable in the qPCR with good efficiencies. Although, this is a theoretical possibility, the WT sample with the C_t-value of 31.5 with an efficiency of 1.78 suggests that this was not the likeliest explanation here. This occurred also for gene 58953, the putative oxidoreductase, with similarly different Ct-value differences between the transformants and the WT sample. However, these relative transcript level calculations were not conducted and displayed to obtain the integrity of the other results of this study.

The other analyzed genes (105221, 105222, 105223 and 120872) showed a different picture:

Gene 105221 transcript levels were lower in both OE transformants and in 72993-syn-2 but slightly higher in 72993-syn-6 compared to the WTs transcripts of 105221, which indicated that this gene was not co-expressed in this BGC.

The transcript levels for the hypothetical protein 105222 were higher in both transformant types (OE & syn) but in three transformants only minorly (< 1.0 $\log_{10}$-fold) compared to the WT. Although showing higher transcript levels (1.66 $\log_{10}$-fold) in one syn transformant (72993-syn-6) compared to the WT, no assumptions regarding co-expression of this gene were made due to the high deviation between both syn transformants (1.66 vs. 0.99 $\log_{10}$-fold) and even smaller regulation difference in the OE transformants (0.75 vs. 0.44 $\log_{10}$-fold).

105223 transcript levels showed mixed regulation signals in the OE and syn transformants. The putative peptidases levels were slightly higher in each one OE (72993-OE-2) and syn (72993-syn-2) and slightly lower in the other two transformants (72993-OE-1 and 72993-syn-6) indicating that this gene was not co-expressed, thus suggesting that 105223 has no to little

function in the formation or posttranslational processing of the corresponding SM(s) in this cluster.

The last analyzed gene's (120872) transcript levels showed only minor regulation differences in OE and syn transformants ranging from -0.03 to 0.27 $\log_{10}$-fold. The putative function of this gene was to code for an acetyltransferase. As there is close to none difference in regulation compared to the WT it was, similarly to 105223, considered that 120872 was likely to play no active part in the formation or the posttranslational processing of the SM. An overview of the results of the co-expression analysis is given in Table 7.

Table 7. Overview of Co-expression analysis of cluster 21.
Mixed … transformants with the same expression cassette inserted showed once higher, once lower transcript levels than the WT; Not calculated but measured … transcript levels were measured but relative expression could not be calculated due to erroneous raw data of the WT sample

Gene-ID	Function	FunORDER – clustered groups	antiSMASH – type	Transcript levels compared to WT
105221	hyp. protein – DUF3328	-	other	Mixed
105222	hyp. protein	-	other	Strong deviation
105223	put. peptidase S41 fam	Green	other	Mixed
76204	put. oxidoreductase	Green	additional biosyn.	Higher
120872	put. GCN5-acetlytransferase	Green	other	Mixed
58285	PKS	Blue	core biosyn.	already analyzed - Higher
58289	put. reductase	Blue	additional biosyn.	Not calculated but measured – Higher
58953	put. P450 Monooxygenase	Blue	core biosyn.	Not calculated but measured – Higher

Overall the analysis of all the predicted core and additional biosynthetic genes (58285, 58289, 58953 and 76204) by antiSMASH showed the expected higher transcript levels in both OE and syn transformants. FunORDER did cluster 58285, 58289 and 58953 together in the wards dendrogram and the PCA as well as in the heatmap indicating strong likelihood of co-evolution for these three. As both bioinformatic tools results overlap with the qPCR analysis, this is another strong indicator that the tools predictions were correct regarding these three genes.

The genes that were analyzed but assigned no specific function by antiSMASH were not as clearly to interpret: While three (105221, 105223 and 120872) of those did show mixed

regulation signals the other one's (105222) standard deviation was too strong. The results pointed out that the prediction of the antiSMASH was correct regarding clusters assignment of the core and biosynthetic genes but incorrect by assigning these as "other" labelled genes to be an actual part of the BGC by the inbuilt ClusterFinder algorithm. This results were not surprising when one considers that not all predicted genes of a cluster by bioinformatic tools prove to be an actual part of the cluster or important (refer to methods of the used algorithms and results regarding training sets of well understood BGCs[63,107,121]). In order to verify the whole antiSMASH predictions, all genes have to be analyzed by qPCR or RNAseq.

3.3.1.2. Cluster 22

The co-expression analysis of Cluster 22 proved to be trickier than 21 because the cluster inherited one PKS (105804) and one hybrid PKS-NRPS (59315). The hybrid PKS-NRPS 59315 transcript levels were higher in the syn-transformants (105805-syn-2/6) while the transcripts of 105804 showed no mentionable regulation difference in all tested transformants. The cluster is schematically shown in Figure 13.

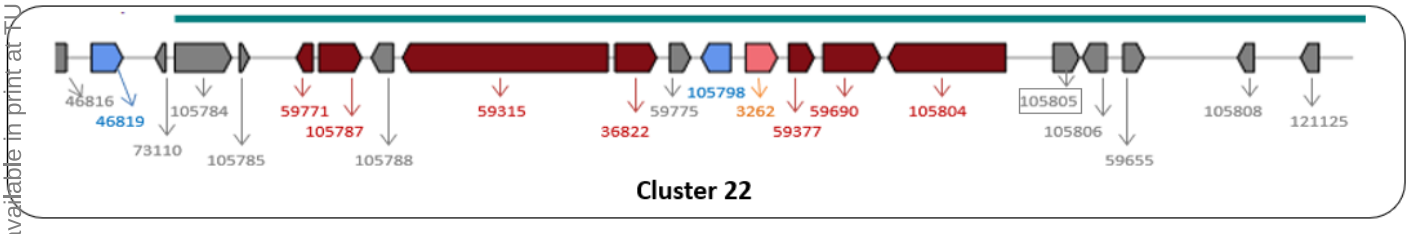

Figure 13. Cluster overview of cluster 22.
Numbers indicate the gene numbers/protein IDs assigned to these sequences after gene prediction by the Joint Genome Institute[106]; Grey arrows … other genes; Blue arrows … transport related genes; Red arrows … core biosynthetic genes; Pink arrow … additional biosynthetic gene; Grey box … putative transcription factors of corresponding cluster; Blue/greenish bar … ClusterFinder[107] association of genes that are predicted to be an actual part of the cluster;

Table 8 gives an overview of predictions both applied algorithms made and the genes that were finally chosen for subsequent qPCR analysis. The detailed results of the FunORDER analysis are given in the supplemental section of this work (see 7.3)

Table 8. Overview of gene predictions by antiSMASH, clustering by FunORDER analysis and chosen target genes. Gene-IDs were derived from after gene prediction by the Joint Genome Institute[106]; Functions were assigned after BLASTP and PANNZER2 analysis; put … putative; hyp. … hypothetical protein; PKS … Polyketide synthase; unchar … uncharacterized; MFS … Major Facilitator Superfamily; Green, Red and Blue clustered groups refer to the analysis results of the FunORDER tool displayed in section 7.3.

Gene-ID	Function	FunORDER – clustered groups	antiSMASH – type	Chosen for qPCR analysis
46816	glycoside hydrolase	-	other	No
46819	put. substrate transporter	Red	transport-related	No
73110	acetyltransferase	Red	other	No
105784	put. TF	-	other	No
105785	unchar. protein	Red	other	No
59771	put. prenyltransferase	Red & Blue	core biosyn.	Yes
105787	put. halogenase	Red	core biosyn.	No
105788	put. protease	Red	other	No
59315	put. PKS-NRPS	Green	core biosyn.	already analyzed
36822	put. hydrolase	Green	core biosyn.	Yes
59775	put. reductase	Blue	other	Yes
105798	MFS transporter	Green	transport-related	Yes
3262	aldehyde dehydrogenase	Green	additional biosyn.	Yes
59377	put. P450 Monooxygenase	Green	core biosyn.	Yes
59690	put. oxidoreductase	Blue	core biosyn.	Yes
105804	PKS	Red	core biosyn.	already analyzed
105805	put. TF	-	other	No
105806	put. acetyltransferase	-	other	No
59655	put. methyltransferase	Green	other	Yes
105808	unchar. Protein	Red	other	No
121125	hyp. protein	Red	other	No

antiSMASH analysis yielded 21 genes to be part of this BGC, whereas the FunORDER tool clustered 20 genes in this BGC. The one which was missing in FunORDER's clustering was 46816, a glycoside hydrolase. ClusterFinder[107] analysis did exclude 46816 already from being an actual part of this cluster as well as 46819 and 73110. Therefore, it was decided to not analyze these three genes by qPCR in respective of limited resources.

The dendrogram of strict distances (Figure 24/B) revealed that six genes were clustered as likely co-evolved by the FunORDER tool (labelled Green in Table 8). 36822 a putative hydrolase,

105798 a major facility superfamily transporter, 59315 the PKS-NRPS, 59655 a putative methyltransferase, 3262 a succinate semialdehyde dehydrogenase and 59377 a putative cytochrome P450 monooxygenase. Three of these were classified by antiSMASH analysis as core biosynthetic genes (36822, 59315 and 59377), one as additional biosynthetic gene (3262), one as transport related gene (105798) and the last one wasn't classified in a specific category (59655 – other genes) although BLASTP analysis revealed that this gene inherits the amino acid sequence of a conserved methyltransferase domain. These six genes were chosen to be subsequently analyzed by qPCR because of the expected upregulation regarding their predicted role. Notably, the TF-105805, that was chosen for the activation of this cluster, was clustered together with the former mentioned genes in the heat map (Green circle in Figure 24/A) but not in the Wards plot (see Figure 24/B) nor in the PCA (see Figure 24/C). It was decided to analyze the "Blue" clustered genes from Table 8 (59690, 59775 and 59771) because two of these genes were predicted by antiSMASH analysis to be core biosynthetic genes (59771, a putative prenyltransferase and 59690, a putative oxidoreductase) and the third (59775) by BLASTP analysis as putative NADPH-binding protein with a conserved reductase domain. The PCA plot of the FunORDER analysis (see Figure 24/C) showed, besides the clustering of the Green clustered genes, that additional nine genes could also be sub summarized by a reduction of dimensions (Red in Table 8) indicating that this cluster might be in fact two clusters that merged. This assumption will be further explained in section 4.3.

Cluster 22 was elucidated in syn transformants only as they were the only ones in which the PKS-NRPS 59315 showed higher transcript levels (see Figure 10/B) compared to the WT. The qPCR results of the selected target genes are displayed in Figure 14.

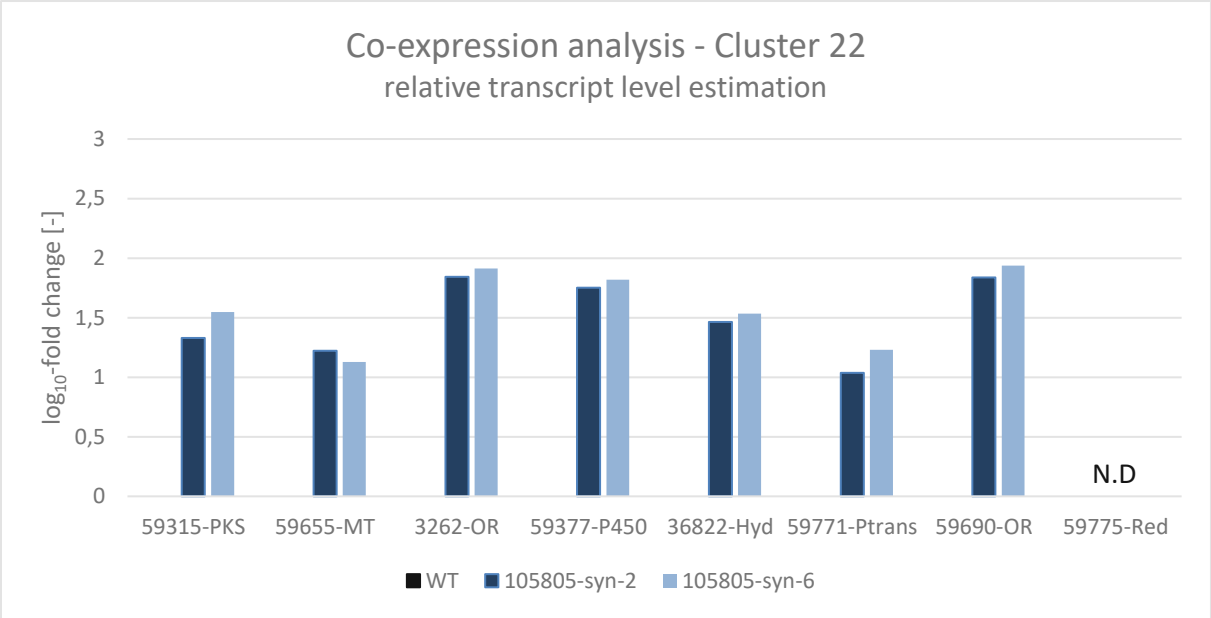

Figure 14. Co-expression analysis – BGC 22.
WT … *T. reesei* QM6a *Δtmus53*; syn … synTFs-cassette transformants; PKS … Polyketide synthase; MT … methyltransferase; OR … oxidoreductase; P450 … Cytochrome P450 like enzyme; Hyd … hydrolase; PTrans … prenyltransferase; N.D … Not Detected
Transformants and WT have been cultivated on glucose as sole carbon source. Samples have been taken after 72 hours. Relative transcript levels of the indicated genes have been normalized to the WT samples using the reference genes *act1* & *sar1*. Mean values of two replicates are given.

Four of the Green clustered genes (59655, 3262, 59377 and 36822) showed higher transcript levels in the syn transformants ranging between 1.03 and 1.91 log_{10}-fold compared to the parental strain. This was expected because these four genes were either identified as core or additional biosynthetic genes. The only Green clustered gene that couldn't be analyzed was 105798, the transport-related one. The transcripts of 105798 were measured and the raw data suggested (C_t-values: 21.7/21.0) that this gene transcript levels were higher compared to the WT sample (Ct-value: 33.6) but only one WT replicate yielded a utilizable signal and therefore the relative transcript levels could not be calculated.

Furthermore, 59771 - the putative prenyltransferase and 59690 the putative oxidoreductase transcript levels were also higher in both the syn transformants compared to the parental strain. The putative prenyltransferase 59771 moderately, ranging between 1.01 - 1.25 log_{10}-fold and the putative oxidoreductase 59690 between 1.75 - 1.80 log_{10}-fold. 59775, a putative reductase, was not detectable (N.D). An overview of the results of the co-expression analysis is given in Table 9.

Table 9. Overview of Co-expression analysis of cluster 22.
Not calculated but measured … transcript levels were measured but relative expression could not be calculated due to erroneous raw data of the WT sample; N.D … Not detectable – no measured transcripts

Gene-ID	Function	FunORDER – clustered groups	antiSMASH – type	Transcript levels compared to WT
59771	put. prenyltransferase	Red & Blue	core biosyn.	Higher
105804	PKS	Red	core biosyn.	already analyzed - Mixed
59775	put. reductase	Blue	other	N.D.
59690	put. oxidoreductase	Blue	core biosyn.	Higher
59315	put. PKS-NRPS	Green	core biosyn.	already analyzed - Higher
59377	put. P450 Monooxygenase	Green	core biosyn.	Higher
36822	put. hydrolase	Green	core biosyn.	Higher
3262	aldehyde dehydrogenase	Green	additional biosyn.	Higher
105798	MFS transporter	Green	transport-related	Not calculated but measured – Higher
59655	put. methyltransferase	Green	other	Higher

The Green clustered genes (36822, 105798, 3262, 59377 and 59655) measured transcripts (see Table 9) were all higher in the tested transformants compared to the wildtype sample, which agreed with antiSMASH analysis. antiSMASH assigned three genes to be core (36822, 59315 and 59377), one to be an additional (3262) and one to be a transport-related gene and therefore, also essential. Interestingly, 59655 the putative methyltransferase, was clustered together with the other core and biosynthetic genes as "Green" in the FunORDER's PCA and dendrogram results but was assigned no core or biosynthetic function by antiSMASH.

On the other hand, genes 59771 and 59690 (Blue clustered group/Table 9) weren't clustered together with the Green group (36822, 105798, 59315, 59655, 3262 and 59377) by FunORDER but still had higher transcripts measured than the wildtype. 59690, a putative oxidoreductase and 59771, a putative prenyltransferase, were both assigned core biosynthetic functions by antiSMASH analysis which matched the qPCR results. These two examples showed why it's always useful to apply more than one bioinformatic analysis and combine their results.

3.3.1.3. Cluster 78

The last cluster that was subject to co-expression analysis was cluster 78. Therefore, an antiSMASH, FunORDER, BLASTP and PANNZER2 analysis was conducted. The cluster is displayed schematically in Figure 15.

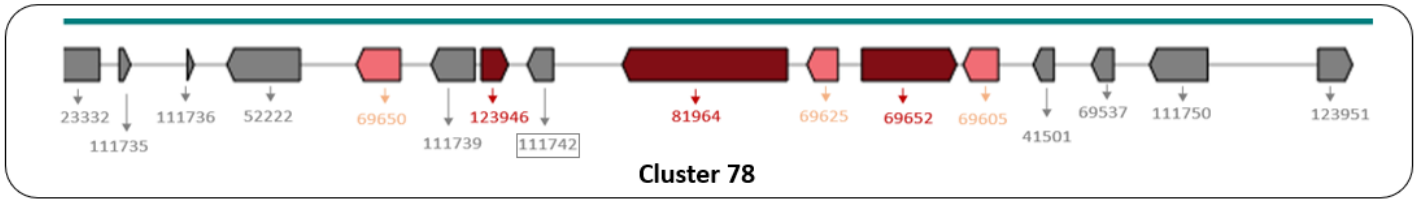

Figure 15. Cluster overview of cluster 22.
Numbers indicate the gene numbers/protein IDs assigned to these sequences after gene prediction by the Joint Genome Institute[106]; Grey arrows … other genes; Red arrows … core biosynthetic genes; Pink arrows … additional biosynthetic gene; Grey box … putative transcription factors of corresponding cluster; Blue/greenish bar … ClusterFinder[107] association of genes that are predicted to be an actual part of the cluster;

Table 10 gives an overview of predictions both applied algorithms made and the genes that were finally chosen for subsequent qPCR analysis. The detailed results of the FunORDER analysis are given in the supplemental section of this work (see 7.3)

Table 10. Overview of gene predictions by antiSMASH, clustering by FunORDER analysis and chosen target genes. Gene-IDs were derived from after gene prediction by the Joint Genome Institute; Functions were assigned after BLASTP and PANNZER2 analysis; put … putative; PKS … Polyketide synthase; unchar … uncharacterized; ER … endoplasmic reticulum; Green clustered group refers to the analysis results of the FunORDER tool displayed in section 7.3.

Gene-ID	Function	FunORDER – clustered group	antiSMASH – type	Chosen for qPCR analysis
23332	unchar. protein	-	other	No
111735	unchar. protein	-	other	No
111736	unchar. protein	-	other	No
52222	unchar. protein	-	other	Yes
69650	put. FAD-binding dehydrogenase/OR	Green	additional biosyn.	Yes
111739	put. regulatory protein	Green	other	No
81964	PKS	Green	core biosyn	already analyzed
69625	put. lactamase/hydrolase	Green	additional biosyn.	Yes
69652	put. methyltransferase/dehydrogenase	Green	core biosyn.	Yes
69605	lyase/cystathionine synthase	-	additional biosyn.	No
41501	unchar. protein	-	other	No
69537	unchar. protein	-	other	No
111750	reductase like transmembrane protein	-	other	No
123951	ER-membrane protein	-	other	No
123946	dehydrogenase	-	core biosyn.	No
111742	put. TF	-	other	No

FunORDER predicted co-evolutionary background for 12 genes of this BGC compared to 16 genes that were assigned to this cluster by antiSMASH analysis. The four genes between which FunORDER could not predict co-evolutionary background were 23222, 111735, 111736 and 123951. 23222, 111735 and 111736 were assigned no function by antiSMASH. 123951 was an ER membrane protein. BLASTP analysis revealed that 23222 contained a conserved HET domain and the other two 111735 and 111736 did not show any similarity other than uncharacterized protein. PANNZER2 analysis yielded the same predictions for these four genes and therefore, they were not further analyzed due to the assumed irrelevance for the BGC.

The heat map of strict distances (see Figure 25/A) showed less and weaker correlated co-evolved genes compared to cluster 22 (see Figure 24/A). While in cluster 21 only two genes (81964 and 69625 – green circle in the left corner of Figure 25/A) were showing a pairwise lower strict distance value of < 0.5, this was the case for seven pairs in cluster 22. The genes

clustered as "Green" in Table 10 were, 111739 a putative regulatory protein, 69652 a putative methyltransferase/dehydrogenase, 69650 a hypothetical protein/putative FAD-binding dehydrogenase, 69625 a putative lactamase/metallo- lactamase and 81964 the PKS of this cluster. The four last mentioned genes were also clustered together in the PCA (see Figure 25/C). antiSMASH analysis assigned 69652 to be a core biosynthetic gene and 69650 and 69625 were assigned to be additional biosynthetic genes. These three genes were chosen to be further analyzed by qPCR due the mentioned overlap from FunORDER and antiSMASH analysis and their possible importance for this cluster. 111739 wasn't assigned a specific function by antiSMASH and not clustered together with 69652, 69650 and 69625 in the PCA of FunORDER and therefore not analyzed. 52222 was chosen for qPCR analysis due to it being clustered together with the TF 111742 in the dendrogram (see Figure 25/B)

The qPCR analysis was performed with two of the four OE-transformants (111742-OE-3/4). Figure 16 shows the results.

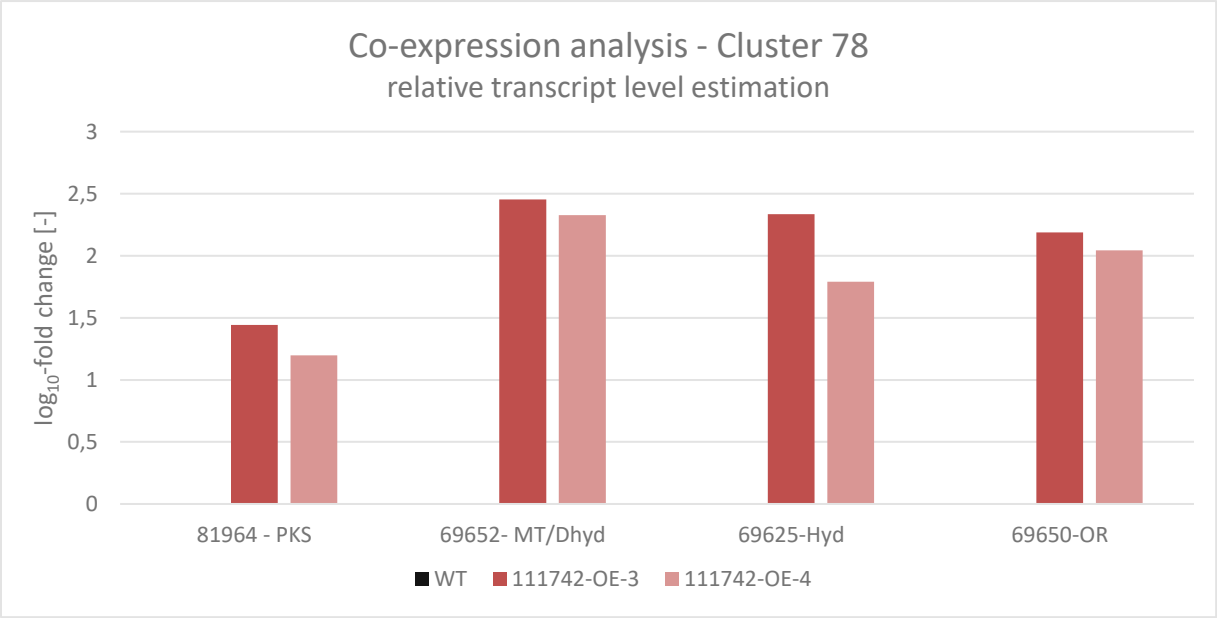

Figure 16. Co-expression analysis – BGC 78
WT … *T. reesei* QM6a *Δtmus53*; OE … overexpression-cassette transformants; PKS … Polyketide Synthase; MT … methyltransferase; OR … oxidoreductase; Hyd … Hydrolase; Dhyd … dehydrogenase
Transformants and WT have been cultivated on glucose as sole carbon source. Samples have been taken after 72 hours. Relative transcript levels of the indicated genes have been normalized to the WT samples using the reference genes *act1* & *sar1*. Mean values of two replicates are given.

The relative transcript levels of the core biosynthetic gene (putative methyltransferase/dehydrogenase 69652) were higher in both OE transformants between 2.3 and 2.4 $\log_{10}$-fold compared to the parental strain. The other two screened genes (69625 and 69650) were assigned to be additional biosynthetic genes and showed similarly high transcript levels as the core biosynthetic genes in both OE transformants. The putative lactamase (69625)

as well as the putative dehydrogenase (69650) transcript levels ranged between 1.7 and 2.3 log_{10}-fold higher compared to the parental strain. This was expected for all three genes (69652, 69625 and 69650). Another gene that was measured but yielded no quantifiable fluorescence signal was gene 52222. This gene was not assigned a function by antiSMASH and PANNZER2 analysis but BLASTP analysis showed a conserved domain of Ankyrin repeats, which aid in signal transduction between membrane proteins and cytoskeletal elements[122]. 52222 was nonetheless chosen for analysis, although it was not clustered in the same clade of the dendrogram nor in the PCA with 69652, 69625 and 69650 (see Figure 25), because this gene was in a clade together with the TF 111742 and therefore possibly co-regulated due to their likely co-evolution. The results suggest otherwise. An overview of the qPCR results is given in Table 11.

Table 11. Overview of Co-expression analysis of cluster 78.
Gene-IDs were derived from after gene prediction by the Joint Genome Institute; Functions were assigned after BLASTP and PANNZER2 analysis; put … putative; unchar … uncharacterized; OR … oxidoreductase; N.D … not detectable; Green clustered group refers to the analysis results of the FunORDER tool displayed in section 7.3.

Gene-ID	Function	FunORDER – clustered groups	antiSMASH – type	Transcript levels compared to WT
69650	put. FAD-binding dehydrogenase/OR	Green	additional biosyn.	Higher
69625	put. lactamase/hydrolase	Green	additional biosyn.	Higher
69652	put. methyltransferase/dehydrogenase	Green	core biosyn.	Higher
81964	PKS	Green	core biosyn	already analyzed - Higher
52222	unchar. protein	-	other	N.D

Four of the five Green clustered genes transcript levels (69652, 69650, 69625 and the PKS 81964) were higher in the OE transformants than in the WT sample. 69652 was assigned to be a core biosynthetic gene and the other two were assigned to be additional biosynthetic genes (69650 and 69625) by antiSMASH analysis. Interestingly, two other genes of possible importance (69605, a cystathionine synthase, and 123946, a dehydrogenase) were clustered together in the PCA (Figure 25/C) by FunORDER and assigned to be an additional and core biosynthetic gene by antiSMASH. They were clustered in the same clade adjacent to the Green clustered group (Figure 25/B) and therefore a potential interesting target but due to minor interest in this cluster for follow up experiments no resources were spent to elucidate this lead. The minor interest in this cluster resulted from the measured C_t-values of the PKS 81964: While

the OE transformants (111742-OE-1/2/3/4) C_t-values were ranging between 18.7 and 21.2, the WTs C_t-values were 24.3 which meant that there was already a relative moderate amount of transcripts of this PKS in the WT. The difference of these C_t-values wouldn't make it easy to find this expressed metabolite when searching through the metabolome in follow up experiments. In case of this cluster it would also be more correct to speak about overexpressing the cluster PKS than stating that the cluster was activated but for similarity and an easier overview, the cluster was declared as "activated".

4. Discussion

4.1. Remarks about qPCR analysis

In consideration of uncountable publications with qPCR as method, the reader might notice that there is difference in the data communication between this project and the most others. While most other publications show standard deviations between at least two (often three) biological replicates[102,123–127], they were missing here. The reason for that was that the tasks of this project were ranging from designing synTFs, over cloning, transforming and finally up to measuring the generated regulation differences and were mainly orientated to find out **if** the concept of synTFs that have proven suitable in another context[102] as well as in *A. nidulans*[98], would also work, in up to this point silent BGCs of *Trichoderma reesei* QM6a. Therefore and because of a strict time table, no biological replicates of the same transformants were measured. Another reason for the missing measurements of biological replicates were the considerable costs for the underlying goal. The Luna qPCR system from NEB is cost intensive and for our cause of finding out if the synthetic transcription factor design works in *T. reesei*'s silent BGCs, one biological sample measured in duplicates was enough.

Nonetheless the qPCR data and the statements derived from it, can be considered powerful when one recapitulates that the transformants with the same cassette inserted (e.g. OE-1 and OE-2) can be regarded as "similar-to-biological" replicates although being two separate transformants. Of course, one must bear in mind that in order to compare two transformants with the same cassette inserted correctly, one would have to make sure that only one copy of the cassette was integrated into the genomes of each transformant with a differential method, like Southern Blotting[128] but for the goal of sheer activation of a cluster that was not necessary and therefore not performed. This was also the reason why measurements in which two or more OE or syn transformants with the same cassette inserted but different regulation behavior ((e.g. one up- (122783-syn-9) and one downregulated (122783-syn-2) regarding the same gene (e.g. PKS 65172)) were judged as insufficient for a strong statement, towards cluster activation. Furthermore, the C_t-values of the raw data were compared and considered when making statements about the activation or non-activation of a cluster as well as the resulting melting curves. Samples where only one reliable C_t-value was measured (see 7.4) were therefore ruled out for interpretation, although the measured transcripts indicated a clear inclination (see Cluster 21; gene: 58289; section 3.3.1.1). These problems would have been preventable by re-

measuring the corresponding samples in technical duplicate or triplicates, but this was no option due to lack of time.

4.2. Cluster activated or not?

The results of cluster 21 were as clear as possible because the transcript levels of the PKS-NRPS 58285 were higher in all 4 transformants of both cassettes (OE & syn) and therefore, the cluster was judged as activated.

Cluster 22 yielded a different picture: While PKS 105804 transcript levels were lower or only slightly higher in any of the four transformants (105805-OE-2/3 and 105805-syn-2/6), the second core biosynthetic genes transcript levels, namely the PKS-NRPS 59315, were higher in the syn transformants (105805-syn-2/6) only compared to the WT. Although the transcript levels of PKS-NRPS 59315 were only half as high (~1.35 – 1.55 $\log_{10}$-fold/see Figure 10/A/B) in transformants 105805-syn-2/6 compared to the PKS-NRPS 58285 in transformants 72993-OE/syn (~ 2.8 $\log_{10}$-fold) of cluster 21, a glimpse on the raw data revealed the difference in transcript level expression was clearly distinguishable in both raw data sets. The measured transcripts for the PKS-NRPS 59315 could be interpreted as very weakly basal expressed in the WT (C_t-value: 28.6) compared to the transcripts of the PKS-NRPS 58285 in the WT (C_t-value: 30.8). However, the difference in C_t-values between WT and the syn transformants (105805-syn-2/6) regarding the PKS-NRPS 59315 was sufficient (WT: 28.6 vs. 22.5/21) and therefore, cluster 22 was regarded activated in the syn transformants. Additionally, it was generally considered a strong indicator for a statement towards activation, if two different, selected and cultivated transformants of the same type (OE or syn) showed the same regulation behavior in a close range (both up- or both downregulated).

There is already one reason prepared in section 3.3 why cluster 52 was not regarded activated in the respective transformants (79725- and 122783 OE and syn). Another reason why cluster 52 wasn't regarded activated was that the raw data differed from the raw data of the as activated declared clusters (21, 22 and 78). Generally speaking, the earlier the cycle in which the fluorescence signal differs significantly from the background fluorescence (C_t-Cycle threshold) in an qPCR experiment, the more transcripts are there from the start and the more robust is the data in the end. Therefore, the resilience of the statements regarding activation can be checked by comparing raw data C_t-values. While the C_t-values for all OE and syn transformants of 79725 and 122783 regarding the regulation of the PKS 65172 were between 28.3 and 32, the WTs C_t-values for the same PKS were 30.7 and 30.8. Although the calculated

$\Delta\Delta C_t$'s were still ranging between log_{10}-fold changes of 0.65 – 0.95 these results couldn't be interpreted as strongly as the regulation of PKS 81964 in cluster 21 where the C_t-values were ranging between 20.7 and 20.8 in 111742 OE transformants compared to 24.3 in the WT, resulting in nearly similar log_{10}-fold changes between 0.77 and 1.44. Without this information it might seem peculiar as to for what reason cluster 52 was regarded not activated while 78 was declared activated. Another reason why cluster 78 was considered activated while 58 was not, was that all four OE transformants (111742-OE-1/2/3/4) showed the same behavior regarding their transcript levels, namely they all had higher transcript levels of the PKS 81964 estimated compared to the WT, although it was acknowledged that regulation differences < 1.0 log_{10}-fold change were ultimately not considered sufficient (111742-OE-1/2 transformants yielded values of < 1.0). 79725 and 122783, with the exception of 122783-OE-1/3, showed mixed regulation signals compared to 11174-OE. To underline the importance of the analysis of the raw data, the C_t-values of the other transformants (72993-OE/syn and 105805-syn) regarding their corresponding PKS-NRPS were all between 19.1 and 21.9, while the WTs PKS-NRPS C_t-values were between 30.8 and 28.6. The aim of this project was to activate clusters in the transformants with the synTF or OE-cassettes in order to express an SM that should be able to be found in follow-up experiments. The search for the SM(s) would be comparable to finding a needle in hay if small regulation differences would have been considered too.

Another aspect of cluster activation was that interestingly, cluster 21 was activated in both transformant types (OE and syn) while cluster 22 was only activated in syn transformants and to further complicate the story, cluster 78 only in OE transformants. The explanation for that is maybe because of the nature of the expression-cassettes. While the plain overexpressing of a transcription factor was not sufficient for target upregulation (111742-syn transformants), Derntl et. al[102] and Grau et. al[98] showed that the usage of a xenomorphic TAD (*ypr1*' and *afoA*) can be necessary for upregulation of a target. This could be the explanation why only syn-cassettes carrying transformants (105805-syn-2/6) in cluster 22 were able to upregulate the target PKS-NRPS 59315 while OE transformants weren't. On the other hand, this would not explain, yet even contradict the result obtained for 111742 transformants carrying the overexpression-cassette: The results on this end showed that this cluster was only activated in the OE transformants. Therefore, we speculate that the presence of the TAD of the native transcription factor 111742 is utmost necessary for target regulation. This would explain why there was no upregulation of the PKS 81964 in the syn transformants. In this case the constitutive expression of the native transcription factor through the expression cassette, was key for the activation of this cluster.

4.3. Comparison of homologous gene clusters

All three clusters that were activated, have known homologous gene clusters in other organisms. The homologous clusters are documented in the MIBIG[129] database. This database allows to deposit the **M**inimum **I**nformation about a **B**iosynthetic **G**ene cluster. Each gene cluster is assigned a unique identification code and provides data like the annotation of the involved genes, the resulting product of that cluster, the publications in which these knowledges were gained and many more. These information are particularly interesting if one tries to find out what the possible product of an unknown gene cluster might look like and to roughly estimate if the following lab work is worth the effort. Cluster 21 and cluster 22 were chosen for follow-up experiments in this research group and the homologous cluster will therefore be elaborated extensively.

For cluster 21 five homologous clusters were found in the MIBIG database. A visualization of these homologous clusters is prepared in Figure 17.

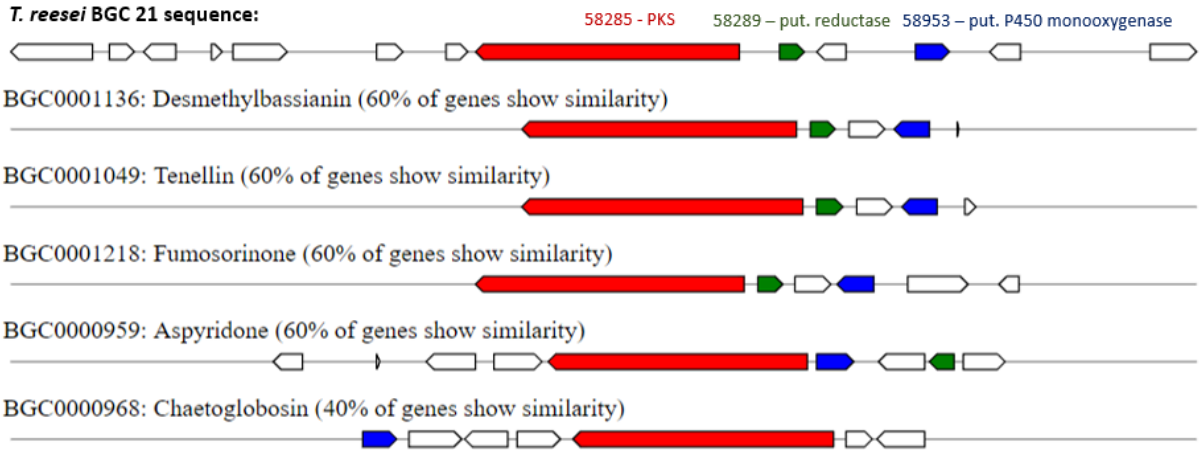

Figure 17. Overview of homologous BGCs to cluster 21.
58285 … PKS-NRPS in *T. reesei*; 58289 … put. reductase *T. reesei*; 58953 … put. P450 Monooxygenase *T. reesei*; BGCxxx … homologous clusters from MIBIG website. Desmethylbassianin, tenellin, fumosorinone, aspyridone A and chaetoglobosin A are the BGCs products.

Four of the five homologous cluster (BGC0001136, BGC0001049; BGC0001218 and BGC0000959) inherit a homolog to the PKS 58285, the putative reductase 58953 and to the putative P450 monooxygenase 58953 of the BGC 21 from *T. reesei*. Thereby, 60 % of the genes showed similarity with the genes in the homologous BGCs. The fifth known homologous BGC (BGC0000968) inherited a homolog to the PKS 58285 and to the putative P450 monooxygenase 58953. 40 % of genes showed similarity with the genes in the homologous BGC0000968. The biosynthetic gene clusters which yielded desmethylbassianin and tenellin as products have been

described in *Beauveria bassiana*[130]. Both BGCs consist of four genes: The desmethylbassianin gene cluster comprises of *dmbA*, which codes for a P450 monooxygenase, *dmbB*, which codes for a second P450 monooxygenase, *dmbC* which codes for a reductase/dehydrogenase and *dmbC* which codes for the PKS-NRPS[131]. The tenellin gene cluster comprises of *tenS*, which codes for the PKS-NRPS, *tenA* which codes for a P450 monooxygenase, *tenB* which codes for a second P450 monooxygenase and *tenC,* which codes for a reductase[132]. The fumosorinone BGC (BGC0001218) was described in *Isaria fumosorosea* and consists of six genes[133]. Four of which could be attributed importance for the formation and synthesis of fumosorinone, namely *fumoA* which codes for a P450 monooxygenase, *fumoB* which codes for the second P450 monooxygenase, *fumoC* which codes for a reductase and *fumoS* which codes for the PKS-NRPS of this cluster[133]. The aspyridone A BGC was described in *Aspergillus nidulans*[134]. It consists of eight genes: *apdA* which codes for PKS-NRPS, *apdB* which codes for a P450 monooxygenase, *apdC* which codes for a reductase, *apdD* which codes for a FAD-dependent monooxygenase, *apdE* which codes for a second P450 monooxygenase, *apdF* which codes for an exporter, *apdR* which codes for a putative Gal4-like Zn_2Cys_6 TF and *apdG* which codes for a dehydrogenase[134]. All four metabolites are 2-pyridone derivates, chemically very similar and differ only by methylation pattern and chain length. The fifth homologous gene cluster yields chaetoglobosin A and only has two homologous genes to *T. reesei's* cluster 21 genes, the PKS-NRPS encoding gene *cheA* and one putative P450 monooxygenase encoding gene *cheG*[135]. The chaetoglobosin cluster inherits another six genes, two of which code for putative TFs (*cheC* and *cheF*), another two coding for putative monooxygenases (*cheD* and *cheE*) and the last one coding for a enoyl reductase (*cheB*)[135]. An overview of all metabolites from the homologous clusters is provided in Figure 18.

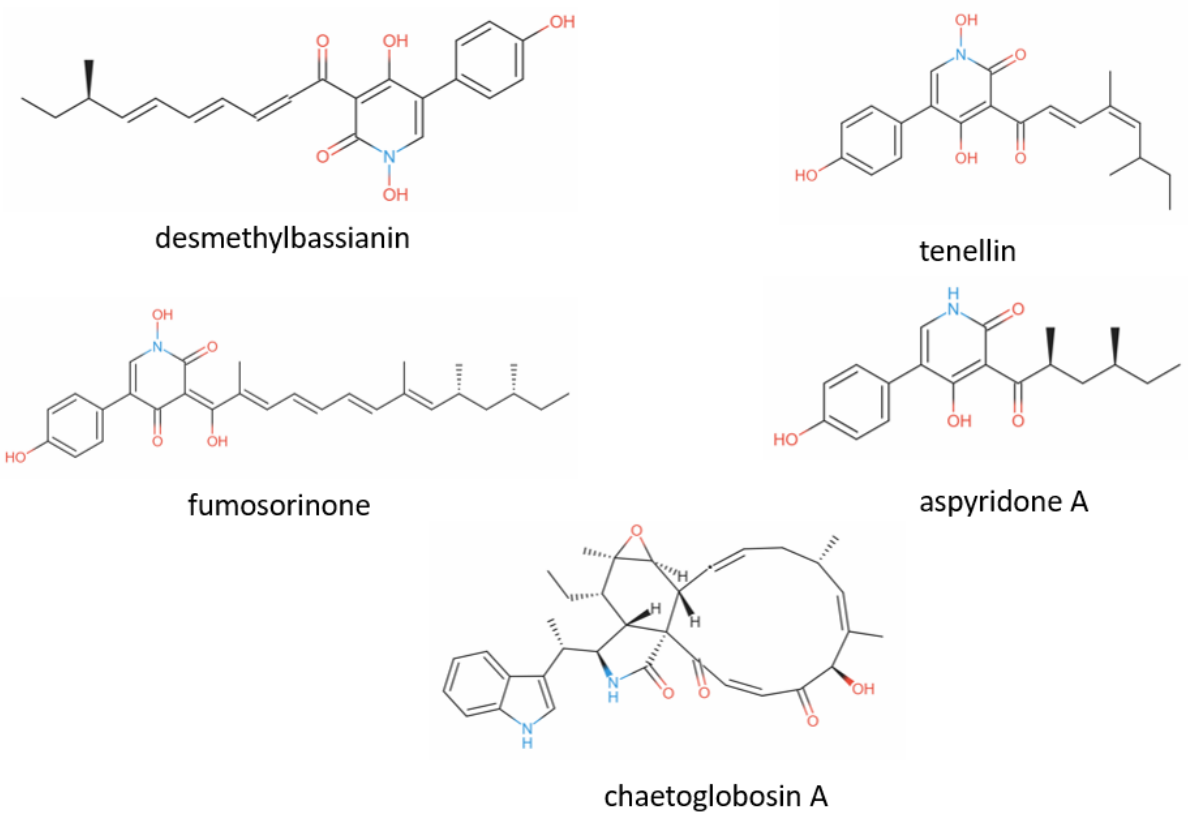

Figure 18. Metabolites of homologous gene clusters to *T. reesei*'s cluster 21.

Interestingly, gene 76204, coding for a putative oxidoreductase in *T. reesei*'s cluster 21, was not matched by the algorithms as homologous to the second P450 monooxygenases in the tenellin, desmethylbassianin, fumosorinone and aspyridone A cluster. However, it might play the role of the second P450 monooxygenase, like *dmbB, tenB, fumoB* and *apdD* do in their respective clusters. The PKS 58285, the putative reductase 58289, the putative P450 monooxygenase 58953 and the putative oxidoreductase 76204 transcript levels were all higher in the transformants than in the WT sample (see Figure 12 and Table 7). Considering this and the homology of the PKS 58285, the reductase 58289 and the P450 monooxygenase 58953 to their respective genes in the homologous clusters, we speculate that the product of the BGC 21 in *T. reesei* will most likely yield a 2-pyridone derivate. We furthermore, conclude that the resulting metabolite of *T. reesei*'s cluster 21 will most likely not be similar to chaetoglobosin A, as the chaetoglobosin A cluster inherits several more additional biosynthetic enzymes (in total three monooxygenases and one enoyl reductase) compared to the other four homologous BGCs and only the PKS 58285 and the P450 monooxygenase 58953 are homologous to two genes in this cluster. However, these assumptions represent a coarse idea about the resulting metabolite of cluster 21, as cluster 21 remains not completely investigated.

Two homologous clusters were found in the MIBIG database for cluster 22. These homologous clusters are displayed in Figure 19.

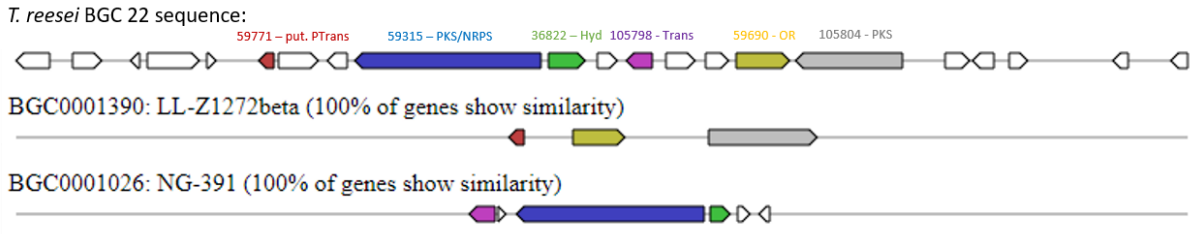

Figure 19. Overview of homologous BGCs to cluster 22.
59771 … put. prenyltransferase (PTrans) in *T. reesei*; 59315 … PKS/NRPS in *T. reesei*; 36822 … put. putative hydrolase (Hyd) in *T. reesei*; 105798 … put transporter (Trans) in *T. reesei*; 59690 … put. oxidoreductase (OR) in *T. reesei*; 105804 … PKS in *T. reesei*; BGCxxx … homologous clusters from MIBIG website: LL-Z1272beta and NG-391 are the BGCs products.

The first homologous cluster BGC0001390 has three homologous genes to *T. reesei*'s cluster 22. The PKS-NRPS 105804, the putative oxidoreductase 59690 and the putative prenyltransferase 59771. The resulting metabolite of the homologous cluster, LL-Z1272-beta was firstly isolated from unclassified *Fusarium* spp[136]. Its biosynthesis was described by Li et. al in 2016 in *Stachybotrys bisbyi* PYH05-7[137]. The cluster consists of three genes: *stbA*, the PKS, *stbB* a NRPS enzyme-like encoding gene and *stbC* a prenyltransferase[137]. The transcript analysis of PKS 105804 revealed that the transcription of PKS 105804 could not be enforced in either tested transformant (see Figure 10) and we therefore assume that the resulting metabolite of cluster 22 will more likely be similar to the metabolite of the second homologous gene cluster, BGC0001026. This second homologous gene cluster inherited five genes in which three were homologous to *T. reesei*'s cluster 22 genes. The PKS-NRPS 59315, the putative hydrolase 36822 and the putative transporter 105798. The metabolite of the homologous cluster is NG-391, a metabolite firstly isolated in *Fusarium* spp[138] then *Metarhizium anisopliae*[139]. The cluster was investigated in 2010 by Donzelli et. al in *M. robertsii* (formerly *M. anisopliae*), yielding six genes to be part of it[140]. The PKS-NRPS *ngs1* (ORF 4) was shown to be vital for the production of biosynthesis of NG-391[140]. The other five genes that were part of this BGC were identified by sequencing of an *M. robertsii* BAAC clone and bioinformatically assessed[140]. ORF3, a putative hydrolase and homologous to 36822 and ORF6, a putative MFS transporter homologous to 105798, were identified this way. All three homologous genes (59315, 36822 and 105798) transcript levels were higher in the transformants where cluster 22 was activated (see Figure 10 and Figure 12). We therefore assume that the core chemical scaffold might point towards the direction of NG-391. However, there were several more important core and tailoring genes were co-expressed in cluster 22 (see Figure 14 and Table 9) and therefore, this assumption only displays a coarse idea of the resulting metabolite rather than a solid prediction. Interestingly, both homologous genes (59771 and 59690) to the LL-Z1272-beta cluster genes,

were also highly co-expressed in cluster 22 too, underlining the vague outlook for the cluster product.

Another outcome of the co-expression analysis of cluster 22 was that the other clearly identified PKS 105804 was not showing higher transcript levels in any of the transformants compared to the WT, independently from the inserted expression-cassette (OE or syn). This led to the speculation that the inserted modified TF (105805) was not the one necessary to induce the expression of PKS 105804. Furthermore, there is a second TF in cluster 22, namely 105784, which also has a Gal4-like Zn_2Cys_6 binuclear conserved motif. It is possible that TF 105784 is the one that regulates the PKS 105804 and we therefore, speculated that cluster 22 could be two merged biosynthetic gene clusters in which PKS 105804 forms a different SM than the one resulting from PKS-NRPS 59315. That assumption was underlined by the fact that the first homologous gene cluster BGC0001390, did not only yield a completely different product than BGC0001026 but also that three different genes were homologous to each cluster (BGC0001390 – 59771, 105804 and 59690; BGC0001026 – 59315, 105798 and 36822). Another hint pointing towards the "two-submerged cluster" assumption, is that the three genes (59771, 59690 and 105804) that were 100% homologous to the genes that were described to produce LL-Z1272-beta, were clustered together as Blue group in FunORDER's PCA (see Figure 24/C). The other three genes (59315, 105798 and 36822) that were 100% homologous to the genes that were described to produce NG-391 were clustered as Green in FunORDER's PCA.

An approach to clarify this puzzle would be to generate an OE & syn-cassette with putative TF 105784 and run the whole procedure of this project with transformants carrying these cassettes. It would be interesting to find out if the assigned core biosynthetic genes that were now co-expressed in the 105805-syn transformants would also be co-expressed in the 105784-OE or syn transformants. Of course, only if the PKS 105804 would be expressed at all in these transformants. Furthermore, all assigned cluster genes could be screened, in then all four types of OE and syn transformants (105805 and 105784) to gain the biggest picture possible of this cluster.

The MIBIG database yielded four homologous gene clusters for cluster 78. An overview is given in Figure 20.

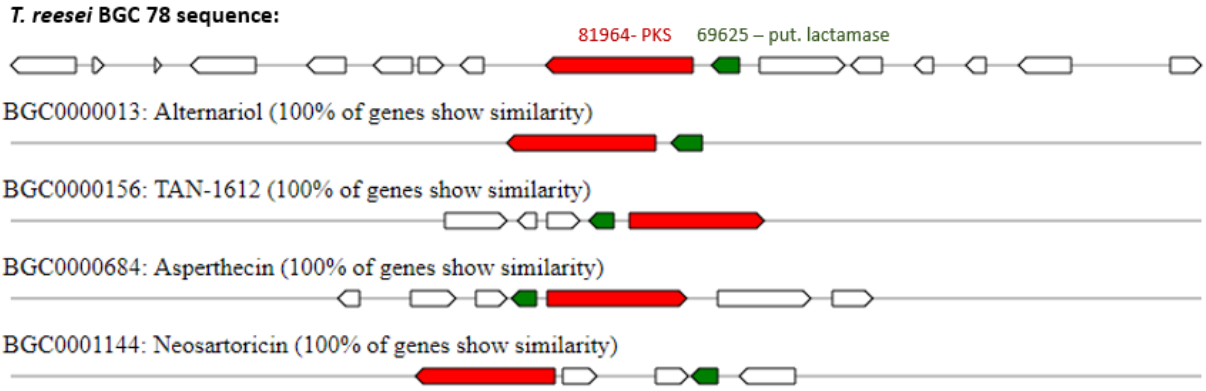

Figure 20. Overview of homologous BGCs to cluster 78.
81964 … PKS in *T. reesei*; 59315 … PKS/NRPS in *T. reesei*; 69625 … put. lactamase; BGCxxx … homologous clusters from MIBIG website: alternariol, TAN-1612, asperthecin and neosartoricin are the products of the homologous clusters.

All four homologous cluster had the same two homologous genes to cluster 78 namely, the PKS 81964 and the putative lactamase 69625. The BGC000013 that resulted alternariol in *A. nidulans* is shown in Figure 20[141]. The authors replaced the original promotor of the PKS with an inducible (*alcaA*) and this resulted in the expression and detection of alternariol[141]. They didn't investigate this locus further because their hypothesis focused on unexplored PKS in general in this *Aspergillus* strain[141]. Therefore, and because cluster 78 in *T. reesei* is comprised of several more core and additional biosynthetic genes that were co-expressed (69650, 69625 and 69652) we assumed that the cluster metabolite from cluster 78 will most likely be severe differently to alternariol. In general, the cluster co-expression analysis for cluster 78 was only performed for three genes (69650, 69625 and 69652) additionally to the PKS 81964. Although all three genes showed higher transcript levels in the tested transformants, there was still one more core (123946) and one more additional biosynthetic (69605) gene that wasn't screened due to lack of time and resources. It would be too speculative to assume the coarse chemical scaffold based on only two homologous genes in each known homologous cluster without the information of these two genes (123946 and 69605) regarding their co-expression behavior.

5. Conclusion and Outlook

In this project we demonstrated that synthetic transcription factors (synTFs), in which the modularity of TF is exploited by fusing the DBD of the gene of interest to the TAD of a different TF, can be constructed and used to activate silent or induce formerly weakly expressed biosynthetic gene clusters in the ascomycete fungus *T. reesei*. Apart from activating three clusters, the biggest achievement, especially in the light of future projects, was that these cluster activations were achieved after cultivation on a cost-efficient carbon source (D-glucose) on minimal medium (see 6.6.1). Future groups that will make use of this approach will not have to experiment with several substrates and incubation durations or methods in order to induce silent BGCs. This will lead to more time-efficient research as well as cost-limiting factor in several research groups and therefore, enable researchers to focus on more important aspects in their projects.

The co-expression analyses revealed that the activated clusters were in fact induced and showed that the most bioinformatic predictions were useful and accurate. However, additional experiments will have to be performed, to determine the boundaries of each activated BGC and to suggest the biosynthesis pathway. More steps to come include the identification of the expressed SM(s), the isolation and the characterization. A glimpse of one these approaches will be presented in 7.5.

Future prospects include the potentially useful properties of the, then identified, SM(s) and the possible utilization of this product.

6. Material and Methods

6.1. *In silico* analyses

Genomic data of *T. reesei* QM6a[105] was used to identify BGCs with antiSMASH analyses in its web server version 4.3. Four out of 91 clusters have been chosen to be further examined. In these four clusters the putative TFs were identified using the BLASTP search tool and PANNZER2 analysis. In activated clusters all genes were subject to PANNZER2 and BLASTP analysis. The numbers that were used to described the genes (e.g. 105805 - TF), were the gene-IDs/protein-IDs, assigned to these genes by the Joint Genome Institute (JGI; https://mycocosm.jgi.doe.gov/Trire2/Trire2.home.html; Department of Energy, USA) after gene prediction[106]. Furthermore, the activated clusters were subject to analysis with FunOrder tool[115].

6.2. Design of synthetic transcription factors

Table 12 displays the chosen clusters and corresponding putative TFs with the respective length of the truncated DBD (DBD').

Table 12. Overview of chosen clusters and corresponding transcription factors.

Cluster	putative Transcription factor	DBD'	Length (bp)
21	72993	1-N242	807
22	105805	1-E184	553
52	122783	1-A119	454
	79725	1-R125	377
78	111742	1-Q133	401

Figure 21 shows how the synthetic transcription factors (synTFs) were designed:

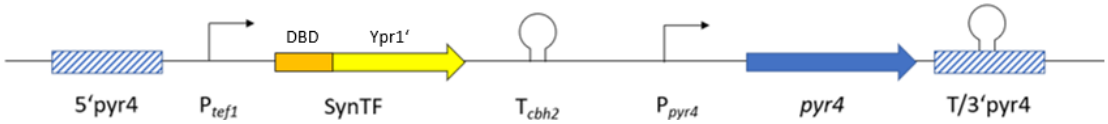

Figure 21. Construction of the synthetic transcription factors (synTFs).
5'pyr4 – 5' homologous flank of *pyr4* locus; P*tef1* – promotor of *tef1;* DBD – DNA-binding domain of truncated TF; *Ypr1'* – truncated Yellow pigment repressor; T*cbh2* – terminator *cbh2*; P*pyr4* – promotor *pyr4; pyr4* – gene;

The aim was to reinsert the synTF-cassettes into the *pyr4* locus from *T. reesei*. Therefore a homologous recombination strategy[113] was pursued by constructing approximately 1000 bp homologous flank regions 5'-upstream from the synTF and 3'-downstream of the reinserted *pyr4* gene. The promotor of *tef1* (translation elongation factor 1-alpha) was chosen for strong constitutive expression and the terminator from *cbh2* (cellobiohydrolase II) for strong termination. The synTFs were constructed as a truncated version of the TF, mainly the truncated DNA-binding domain fused to a truncated transactivation domain from the Yellow pigment repressor (Ypr1) from *T. reesei*. This TAD' was chosen because of its ability to activate genes nearly carbon source independently[102] and the plasmid pRP4_TX that inherited all the above-mentioned inserts, only lacking the synTF, was constructed prior to this work in this research group[102].

Furthermore, overexpression-cassettes (OE) have been constructed additionally to the synthetic transcription factor cassettes. The structure is identical to the synTFs except that the whole gene of the transcription factors and not a truncated version of it, was inserted between the *tef1* promotor and the *cbh2* terminator. The design is displayed in Figure 22.

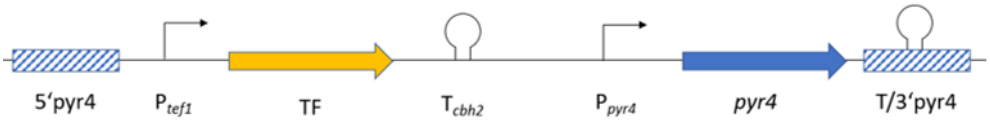

Figure 22. Overexpression-cassette (OE) design.
5'pyr4 – 5' homologous flank of *pyr4* locus; P*tef1* – promotor of *tef1*; TF – transcription factor; T*cbh2* – terminator *cbh2*; P*pyr4* – promotor *pyr4; pyr4* – gene;

These OE-cassettes have been designed and generated previously by colleagues in this research group but were analyzed in this project.

6.3. Cloning

The truncated versions of the transcription factors were amplified using Q5®-High Fidelity-Polymerase (NEB, Ipswich, MA, USA, #M0491L) in a 50 µl standard protocol approach according to the manufacturer. The primers that were used are listed in Table 13. PCR-products were then ligated into pJET1.2 (Thermo Fisher Scientific Inc., Waltham, MA, United States; CloneJET PCR Cloning Kit #K1231,) with a ratio of 1:2, applying 200 U T4 DNA ligase. Ligated pJET1.2 was transformed into *E. coli* Top10 (Thermo Fisher Scientific Inc.) competent cells and selected on Lsyogeny Broth medium[142] (LB - see Table 17) plates with 0.1 % (v/v) of

0.1 g/L ampicillin as selection pressure. Plasmids were isolated after incubation of transformed *E. coli* cells at 37 °C overnight (O/N) with QIAGEN® Plasmid Mini Kit 25 (Hilden, Germany). Isolated plasmid was then sequenced at Microsynth AG (Balgach, Switzerland).

The transformation protocol into *E. coli's* Top 10 cells consisted of the following steps and was performed for transformations of vectors pJET1.2, pCD_pXY1 and pRP4_TX.

1. Incubate ligation reaction for 10 minutes at room temperature (RT) after mixing the ligase with the digested plasmid and respective DNA.
2. Addition of 100 µl *E. coli* Top 10 cells and immediate transfer of reaction tubes on ice for 40 minutes.
3. Heat shock application to the reaction tube at 42 °C for 2 minutes
4. Addition of 500 µl LB medium and transfer of reaction tube to 37 °C oven for 40 minutes.
5. Plating of 200 µl reaction suspension onto LB plates with respective selection pressure and incubation of plates for 1-2 days at 37 °C.

After verification of correct amplification, 30 µg of pJET1.2 were subject to 2 hours restriction digestion at 30 °C with 10 U of AflII & MunI separation on an 1.5 % agarose gel (solved in 1X TAE) for 50 minutes at 100 V, followed by gel elution (NEB Monarch® Gel Extraction Kit #T1020L). The DBD' were cloned into AflII and SpeI-digested pCD_pXY1 (see 7.1) plasmid with a 3:1 ratio (synTF:vector), using 200 U T4 DNA Ligase. Ligated pCD_pXY1 was then transformed into *E. coli* Top 10 competent cells and selected on LB-plates with 0.1 % (v/v) of 0.1 g/L spectinomycin as selection pressure. Transformed *E. coli* colonies were subject to verification of the correct insert with Colony-PCR and test restriction digest.
In the following 30 µg pCD_pXY1 were digested with 10 U AflII & SpeI, releasing the assembled synTFs (DBD' + Ypr1') for the final ligation into pRP4_TX. The synTFs were ligated with a 3:1 ratio (synTF:vector) using 200 U T4 DNA Ligase (NEB) and transformed into *E. coli* Top10 cells. Colonies were checked for the correct insert via Colony-PCR and test restriction digest. Both were carried out as described in the following (also for previously mentioned steps):
Colony-PCR was carried out with one colony each and OneTaq-DNA Polymerase (NEB, #M0509L) in a 25 µl standard protocol approach according to the manufacturers protocol. Test restrictions were carried out by digesting 1 µg plasmid with 5 U AflII & SpeI for 90 minutes at 37 °C. Colony-PCR products and test restriction digests were separated on an 1.5 % agarose

gel (solved in 1X TAE and 100X SybrSafe™ stain, Thermo Fisher Scientific Inc.) at 100 V for 30 minutes. Bands were highlighted with BioRad® Gel Doc XR with Image Lab Software.
All primer sequences that were used in this project are summarized in the supplementary section (see 7.1). All restriction digestions in this project have been performed with enzymes from New England Biolabs® (NEB Ipswich, MA, USA).

6.4. Fungal Transformation

6.4.1. Fungal strain and cultivation

Trichoderma reesei QM6a *Δtmus53 Δpyr4* was used in this project[143]. The strain was kept on malt extract (MEX) plates and 5 mM uridine were added where applicable. From these plates a light green spore suspension was prepared in 1 ml 0.8 % (w/v) NaCl / 0.05 % (v/v) Tween80 and thereof, 100 µl were incubated on plates containing MEX and 5 mM uridine for around 15 hours at 30 °C. The plates were prepared prior to incubating with cellophane discs, to ease the harvest of the mycelium for the following protoplast generation.

6.4.2. Transformation and selection

150 mg of lysing enzymes from *T. harzianum* (Glucanex® Merck; L1412-5G) and 60 mg β-glucanase were dissolved in 15 ml transformation buffer A (see Table 17) and protoplasts were generated by incubating mycelium in this solution at 30 °C and 60 rounds per minute (rpm) in an rotary shaker for approximately four hours.
Protoplasts were identified with a microscope and approximately 10^7 were further processed per approach. Protoplasts were filtrated and washed with 1.2 M D-sorbitol before suspending them in 1 ml transformation buffer B (see Table 17).
20 µg of pRP4_TX plasmids (see 6.3), were linearized O/N with 30 U of either SpeI or AflII in 10 X CutSmart buffer (NEB®) at 37 °C. Linearized DNA was precipitated in 96 % ethanol and after washing with 70 % ethanol, dissolved in 15 µl double distilled water (ddH_20). This solution was mixed with 100 µl protoplast suspension, 135 µl transformation buffer B and 100 µl 20 % polyethylene-glycol (PEG - see Table 17). This mixture was incubated on ice for 30 minutes and afterwards 750 µl 60 % PEG was added subsequently. After 20 minutes of incubation at room temperature, 4.1 ml of transformation buffer C (see Table 17) were added and different amounts (100 µl, 300 µl, 1 ml and the rest) of this mixture were poured into petri

dishes with 15 ml of 50 °C warm molten Mandels-Andreotti (MA)[144] agar without peptone but 1 M D-sorbitol . The plates were incubated at 30 °C in steady light for several days until colonies were visible.

Colonies were transferred to plates containing MA agar without peptone and after another two to three days transferred to MA plates containing 0,1 % (w/v) IGEPAL® (Merck, 18896-50ML) for homokaryon streaking. After one to two days single colonies were picked and transferred to MA plates without peptone and left to grow at 30 °C until sporulation. All buffers and media that were used in this project are summarized in section 7.2 Table 17.

6.5. Genotyping

The genotyping procedure included polymerase chain reaction (PCR) amplification, gel electrophoresis and visualization on the BioRad® Gel Doc XR with Image Lab Software.

6.5.1. DNA Extraction

Spores from the *T. reesei* transformants, with either synTF or OE-cassettes, were incubated in 1.5 ml MEX medium for two to three days and the mycelium was subsequently processed for DNA extraction with the phenol-chloroform method. In short: approximately 50 mg of mycelium was homogenized in 1 ml CTAB buffer (Table 17) with glass beads of different sizes (0.37 g - 0.1 mm diameter; 0.25 g - 0.25 mm diameter and 1 bead with 5 mm diameter) using the MP Biomedical™ FastPrep -24™ Classic device at level 6 for 30 seconds. The homogenizate was incubated for 20 minutes at 65 °C and 800 µl of supernatant was mixed with 800 µl phenol:chloroform:isoamylalcohol mixture (25:24:1). Aqueous top phase was transferred to a new tube and mixed with chloroform. After 10 minutes incubation at RT samples were centrifuged for 10 min at 4 °C. Aqueous top phase was transferred to a new tube and RNA was degraded with 50 U RNase A at 37 °C for 45 minutes. Genomic DNA was precipitated in 350 µl isopropanol. The resulting pellet was washed with 70 % ethanol and dissolved in 50 µl TE-buffer (pH = 8). Concentration and purity were determined via measurement on the NanoDrop™ ONE device (Thermo Fisher Scientific Inc.).

6.5.2. Primer design and PCR protocol

The primers for genotyping were designed according to the following characteristics:

1. Length between 20 and 22 base pairs
2. Approximately 50 % GC-content
3. No A or T stretch at the 3' ending
4. Few to none secondary structures or primer dimers forming

PCR was carried out with 10 ng/µl genomic DNA in a 25 µl standard reaction with OneTaq® DNA polymerase (NEB, #M0509L) according to the manufacturer protocol and separated as well as visualized as described previously (see 6.4).

6.6. Quantitative PCR (qPCR)

6.6.1. Cultivation and cDNA synthesis

Transformants that were suitable for qPCR analysis (see Figure 9) were incubated in MA medium with 1% D-glucose as sole carbon source without peptone (see Table 17). Instead of citrate the pH was adjusted using 5 M HCl. Spore suspensions were added to the medium to maintain an equal starting optical density (OD) of 0.05 and the incubation trays were incubated at 30 °C. After 72 hours mycelium was harvested and flash frozen in liquid nitrogen. Cultivations were carried out in biological triplicates but only one sample per transformant was measured. The other mycelia were stored in -80 °C freezer.

Following, approximately 100 mg fungal mycelium was homogenized with the Fast Prep device as prescribed previously (6.5.1). Total RNA was isolated using the Direct-zol™ RNA Miniprep Kit (Zymo Research, #R2050) according to the manufacturers protocol. DNA was digested during the extraction process with 360 U DNAse I at RT for 15 minutes. RNA concentration and purity were measured with NanoDrop™ ONE device and the integrity of the extracted RNA was measured once, for evaluation of the extraction process, with the 5200 Fragment Analyzer system (Agilent, Santa Clara, CA, USA). RNA purities in samples was accepted at a A_{260}/A_{280} ratio of at least 2.0 and A_{260}/A_{230} ratio of at least 2.25. RNA integrity had an RQN of 7.8 for one WT extraction and an RQN of 6.8 for one 105805-syn-6 extraction. This was regarded enough for qPCR analysis.

500 ng RNA were subsequently reverse transcribed into complementary DNA (cDNA) with LunaScript™ RT SuperMix (NEB, #E3010) according to the manufacturers protocol. cDNA was stored in -20 °C freezer.

6.6.2. qPCR and transcript level estimation

cDNA was diluted 1:50 and 2 µl of the dilution were used as template in a 20 µl reaction. The pipetting was carried out with the QIAgility robot (QIAGEN, Hilden, Germany, ID:9001903) to minimize errors. qPCR was performed with Luna Universal qPCR Master Mix (NEB, #M3003L) and measurements were carried out as duplicates with no template controls (NTC – water) carried out in triplicates. The measurements were carried out on a Rotor-Gene® Q 5plex platform (Qiagen) and the qPCR protocol was set up as following:

1. Initial Denaturation 95 °C - 60 seconds
2. Denaturation 95 °C - 15 seconds
3. Extension 60 °C - 30 seconds (+ plate read)
4. Melt Curve 60-95 °C – 1 °C/minute

Steps two and three were repeated for 40 cycles. Step one and four were carried out once. Fluorescence was measured at 510 nm (SYBR® Green I). Data was primary analyzed with the inbuilt software of Qiagens Rotor Gene (Rotor-Gene Q Series Software) and relative expression levels were calculated using Pfaffl et. al's method[145]. Selected genes of each cluster were normalized to the wildtype samples (WT – *T. reesei* QM6a *Δtmus53*) and *act1* (encoding actin) and *sar1* (encoding a small GTPase) were chosen as reference genes[146]. Quantitative fluorescence signals in negative controls were either not measured at all (Efficiency = 0) or beyond what was considered impactful (C_t > 35). C_t-values above 30 were considered less impactful, especially when WT and sample strains had similar results in this range. Generally, regulation differences lower than 0.5 $\log_{10}$-fold were not considered impactful as biological and methodical deviations range in this area but were nonetheless mentioned to underline the difference to strongly elevated transcript levels (1.5 - 3.0 $\log_{10}$-fold change). Efficiencies were ranging between 1.75 and 1.95 and were averaged genewise. Raw data of all performed qPCR is provided in an extra file and can be requested at the author.

7. Supplementary Material

7.1. Primer- and plasmid sequences

In this section of the supplementary all used primer sequences and the sequence of the plasmid pCD_TXY will be displayed.

pCD_TXY sequence:

CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTA
ACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGT
GGCCGCTACAGGGCGCTCCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGTTTCGGTGC
GGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAA
CGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAGCGCGACGTAATACGACTCA
CTATAGGGCGAATTGAGTGAAGGCCGTCAAGGCCTAGGCGCGCCAGCGGCCGCCGTTCCGGAACC
CCTAGGCTCGAGATAACGGTGAGACTAGCGGCCGGTCCCTTATCCAGCTGTCCACGTTGGCCTGCC
CTCAGTTAGCGCTCAACTCATGCCCTCACTGGCGAGGCGAGGGCAAGGATGGAGGGGCAGCATCG
CCTGAGTTGGAGCAAAGCGCCGCCATGGGAGCAGCGAACCAACGGAGGGATGCCGTGCTTTGTCG
TGGCTGCTGTGGCCAATCCGGGCCCTTGGTTGGCTCACAGAGCGTTGCTGTGAGACCATGAGCTAT
TATTGCTAGGTACAGTATAGAGAGAGGAGAGAGAGAGAGAGAGTGTGAGAGGGAAAAGGTGAGG
TTGAAGTGAGGTAGTAAGTGAAAAAAAAAAAAAAATCAACACTGACGGCTGCCGTCTGCCACCCT
CCTCCACCCAGCCACCTGCACACTCAGCGCGCAGCATCACCTAATCTTGGCTCGCCTTCCGCAGCTC
AGGTTGTTTTTTTTTTTCTCTCTCCTCGTCGAAGCCGCCCTTGTTCCCTTATTTATTTCCCTCTCCTA
CCTTGTCTGCCTTTGGTCCATCTGCCCCTTTGTCTGCATCTCTTTTGCACGCATCGCCTTATCGTCGT
CTCTTTTTTCACTCACGGGAGCTTGACGAAGACCTGACTCGTGAGCCTCACCTGCTGATTTCTCTCC
CCCCCTCCCGACCGGCTTGACTTTTGTTTCTCCTCCAGTACCTTATCGCGAAGCCGGAAGAACCCTC
TTAACCTCTAGACTTAAGATCCATATGTTGTCCAATCCTCTCCGTCGCTATTCTGCCTACCCCGACA
TCTCCTCGGCGTCATTTGACCCGAACTACCATGGCTCACAGTCGCATCTCCACTCGATCAACGTCAA
CACATTCGGCAACAGCCACCCCTATCCCATGCAGCACCTCGCACAGCATGCGGAGCTTTCGAGTTC
ACGCATGATAAGGGCCAGTCCGGTGCAGCCAAAGCAGCGCCAGGGCTCTCTTATTGCTGCCAGGA
AGAATTCAACGGGTACTGCTGGGCCCATTCGGCGGAGGATCAGTCGCGCTTGTGACCAGTGCAACC
AGCTTCGTACCAAGTGCGATGGCTTACACCCATGTGCCCATTGTATAGAATTCGGCCTTGGATGCG
AATATGTCCGAGAGAGAAAGAAGCGTGGCAAAGCTTCGCGCAAGGATATTGCTGCCCAGCAAGCC
GCGGCGGCTGCAGCACAACACTCCGGCCAGGTCCAGGATGGTCCAGAGGATCAACATCGCAAACT
CTCACGCCAGCAAAGCGAATCTTCGCGTGGCAGCGCTGAGCTTGCCCAGCCTGCCCACGACCCGCC
TCATGGCCACATTGAGGGCTCTGTCAGCTCCTTCAGCGACAATGGCCTTTCCCAGCATGCTGCCATG
GGCGGCATGGATGGCCTGGAAGATCACCATGGCCACGTCGGAGTTGATCCTGCCCTGGGCCGAACT
CAGCTGGAAGCGTCATCAGCAATGGGCCTGGGCGCATACGGTGAAGTCCACCCCGGCTATGAGAG
CCCCGGCATGAATGGCCATGTGATGGTGCCCCCGTCGTATGGCGCGCAGACCACCATGGCCGGGTA
TTCCGGTATCTCGTATGCTGCGCAAGCCCCGAGTCCGGCTACGTATAGCAGCGACGGTAACTTTCG
ACTCACCGGTCACATCCATGATTACCCGCTGGCAAATGGGAGCTCGCCCTCATGGGGAGTCTCGCT
GGCCTCGCCTTCGAACCAGTTCCAGCTTCAGCTCTCGCAGCCCATTAATGGCGACGTCGGAGCAAT
TGTTCTTACTCCACAGTCGACAACGGCCCCGAGCCGCATGCCGTCCATATCACTTGCTGCTGCGGAC
CCAATCATGCCCGACGTGGTATTGATGGAACTCGATCTACTATACTTTGAGAGAGTCCATCCCTTTG
CGCCCATGGTCCACAAGAGGCGGTACTATGCCTGGGCCAGCGACCCAAACGCCTCGCCGGCCCGTA
CCGCTCTGCGCTCCGCCATGAGGACCATTGCGTCGGCCATGTCGCCTCAATTCTGCGATATAGGAC
ACGTCATGTATGCTTCCACGCGACGCATGCTCGAGACACAGGACGCATGCCCGGAAACCGGCCTCC
CCTGGATGACGAGGCTCAAGCCACCTCATGAGCAGCGGAAGATGCACCACGAGCGCATACAGGCT
TGGCTCTTGCTGGCGTACTATGATGTCTTGCGTAAATCAGAGCATCAGGCCTTTATCACCGCCCGCC
GAGCATTTCGCCTTCTCCGCCTCTCTGGCCTGTGCGAGATGGACATAGACGCCTACAGCGGCAGAC
AGGGCGATGTCGCGACATGCACTCCGCCTGCTGAGACAACGTGGAATATGCAGATGTGTTCGTCAG
AGAATGGTGCAGATGAGGCTGTTTTGCAGCAGGATTGGATCTCCGTGGAGGAGAGGAGGCGGACT
GTCTGGAGCGCCTTTCTCCTCGACCGCCTCTCGACAATGGTCAACGACCAGCCCACGATGCTGATG
GAAGAAGCGTTCTACACACGACTTCCCATGTCAGAGGCCGAGTTCCAGAGCGGTACCCAAGAGCC
CGTCTCGCAGATGGGCTTCCTGATCGAGGCAACAGACGGGATCAAAGCCTCCAACAGCATCCAACC

ATTGCCCCCATTCGCCCATTGCGTCGTCGTGGCCAACCTCTTTGCACGCTGCATGACACACTGCAAG
ATGGCCATGCAATCACCGCCCATGTCTGCACCTGAGGCCCACGACTTTTGGATACGACATCAGTGG
CTGGCATCCGCGGCTGCAAACGCCTGCGAGTCTACGGAGACGCGATGTGACCCCATGCTGGTCTTC
ACGCGCATATTGGCTTACAGCGCGAGCTTGTCACTGTGCAGTACGGCAAATGCTACGTCGTGGCAG
ACGCTGGACCACCACTTGATGGCGATGGCTTGTAAGCCGGCTGCGCATCAGGCCGCTTCCGAGGTG
GTGCGCATCATCAAGACGGCCCCGCGAATTGCCTTCTTCAAGATGCACCCCTTCTTCCCTAATGCCA
TCGCACTCGTTACCAGTTTCCTCAACGCCGACGTGCCTTATCTGCCTTCGACACGAGGTGGCAATGC
CATGGACGCCATTCAGGAGAGACAGGATGCTGTGAATGAATTGCTTGCGGCGCTGCGAAGATCAA
GTCAGGTGAACAACTTGGCTGCAGAGCTATTGTGCAAGCTGGAACTGGACATTGGACAGGCGGCG
AGCGATGGGAGCATTTACGGCTAGACTAGTTGATAAGCTATTAACTGGCCTCATGGGCCTTCACTT
CACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAACATGGTCATAGCTGTTTCCT
TGCGTATTGGGCGCTCTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGGTAAAGCCT
GGGGTGCCTAATGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCG
TTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA
AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTT
CCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAT
AGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAA
CCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGAC
ACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGT
GCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGC
GCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC
GCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAA
GATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGG
TCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT
CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTATTATTTACCCACCACTTTGGTAATTTCGCCT
TTCACGTAATGCACAAATTCTTCCAGCTGATCCGCACGGCTGGCCAGACGATCTTCTTCCTGGCCCA
GATACGCCTGACGCGCTTCCAGAATCACCGGCTGATACTGCGCCGGCAGACGTTCCATGGCCCAAT
CCGCCGCCACATCTTTCGGCGCAATTTTGCCGGTCACCGCGCTATACCAAATACGGCTCAGGGTCA
GCACCACGTTGCGTTCATCACCCGCCCAATCCGGCGGGCTGTTCCACAGGGTCAGGGTTTCGTTCA
GCGCTTCAAACAGGTCCTGTTCCGGCACCGGATCAAACAGTTCTTCCGCTGCCGGACCAACCAGCG
CCACGCTATGTTCACGCGCTTTGGTCAGCAGAATGGCCAGATCGATATCAATGGTCGCCGGTTCAA
AAATGCCGGCCAGAATATCGTTACGCTGCCATTCGCCAAACTGCAGCTCACGTTTCGCCGGATAAC
GCCACGGAATAATGTCATCATGCACCACGATGGTCACTTCCACCGCACGCAGAATTTCGCTTTCAC
CCGGGCTCGCGCTGGTTTCCAGCAGATCGTTAATCAGCGCACGACGGGTGGTTTCATCCAGACGCA
CGGTCACGGTCACCAGCAGATCAATATCGCTATGCGGTTTCAGGCCGCCATCCACCGCGCTGCCAT
ACAGATGCACGGCCAGCAGGGTCGGTTCCAGATGACGTTCAATCACGCCCACCACTTCAGACAGCT
GGGTGCTCACTTCCGCAATCACCGCTTCACGCATAATGTTCAGTTTGGTTTTCGGACGGCTGCCCTG
CTGGGTCACATCGTTGCTGCTCCACAGCATCAGGCAACGGCCCACCGCATAACGCGCCTGCTGTTT
GCTCGCACGCGGCATGCTCTGCACGCCAAAAAAGCAATCATAGCACGCCATAAACACCGCCACCG
CACCATTACCGCCGCTACGTTCGGTCAGGGTACGGCTCCAATTACGGCTACGCATACTCTTCCTTTT
TCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAG
AAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC

Table 13. Primers used for the generation of the truncated DBD.

Primersequences for truncated DBD				
TF		Sequence 5' → 3'	Name	Number
72993	forward	CTTAAGCGCAAGAGCATCCACAAATGG	72993_fwd-AflII	962
	reverse	TCGACTGTGGAGTAAGAACAATTGCGTTGACAGATGGATCGAGG	72993_N242r-ypr1	B181
105805	forward	CTTAAGATGCTGCTGTTCGGTGGC	105805_fwd_AflII	956
	reverse	TCGACTGTGGAGTAAGAACAATTGCCTCGTCATGATAGTTGTTGC	105805_E181r-ypr1	B173
122783	forward	CTTAAGCAGACTTTGGCCTAGAAGATATG	122783_fw_AlfII	989
	reverse	TCGACTGTGGAGTAAGAACAATTGGAGCATCTTGGCCGATCATTG	44995_A119r-ypr1	B177
79725	forward	CTTAAGAATGGAAACCAAGGCACCC	79725_fwd_AflII	958
	reverse	TCGACTGTGGAGTAAGAACAATTGGTCTGCTTGTCAGTAGGTGGC	79725_R125r-ypr1	B175
111742	forward	CTTAAGCATGGCGTCCTCTTACGGCAC	111742_fwd_AflII	960
	reverse	TCGACTGTGGAGTAAGAACAATTGCTTGGCAAGATGGTGGTACAG	111742_Q133r-ypr1	B179

Table 14. Primers used for colony-PCR

Primersequences for colony-PCR				
TF		Sequences 5' → 3'	Name	Number
Ptef	forward	TCCGGAGAGTTGGGCAAAATCAGGC	Ptef_fwd_BspEI	135
72993	reverse	TCGACTGTGGAGTAAGAACAATTGCGTTGACAGATGGATCGAGG	72993_N242r-ypr1	B181
105805	reverse	TCGACTGTGGAGTAAGAACAATTGCCTCGTCATGATAGTTGTTGC	105805_E181r-ypr1	B173
122783	reverse	TCGACTGTGGAGTAAGAACAATTGGAGCATCTTGGCCGATCATTG	44995_A119r-ypr1	B177
79725	reverse	TCGACTGTGGAGTAAGAACAATTGGTCTGCTTGTCAGTAGGTGGC	79725_R125r-ypr1	B175
111742	reverse	TCGACTGTGGAGTAAGAACAATTGCTTGGCAAGATGGTGGTACAG	111742_Q133r-ypr1	B179

Table 15. Primers used for genotyping.

Primersequences for genotyping				
		Sequences 5' → 3'	**Name**	**Number**
5' flank	forward	CCAGACGGTGATTCACATATACG	5_pyr4_fwd3	338
	reverse	CTTAAGTGTGATGTAGCGTGAGAGCTG	Ptef_rev_BspTI	136
3' flank	forward	TGCCTTTATCCACATGACGC	pyr4_3fwd3	B911
	reverse	CAGGAAGCTCAGCGTCGAG	Tpyr4_rev2	379

Table 16. Primers used for qPCR analysis.

Primersequences for qPCR				
		Sequences 5' → 3'	**Name**	**Number**
Reference genes				
act1	forward	TGAGAGCGGTGGTATCCACG	act1f	A774
	reverse	GGTACCACCAGACATGACAATGTTG	act1r	A775
sar1	forward	TGGATCGTCAACTGGTTCTACGA	sar1fw	A772
	reverse	GCATGTGTAGCAACGTGGTCTTT	sar1rev	A773
Cluster 21				
Genes			**Name**	**Number**
58285	forward	GCTGAACATGATGACCAAAGAG	58285_q1f	A318
	reverse	TTCGTTTTCAGGGAGTAGATGC	58285_q1r	A319
58289	forward	AGCAGAAGCCTGGTTTGACTAC	58289_q1f	B956
	reverse	GCTTTGCCGAACTGTGTAGG	58289_q1r	B957
76204	forward	CCGGAAATTCTCGTTGATAGG	76204_q1f	B958
	reverse	ATTGAGCAGGCTCCATTTGTC	76204_q1r	B959
58953	forward	CGTCAGAGATGAGGAGGAACTG	58953_q1f	B960
	reverse	GGAATCGTTGAGGCAAGCAC	58953_q1r	B961
105221	forward	TCTTTCATCAGCTTCACTGCC	105221_q1f	B962
	reverse	GAGCCCACTCGACAATCTTG	105221_q1r	B963
105222	forward	GACTCTTCCGAGGTTGCAGC	105222_q1f	B964
	reverse	CGGAGGCGTCTGTTTGATAG	105222_q1r	B965
105223	forward	TTTCCACACTGTCGTCACGTC	105223_q1f	B966
	reverse	GTTGGCAAGAAATACACAGCTAC	105223_q1r	B967
120872	forward	TTCACGTATCCAAACCAGACTG	120872_q1f	B968
	reverse	CACTCTATGACCCTGCTTTGG	120872_q1r	B969

Table 16. Primers used for qPCR analysis (continued)

Cluster 22				
Genes			**Name**	**Number**
105804	forward	TTGCCCATTAGCCTAGACTATCG	105804_q1f	314
	reverse	CCAAAGGAGGAGGCAAACC	105804_q1r	315
105798	forward	GGACGTTTTTGCGCAATCTG	105798_q1f	B970
	reverse	AGCACGACACCAAAGGAGGC	105798_q1r	B971
59655	forward	TATGCCGGCCTTGTGGATGC	59655_q1f	B972
	reverse	ATCGCGAATGCTGTCCGTCG	59655_q1r	B973
3262	forward	GGGATACGCGATGGGGAG	3262_q1f	B974
	reverse	TCCTGGCAATGTCTTCGGC	3262_q1r	B975
59377	forward	GCCTCACGCAACCGAATTC	59377_q1f	B976
	reverse	GCGCAACCCGAAACACTG	59377_q1r	B977
59775	forward	AACATCGTCGGGCAATCATC	59775_q1f	B978
	reverse	GAGAGTGTCGTCGGCAAAGAT	59775_q1r	B979
59771	forward	TTTCTTGCCAACGACCTGC	59771_q1f	B980
	reverse	TGCAACGGACAAGCCAAAG	59771_q1r	B981
36822	forward	TAAGACACCCGAGTTCCTGG	36822_q1f	B982
	reverse	AGCCATTTCTCGACAACTCC	36822_q1r	B983
59690	forward	ACTGTTTCCTGGATCCCGG	59690_q1f	B984
	reverse	TTTCGGCCGTTGATGTGAC	59690_q1r	B985
Cluster 52				
Gene			**Name**	**Number**
65172	forward	TAGCCAGGCGAATTATGCG	65172_q1f	305
	reverse	GGTCCTGGGTATCGCCTATC	65172_q1r	306
Cluster 78				
Genes			**Name**	**Number**
81964	forward	TACCAACTCTTCGCAAACGTG	81964_q1f	324
	reverse	CCAGGAGTGATCCAGAAGAAGTC	81964_q1r	325
69652	forward	CAGTATTCGTGGACATTGGC	69652_q1f	B986
	reverse	CCCAATTGTGAAGAATAGCG	69652_q1r	B987
52222	forward	CAAATGGGCTCGACCGTC	52222_q1f	B988
	reverse	GCCAAGCAGCTCGCAATAG	52222_q1r	B989
69625	forward	GCTTGGGATGTGCTATTGG	69625_q1f	B990
	reverse	TGATGTCCGTCGTCGAGAG	69625_q1r	B991
69650	forward	TCTCAAAGGTCGAAGGGCTC	69650_q1f	B992
	reverse	TGAGGTAGATGAAGTCGACAAGG	69650_q1r	B993

7.2. Media and buffers

Table 17. Media & buffers

Medium/buffer	Composition	Solved in
Malt extract (MEX)	3 % (w/v) malt extract (Merck, CAS 8002-48-0)	Tap water
	0.1 % (w/v) peptone (Merck, CAS 91079-40-2)	
	1.5 % (w/v) agar (Merck, CAS 9002-18-0)	
Mandels Andreotti (MA) modified	2 % (v/v) trace elements solution	ddH_20
	25 % (v/v) mineral salt solution	
	50 % (v/v) 0.1 M phosphate citrate buffer	
	0.1 % 5 M urea	
	1 % D-glucose	
	0.1 % peptone (if applicable)	
trace element solution	0.025 % (w/v) $FeSO_4*7\ H_20$	ddH_20 adjust pH to 2.0 with 1M HCl
	0.0085 % (w/v) $MnSO_4*H_20$	
	0.007 % (w/v) $ZnSO_4*7\ H_20$	
	0.01 % (w/v) $CoCl_2*6\ H_20$	
	0.01 % (w/v) $CuSO_4*5\ H_20$	
	0.005 % (w/v) $Na_2SO_4*2\ H_20$	
mineral salt solution	0.56 % (w/v) $(NH_4)_2SO_4$	ddH_20
	0.80 % (w/v) KH_2PO_4	
	0.12 % (w/v) $MgSO_4*7\ H_20$	
	0.16 % (w/v) $CaCl_2*2\ H_20$	
0.1M phosphate citrate buffer	1.78 % (w/v) $Na_2HPO_4*2\ H_20$	ddH20 adjust pH to 5.0 (5.7 for protoplasts) with 1 M citrate
Transformation buffer A	1.36 % (w/v) KH_2P0_4	ddH20 adjust pH to 5.6 with 1 M KOH
	21.86 % (w/v) D-sorbitol	
Transformation buffer B	18.22 % (w/v) D-sorbitol	ddH_20
	0.156 % (w/v) Tris.Cl (pH = 7.5)	
	0.278 % (w/v) $CaCl_2$	
Transformation buffer C	18.22 % (w/v) D-sorbitol	ddH_20
	0.156 % (w/v) Tris.Cl (pH = 7.5)	
60 % PEG solution	60 % (w/v) PEG 4000 (Merck, CAS 25322-68-3|817006)	ddH_20
	1 % (v/v) 1 M Tris.Cl (pH = 7.5)	
	1 % (v/v) 1 M $CaCl_2$	
20 % PEG solution	33.66 % (v/v) 60 % PEG	-
	66.66 % (v/v) transformation buffer B	
Lysogeny Broth (LB)	0.5 % (w/v) yeast extract (Merck, CAS 8013-01-2)	dH_20
	1 % (w/v) NaCl	
	1 % (w/v) peptone (Merck, CAS 91079-40-2)	
CTAB (cetyltrimethylammonium bromide) buffer	8.12 % (w/v) NaCl	ddH_20
	0.156 % (w/v) Tris.Cl (pH = 8.0)	
	2.0 % (w/v) CTAB (Merck, CAS-57-09-0)	
	1% (w/v) poly vinyl pyrrolidone	

7.3. FunORDER results

The results of the FunORDER analyses are gathered in this section.

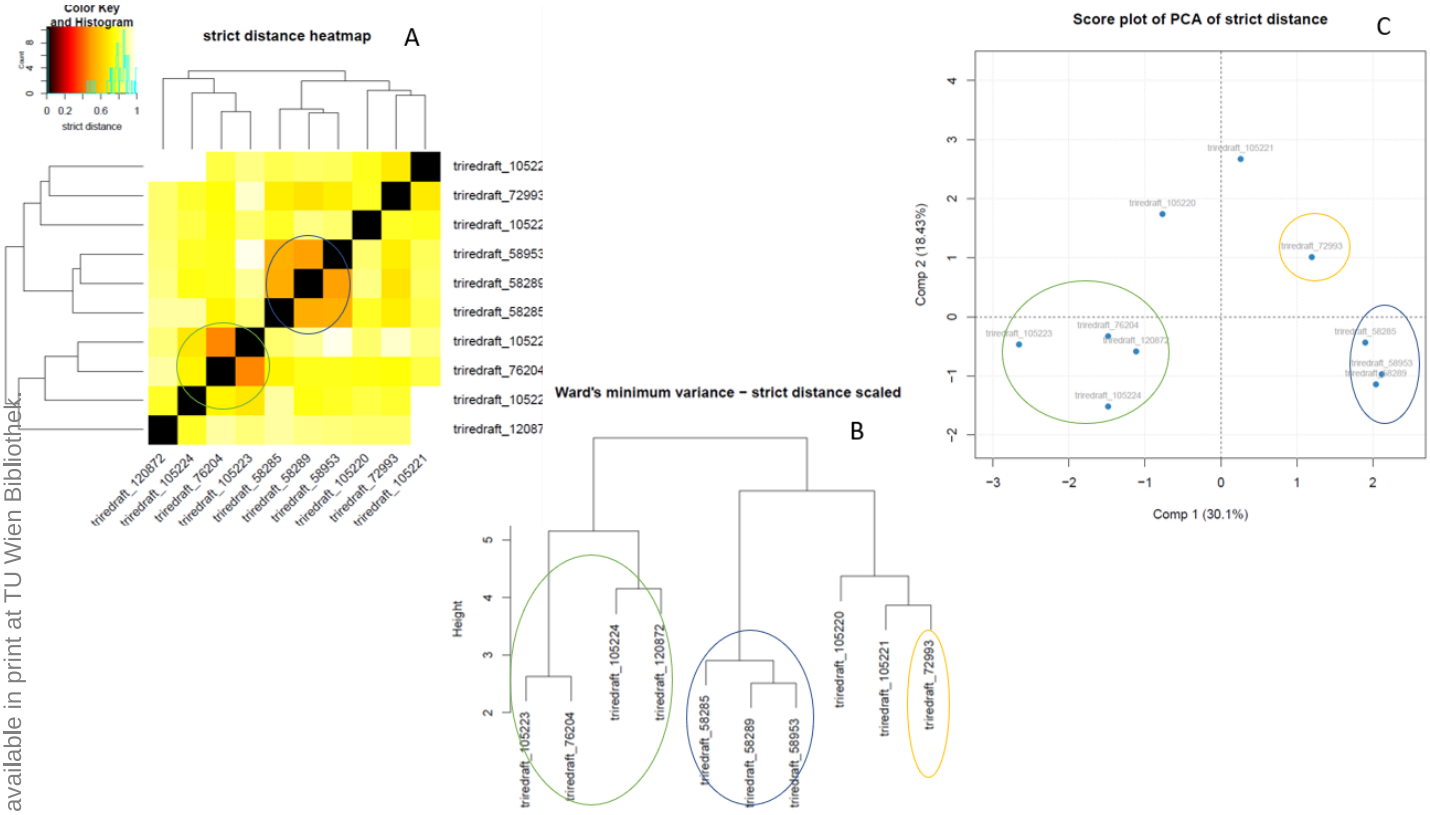

Figure 23. FunORDER analysis of cluster 21.
A … heatmap with strict distance as parameter; B … Ward's plot clustering calculated with strict distance; C … Score plot of Principal component analysis (PCA) of strict distances;
Green circle includes genes: 105223, 105224, 120872 and 76204; Blue circle includes genes: 58285, 58289 and 58953; Yellow circle: transcription factor 72993; Green and blue circles indicate likely co-evolution of included genes.

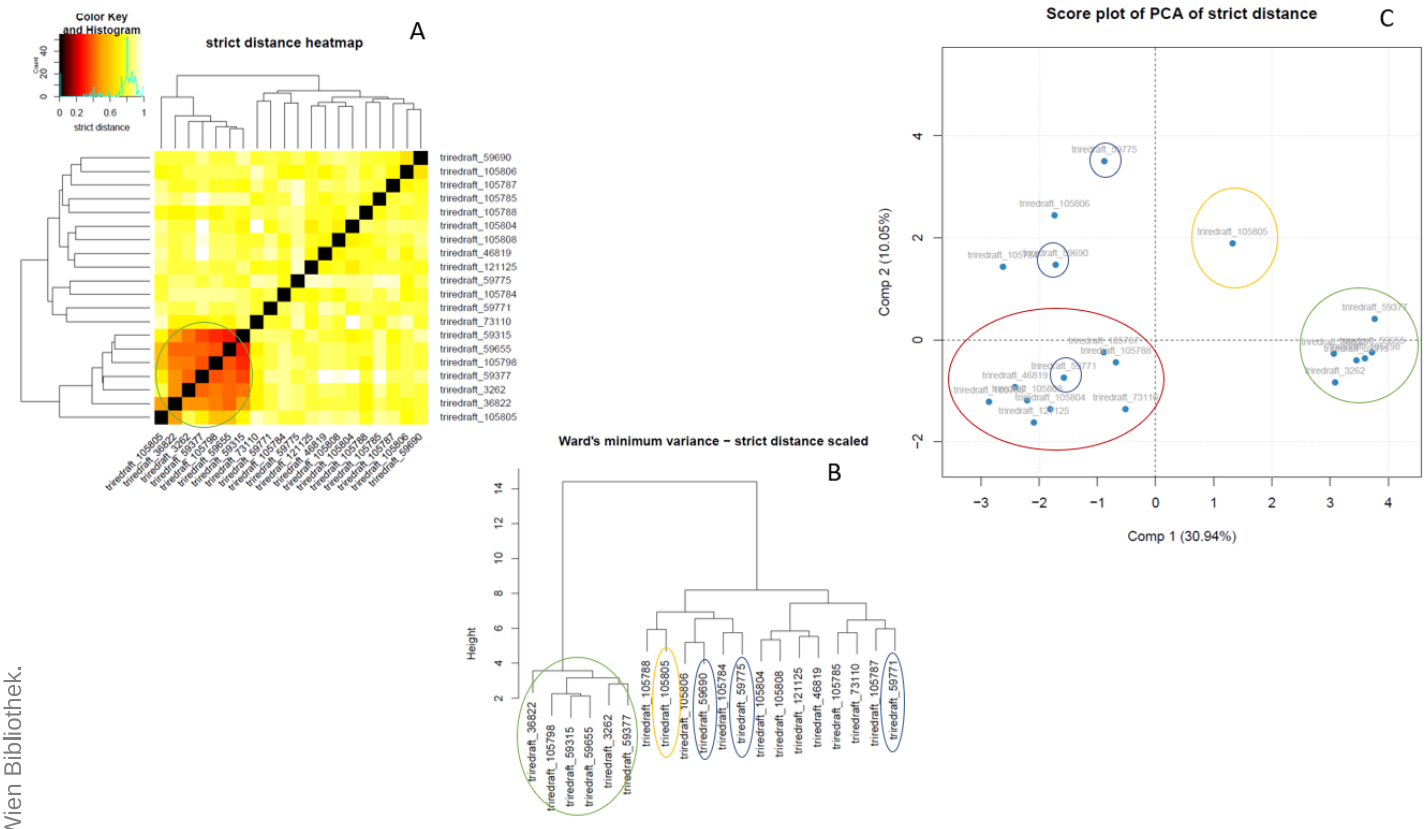

Figure 24. FunORDER analysis of cluster 22.
A … heatmap with strict distance as parameter; B … Ward's plot clustering calculated with strict distance; C … Score plot of Principal component analysis (PCA) of strict distances;
Green circle includes genes: 36822, 105798, 59315 (PKS-NRPS), 59655, 3262 and 59377; Blue circles include genes: 59690, 59775 and 59771; Yellow circle includes TF 105805; Red circle includes genes 105787, 105788, 73110, 59771, 46819, 105808, 105785, 105804 (PKS) and 121125; Green circle indicated likely co-evolution of included genes.

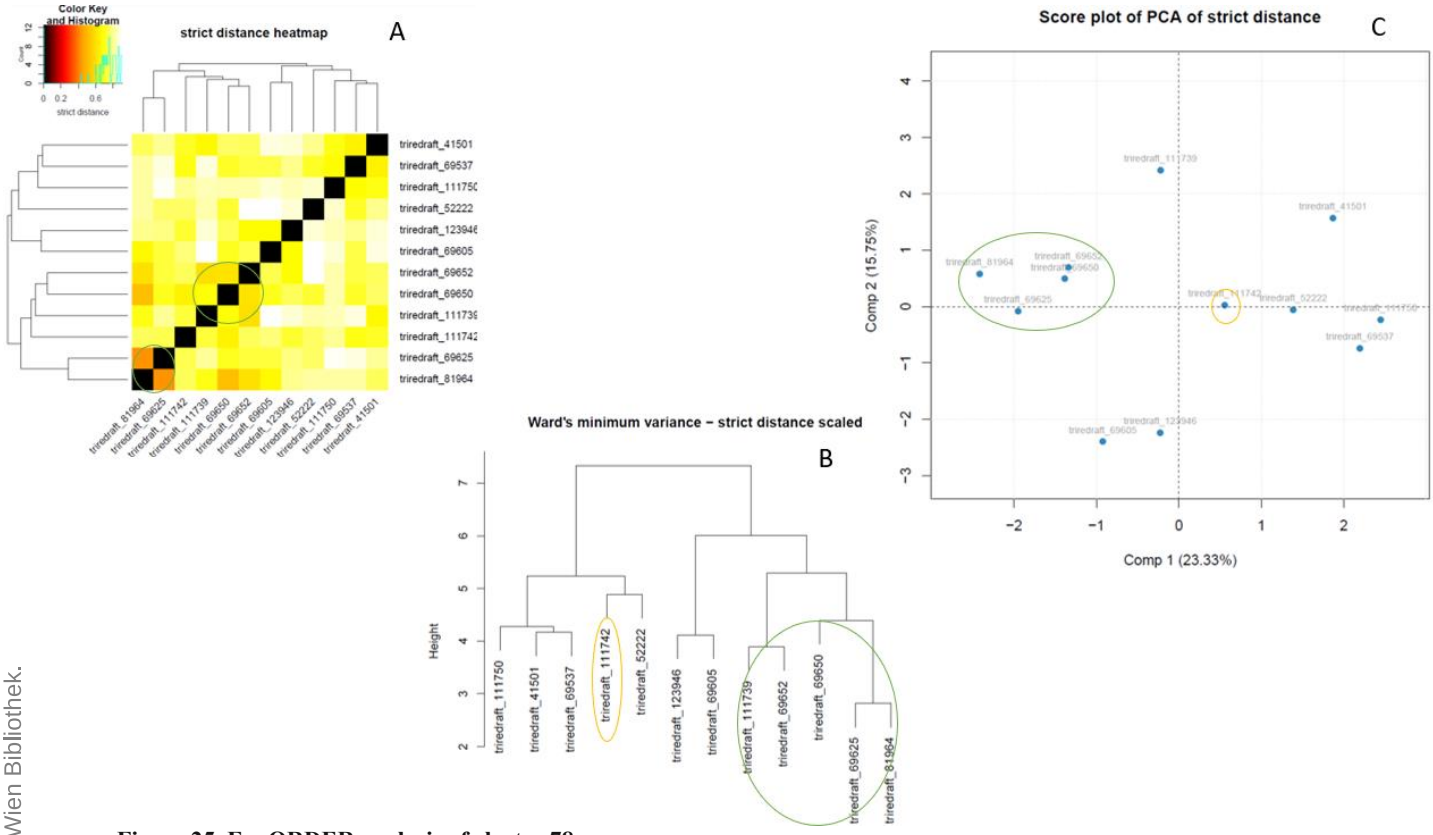

Figure 25. FunORDER analysis of cluster 78.
A … heatmap with strict distance as parameter; B … Ward's plot clustering calculated with strict distance; C … Score plot of Principal component analysis (PCA) of strict distances;
Green circle includes genes: 111739, 69652, 69650, 69625, 81964-PKS; Yellow circle includes TF 111742; Green circle indicated likely co-evolution of included genes.

7.4. Exemplary flawed raw data - qPCR

The following Table shows exemplary what was described as "flawed data" in 3.3.1.

Table 18. Example of flawed qPCR raw data.
Av … Average; Rep … Replication; WT … Wildtype – *T. reesei* QM6a *Δtmus53;* NTC … No Template Control; C_t … Cycle threshold

Sample	Gene	C_t	Av. C_t	Efficiency (EFF)	Rep. EFF	Average EFF
72993-OE-1	58289	21.5	21.5	1.71	1.71	1.73
72993-OE-1	58289	21.4		1.72		
72993-OE-2	58289	19.9	20	1.74	1.73	
72993-OE-2	58289	20.1		1.71		
72993-syn-2	58289	22.6	22.6	1.72	1.73	
72993-syn-2	58289	22.6		1.74		
72993-syn-6	58289	22.2	22.2	1.74	1.74	
72993-syn-6	58289	22.2		1.73		
WT	58289	11.1	21.3	0	0.89	
WT	58289	**31.5**		1.78		
NTC	58289	18.5	-0.16	0.28	0.06	
NTC	58289	11		-0.16		

7.5. Outlook on follow-up experiments

In this section of the supplementary material a follow-up experiments based on the knowledge gained in this project will be presented shortly:

Two of the most promising transformants (72293-OE-1 and 105805-syn-2) were analyzed via untargeted tracer-derived metabolome approach[147].

The method comprises the transformants to be cultivated with ^{13}C-labbeled glucose (tracer) as sole carbon source on minimal medium and subsequent analysis with LC-ESI-HRMS. The principle behind this method is coarsely speaking that the ^{13}C carbons get metabolized by the organism to an equal extent as native ^{12}C carbons[148] and sophisticated algorithms compare tracer-derived chromatograms and spectra with those of the native ones and compare both between the transformants and WT to get a decent overview of the metabolome. Metabolites that are expressed only by the transformants and WT can be found this way. A small part of the results is summarized and displayed in Figure 26.

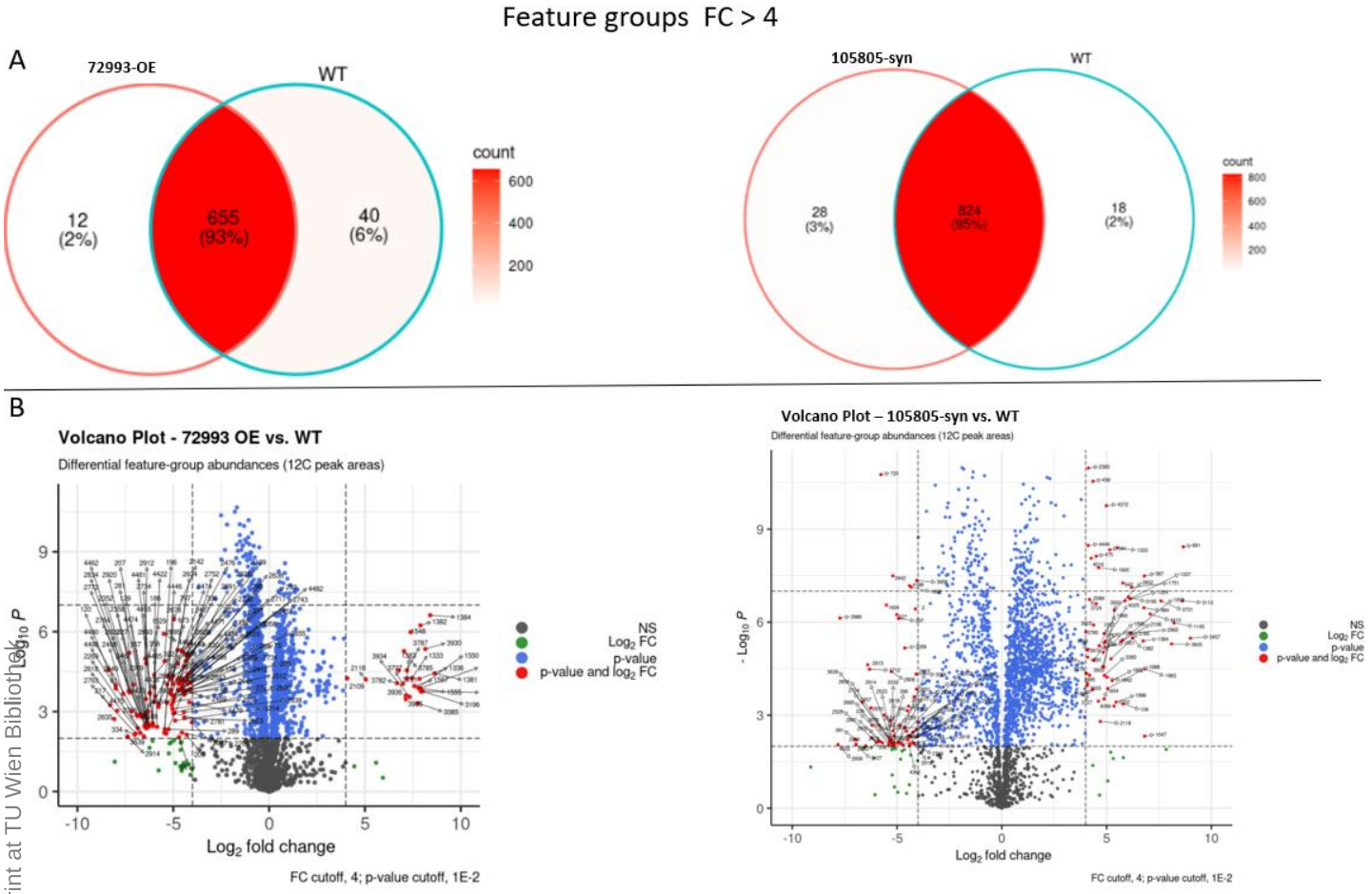

Figure 26. Venn diagram and volcano plots of 72993-OE and 105805-syn transformants compared to the WT.
A … Feature groups correspond to a combination of features found in either positive of negative ionization mode in the mass spectrometry and the corresponding peak in the liquid chromatography and stand for a distinct metabolite[147]. FC stands for fold change and the cutoff for the Venn diagrams and the volcano plots was chosen > ± 4; B … Volcano plot of all measured feature pairs detected in 72993-OE and 105805-syn compared to the WT. p-value cutoff was chosen < 0.01; Grey dots were insignificantly different feature pairs that were also insufficiently different regarding their fold-change compared to the WT. Blue dots indicated significantly different feature pairs but insufficiently difference in fold-change. Green dots indicated sufficiently different fold change of the feature pairs but insignificant statistical difference. Red dots indicate strong statistically significance and strong difference in fold change.

Figure 26/A shows that 12 distinct metabolites were only formed in 72993-OE while 655 were formed by both, 72993-OE and the WT. 28 distinct metabolites were only formed in 105805-syn, and 824 by both 105805-syn and the WT. The fold change was set to > ± 4 to increase the clarity in the distinction and elaborate actual different metabolites formed only by the transformants. Figure 26/B shows the volcano plots of both transformants compared to the WT. Every dot stands for a feature pair detected in the LC-ESI-HRMS and was sorted by the used algorithms (A thorough description of how these algorithms work can be reviewed in the original publication[147]). Grey dots were insignificant and insufficiently different ($-\log_{10}$-p < 2 and $\log_2$-fold change < ±4) feature pairs between WT and transformants. Blue dots were statistically significant, but the log fold change was insufficiently different ($-\log_{10}$-p > 2 and $\log_2$-fold change < ±4) feature pairs between WT and transformants. Green dots were

insignificant, but the log fold change was sufficiently different ($-\log_{10}$-p < 2 and $\log_2$-fold change > ±4) feature pairs between WT and transformants. And finally, red dots were significant and sufficiently different feature pairs between WT and transformants ($-\log_{10}$-p > 0.01 and $\log_2$-fold change > ±4). The volcano plot of 72993-OE vs WT has several feature pairs (red dots) which pointed towards the significantly different expression of metabolites. The same can be stated for 105805-syn vs WT. Furthermore, the latter mentioned had over 20 feature pairs which were highly significant and different ($-\log_{10}$-p > 7 and $\log_2$-fold change > ±4) between 105805-syn and WT. This points towards many metabolites that 105805-syn builds exclusively and to a massively different amount.

Please note that the results which were presented here lastly (Figure 26), were obtained by collaborators of the research group after the main project and the author did not have any part in performing the experiment nor data analysis but only in interpreting of the diagrams. The data was kindly left to the author by the research group afterwards and are only shown for readers to get a better grasp of the work which awaits. The cultivation was performed by the author. Sampling was performed by Univ. Ass. Dr.rer.nat. Mag.rer.nat. Christian Derntl and the analytical measurements and data analyses were performed by our collaborators from the plant-microbe metabolomics research group of the IFA (Tulln, Austria). These results represented only a small fraction of the metabolome analysis. The full outcome, like the search for specific metabolites, carbon enrichment, glucose uptake and more will be published elsewhere.

Acknowledgements

I am most grateful for the opportunity to accomplish my master thesis in the research group of Assistant Prof. Dr. techn. Mach-Aigner in the exciting field of synthetic biology. This young research field has incredible potential and advances will have a huge impact in our modern day lives in the future. I couldn't have done this project without the endless and immediate support from Univ.Ass. Dr.rer.nat Derntl throughout the whole period working there. His support ranged from drafting the concept of this project over sharing his methodical tool box up to answering all the (for sure sometimes tiring) questions I had. I am glad and grateful that I got him as direct supervisor as I couldn't think of someone else more suitable. I'd like to especially thank Univ.Ass. Nadine Hochenegger MSc, who prepared me extraordinary well for this thesis in a former internship and always lent me a hand when needed during this thesis and Univ.Ass. Irene Tomico Cuenca MSc, whose discussions about biology helped me more than once and Univ.Ass. Mag. Pharm. Alexander Vignolle who conducted the bioinformatic analyses, programmed the FunORDER tool and helped immensely in interpreting these results. Furthermore, I'd like to thank everyone who is part of the synthetic and molecular biology research group at the Technical University of Vienna. The working atmosphere in this research group was outstanding and unbelievably helpful. Everyone was pleased to share knowledge disregarding personal advantages which is often a unique feature in research.

Finally, I want to express my deepest gratitude to my family who supported me for longer than they would have to without hesitation. My father Nikos, who never questioned my choice, my mother Yvette, who always knew the right words when times were tough, my brother Alexander, who was always proud of me even when no one should have been and girlfriend Katharina, who stayed beside me from the start to the end and without who I would have quitted countless times. This work belongs to all of us.

Thank you,

Lukas

8. References

(1) Verpoorte, R. Secondary Metabolism. In *Metabolic Engineering of Plant Secondary Metabolism*; Verpoorte, R., Alfermann, A. W., Eds.; Springer Netherlands: Dordrecht, 2000; pp 1–29. https://doi.org/10.1007/978-94-015-9423-3_1.

(2) Muñoz-Elías, E. J.; McKinney, J. D. Carbon Metabolism of Intracellular Bacteria. *Cellular Microbiology* **2006**, *8* (1), 10–22. https://doi.org/10.1111/j.1462-5822.2005.00648.x.

(3) Fernie, A. R.; Carrari, F.; Sweetlove, L. J. Respiratory Metabolism: Glycolysis, the TCA Cycle and Mitochondrial Electron Transport. *Current Opinion in Plant Biology* **2004**, *7* (3), 254–261. https://doi.org/10.1016/j.pbi.2004.03.007.

(4) Ronne, H. Glucose Repression in Fungi. *Trends in Genetics* **1995**, *11* (1), 12–17. https://doi.org/10.1016/S0168-9525(00)88980-5.

(5) Bömke, C.; Tudzynski, B. Diversity, Regulation, and Evolution of the Gibberellin Biosynthetic Pathway in Fungi Compared to Plants and Bacteria. *Phytochemistry* **2009**, *70* (15), 1876–1893. https://doi.org/10.1016/j.phytochem.2009.05.020.

(6) Isah, T. Stress and Defense Responses in Plant Secondary Metabolites Production. *Biological Research* **2019**, *52*.

(7) Goyal, S.; Lambert, C.; Cluzet, S.; Mérillon, J. M.; Ramawat, K. G. Secondary Metabolites and Plant Defence. In *Plant Defence: Biological Control*; Mérillon, J. M., Ramawat, K. G., Eds.; Springer Netherlands: Dordrecht, 2012; pp 109–138. https://doi.org/10.1007/978-94-007-1933-0_5.

(8) Zimand, G.; Elad, Y.; Chet, I. Effect of Trichoderma Harzianum on Botrytis Cinerea Pathogenicity. *Phytopathology* **1996**, *86* (11), 1255–1260.

(9) Elad, Y. Mechanisms Involved in the Biological Control OfBotrytis Cinerea Incited Diseases. *European Journal of Plant Pathology* **1996**, *102* (8), 719–732. https://doi.org/10.1007/BF01877146.

(10) Katz, L.; Baltz, R. H. Natural Product Discovery: Past, Present, and Future. *Journal of Industrial Microbiology and Biotechnology* **2016**, *43* (2–3), 155–176. https://doi.org/10.1007/s10295-015-1723-5.

(11) Gaynes, R. The Discovery of Penicillin—New Insights After More Than 75 Years of Clinical Use. *Emerg Infect Dis* **2017**, *23* (5), 849–853. https://doi.org/10.3201/eid2305.161556.

(12) DePestel, D. D.; Benninger, M. S.; Danziger, L.; LaPlante, K. L.; May, C.; Luskin, A.; Michael, P.; Hadley, J. A. Cephalosporin Use in Treatment of Patients with Penicillin Allergies. *Journal of the American Pharmacists Association* **2008**, *48* (4), 530–540. https://doi.org/10.1331/JAPhA.2008.07006.

(13) Griffith, R. S.; Black, H. R. Erythromycin. *Medical Clinics of North America* **1970**, *54* (5), 1199–1215. https://doi.org/10.1016/S0025-7125(16)32587-1.

(14) Yilmaz, N.; Agus, N.; Bayram, A.; Samlioglu, P.; Sirin, M. C.; Karaca Derici, Y.; Yilmaz Hanci, S. Antimicrobial susceptibilities of Escherichia coli isolates as agents of community-acquired urinary tract infection (2008-2014). *Turkish Journal of Urology* **2016**, *42* (1), 32–36. https://doi.org/10.5152/tud.2016.90836.

(15) Moore, C. E.; Sona, S.; Poda, S.; Putchhat, H.; Kumar, V.; Sopheary, S.; Stoesser, N.; Bousfield, R.; Day, N.; Parry, C. M. Antimicrobial Susceptibility of Uropathogens Isolated from Cambodian Children. *Paediatrics and International Child Health* **2016**, *36* (2), 113–117. https://doi.org/10.1179/2046905515Y.0000000008.

(16) Sibero, M. T.; Sabdaningsih, A.; Cristianawati, O.; Nuryadi, H.; Radjasa, O. K.; Sabdono, A.; Trianto, A. Isolation, Identification And Screening Antibacterial Activity from Marine Sponge-Associated Fungi Against Multidrug-Resistant (MDR) *Escherichia Coli*. *IOP Conf. Ser.: Earth Environ. Sci.* **2017**, *55*, 012028. https://doi.org/10.1088/1755-1315/55/1/012028.

(17) Belknap, K. C.; Park, C. J.; Barth, B. M.; Andam, C. P. Genome Mining of Biosynthetic and Chemotherapeutic Gene Clusters in Streptomyces Bacteria. *Scientific Reports* **2020**, *10* (1), 2003. https://doi.org/10.1038/s41598-020-58904-9.

(18) Devi, R.; Kaur, T.; Guleria, G.; Rana, K. L.; Kour, D.; Yadav, N.; Yadav, A. N.; Saxena, A. K. Chapter 9 - Fungal Secondary Metabolites and Their Biotechnological Applications for Human Health. In *New and Future Developments in Microbial Biotechnology and Bioengineering*; Rastegari, A. A., Yadav, A. N., Yadav, N., Eds.; Elsevier, 2020; pp 147–161. https://doi.org/10.1016/B978-0-12-820528-0.00010-7.

(19) Abo-Kadoum, M. A.; Abo-Dahab, N. F.; Awad, M. F.; Abdel-Hadi, A. M. Marine-Derived Fungus, Penicillium Aurantiogriseum AUMC 9757: A Producer of Bioactive Secondary Metabolites. *J. Basic Appl. Mycol.* **2013**, *4*, 77–83.

(20) Zou, W. X.; Meng, J. C.; Lu, H.; Chen, G. X.; Shi, G. X.; Zhang, T. Y.; Tan, R. X. Metabolites of Colletotrichum Gloeosporioides, an Endophytic Fungus in Artemisia Mongolica. *Journal of Natural Products* **2000**, *63* (11), 1529–1530. https://doi.org/10.1021/np000204t.

(21) Qin, J.-C.; Zhang, Y.-M.; Hu, L.; Ma, Y.-T.; Gao, J.-M. Cytotoxic Metabolites Produced by Alternaria No.28, an Endophytic Fungus Isolated from Ginkgo Biloba. *Natural Product Communications* **2009**, *4* (11), 1473–1476. https://doi.org/10.1177/1934578x0900401106.

(22) Mapari, S. A. S.; Meyer, A. S.; Thrane, U. Photostability of Natural Orange−Red and Yellow Fungal Pigments in Liquid Food Model Systems. *J. Agric. Food Chem.* **2009**, *57* (14), 6253–6261. https://doi.org/10.1021/jf900113q.

(23) Durán, N.; Teixeira, M. F. S.; De Conti, R.; Esposito, E. Ecological-Friendly Pigments from Fungi. *Critical Reviews in Food Science and Nutrition* **2002**, *42* (1), 53–66. https://doi.org/10.1080/10408690290825457.

(24) Keller, N. P. Fungal Secondary Metabolism: Regulation, Function and Drug Discovery. *Nature Reviews Microbiology* **2019**, *17* (3), 167–180. https://doi.org/10.1038/s41579-018-0121-1.

(25) Chiang, Y.-M.; Oakley, B. R.; Keller, N. P.; Wang, C. C. C. Unraveling Polyketide Synthesis in Members of the Genus Aspergillus. *Appl Microbiol Biotechnol* **2010**, *86* (6), 1719–1736. https://doi.org/10.1007/s00253-010-2525-3.

(26) Hertweck, C.; Luzhetskyy, A.; Rebets, Y.; Bechthold, A. Type II Polyketide Synthases: Gaining a Deeper Insight into Enzymatic Teamwork. *Nat. Prod. Rep.* **2007**, *24* (1), 162–190. https://doi.org/10.1039/B507395M.

(27) Austin, M. B.; Noel, J. P. The Chalcone Synthase Superfamily of Type III Polyketide Synthases. *Nat. Prod. Rep.* **2003**, *20* (1), 79–110. https://doi.org/10.1039/b100917f.

(28) Katsuyama, Y.; Ohnishi, Y. Type III Polyketide Synthases in Microorganisms. In *Methods in Enzymology*; Elsevier, 2012; Vol. 515, pp 359–377. https://doi.org/10.1016/B978-0-12-394290-6.00017-3.

(29) Hopwood, D. A.; Sherman, D. H. MOLECULAR GENETICS OF POLYKETIDES AND ITS COMPARISON TO FATTY ACID BIOSYNTHESIS. *Annu. Rev. Genet.* **1990**, *24* (1), 37–62. https://doi.org/10.1146/annurev.ge.24.120190.000345.

(30) Staunton, J. The Extraordinary Enzymes Involved in Erythromycin Biosynthesis. *Angewandte Chemie International Edition in English* **1991**, *30* (10), 1302–1306. https://doi.org/10.1002/anie.199113021.

(31) Kennedy Jonathan; Auclair Karine; Kendrew Steven G.; Park Cheonseok; Vederas John C.; Richard Hutchinson C. Modulation of Polyketide Synthase Activity by Accessory Proteins During Lovastatin Biosynthesis. *Science* **1999**, *284* (5418), 1368–1372. https://doi.org/10.1126/science.284.5418.1368.

(32) Staunton, J.; Weissman, K. J. Polyketide Biosynthesis: A Millennium Review. *Nat. Prod. Rep.* **2001**, *18* (4), 380–416. https://doi.org/10.1039/A909079G.

(33) Shen, B. Biosynthesis of Aromatic Polyketides. 242.

(34) Wang, B.; Guo, F.; Huang, C.; Zhao, H. Unraveling the Iterative Type I Polyketide Synthases Hidden in Streptomyces. *Proc Natl Acad Sci USA* **2020**, *117* (15), 8449. https://doi.org/10.1073/pnas.1917664117.

(35) J. Rawlings, B. Biosynthesis of Polyketides (Other than Actinomycete Macrolides). *Nat. Prod. Rep.* **1999**, *16* (4), 425–484. https://doi.org/10.1039/A900566H.

(36) Abe, I.; Morita, H. Structure and Function of the Chalcone Synthase Superfamily of Plant Type III Polyketide Synthases. *Nat. Prod. Rep.* **2010**, *27* (6), 809–838. https://doi.org/10.1039/B909988N.

(37) Hashimoto, M.; Nonaka, T.; Fujii, I. Fungal Type III Polyketide Synthases. *Nat. Prod. Rep.* **2014**, *31* (10), 1306–1317. https://doi.org/10.1039/C4NP00096J.

(38) Funa, N.; Ozawa, H.; Hirata, A.; Horinouchi, S. Phenolic Lipid Synthesis by Type III Polyketide Synthases Is Essential for Cyst Formation in Azotobacter Vinelandii</Em>. *Proc Natl Acad Sci USA* **2006**, *103* (16), 6356. https://doi.org/10.1073/pnas.0511227103.

(39) Funa, N.; Funabashi, M.; Ohnishi, Y.; Horinouchi, S. Biosynthesis of Hexahydroxyperylenequinone Melanin via Oxidative Aryl Coupling by Cytochrome P-450 in Streptomyces Griseus. *J Bacteriol* **2005**, *187* (23), 8149–8155. https://doi.org/10.1128/JB.187.23.8149-8155.2005.

(40) Hur, G. H.; Vickery, C. R.; Burkart, M. D. Explorations of Catalytic Domains in Non-Ribosomal Peptide Synthetase Enzymology. *Nat. Prod. Rep.* **2012**, *29* (10), 1074. https://doi.org/10.1039/c2np20025b.

(41) Lee, B.-N.; Kroken, S.; Chou, D. Y. T.; Robbertse, B.; Yoder, O. C.; Turgeon, B. G. Functional Analysis of All Nonribosomal Peptide Synthetases in Cochliobolus Heterostrophus Reveals a Factor, NPS6, Involved in Virulence and Resistance to Oxidative Stress. *Eukaryot Cell* **2005**, *4* (3), 545–555. https://doi.org/10.1128/EC.4.3.545-555.2005.

(42) van Gulik, W. M.; de Laat, W. T. A. M.; Vinke, J. L.; Heijnen, J. J. Application of Metabolic Flux Analysis for the Identification of Metabolic Bottlenecks in the Biosynthesis of Penicillin-G. *Biotechnology and Bioengineering* **2000**, *68* (6), 602–618. https://doi.org/10.1002/(SICI)1097-0290(20000620)68:6<602::AID-BIT3>3.0.CO;2-2.

(43) Keating, T. A.; Marshall, C. G.; Walsh, C. T. Reconstitution and Characterization of the Vibrio Cholerae Vibriobactin Synthetase from VibB, VibE, VibF, and VibH. *Biochemistry* **2000**, *39* (50), 15522–15530. https://doi.org/10.1021/bi0016523.

(44) Xu, L.; Huang, H.; Wei, W.; Zhong, Y.; Tang, B.; Yuan, H.; Zhu, L.; Huang, W.; Ge, M.; Yang, S.; Zheng, H.; Jiang, W.; Chen, D.; Zhao, G.-P.; Zhao, W. Complete Genome Sequence and Comparative Genomic Analyses of the Vancomycin-Producing Amycolatopsis Orientalis. *BMC Genomics* **2014**, *15* (1), 363. https://doi.org/10.1186/1471-2164-15-363.

(45) Levine, D. P. Vancomycin: A History. *Clinical Infectious Diseases* **2006**, *42* (Supplement_1), S5–S12. https://doi.org/10.1086/491709.

(46) Griffiths, G. L.; Sigel, S. P.; Payne, S. M.; Neilands, J. B. Vibriobactin, a Siderophore from Vibrio Cholerae. *Journal of Biological Chemistry* **1984**, *259* (1), 383–385. https://doi.org/10.1016/S0021-9258(17)43671-4.

(47) Schmidt-Dannert, C. Biosynthesis of Terpenoid Natural Products in Fungi. In *Biotechnology of Isoprenoids*; Schrader, J., Bohlmann, J., Eds.; Springer International Publishing: Cham, 2015; pp 19–61. https://doi.org/10.1007/10_2014_283.

(48) Berg, J. M.; Tymoczko, J. L.; Stryer, L.; Stryer, L. 26.2. Cholesterol Is Synthesized from Acetyl Coenzyme A in Three Stages. In *Biochemistry*; W.H. Freeman: New York, 2007; pp 739–741.

(49) Tomoda, H.; Tabata, N.; Nakata, Y.; Nishida, H.; Kaneko, T.; Obata, R.; Sunazuka, T.; Ōmura, S. Biosynthesis of Pyripyropene A. *J. Org. Chem.* **1996**, *61* (3), 882–886. https://doi.org/10.1021/jo951424s.

(50) Itoh, T.; Tokunaga, K.; Matsuda, Y.; Fujii, I.; Abe, I.; Ebizuka, Y.; Kushiro, T. Reconstitution of a Fungal Meroterpenoid Biosynthesis Reveals the Involvement of a Novel Family of Terpene Cyclases. *Nature Chemistry* **2010**, *2* (10), 858–864. https://doi.org/10.1038/nchem.764.

(51) Das, A.; Davis, M. A.; Tomoda, H.; Ômura, S.; Rudel, L. L. Identification of the Interaction Site within Acyl-CoA:Cholesterol Acyltransferase 2 for the Isoform-Specific Inhibitor Pyripyropene A. *Journal of Biological Chemistry* **2008**, *283* (16), 10453–10460. https://doi.org/10.1074/jbc.M709460200.

(52) Lincke, T.; Behnken, S.; Ishida, K.; Roth, M.; Hertweck, C. Closthioamide: An Unprecedented Polythioamide Antibiotic from the Strictly Anaerobic Bacterium Clostridium Cellulolyticum. *Angewandte Chemie* **2010**, *122* (11), 2055–2057. https://doi.org/10.1002/ange.200906114.

(53) Gomez-Escribano, J. P.; Song, L.; Fox, D. J.; Yeo, V.; Bibb, M. J.; Challis, G. L. Structure and Biosynthesis of the Unusual Polyketide Alkaloid Coelimycin P1, a Metabolic Product of the Cpk Gene Cluster of Streptomyces Coelicolor M145. *Chem. Sci.* **2012**, *3* (9), 2716. https://doi.org/10.1039/c2sc20410j.

(54) Hosaka, T.; Ohnishi-Kameyama, M.; Muramatsu, H.; Murakami, K.; Tsurumi, Y.; Kodani, S.; Yoshida, M.; Fujie, A.; Ochi, K. Antibacterial Discovery in Actinomycetes Strains with Mutations in RNA Polymerase or Ribosomal Protein S12. *Nature Biotechnology* **2009**, *27* (5), 462–464. https://doi.org/10.1038/nbt.1538.

(55) Lazarus, C. M.; Williams, K.; Bailey, A. M. Reconstructing Fungal Natural Product Biosynthetic Pathways. *Nat. Prod. Rep.* **2014**, *31* (10), 1339–1347. https://doi.org/10.1039/C4NP00084F.

(56) Rokas, A.; Wisecaver, J. H.; Lind, A. L. The Birth, Evolution and Death of Metabolic Gene Clusters in Fungi. *Nat Rev Microbiol* **2018**, *16* (12), 731–744. https://doi.org/10.1038/s41579-018-0075-3.

(57) Guo, L.; Winzer, T.; Yang, X.; Li, Y.; Ning, Z.; He, Z.; Teodor, R.; Lu, Y.; Bowser, T. A.; Graham, I. A.; Ye, K. The Opium Poppy Genome and Morphinan Production. *Science* **2018**, *362* (6412), 343. https://doi.org/10.1126/science.aat4096.

(58) Winzer, T.; Gazda, V.; He, Z.; Kaminski, F.; Kern, M.; Larson, T. R.; Li, Y.; Meade, F.; Teodor, R.; Vaistij, F. E.; Walker, C.; Bowser, T. A.; Graham, I. A. A Papaver Somniferum</Em> 10-Gene Cluster for Synthesis of the Anticancer Alkaloid Noscapine. *Science* **2012**, *336* (6089), 1704. https://doi.org/10.1126/science.1220757.

(59) Qi, X.; Bakht, S.; Leggett, M.; Maxwell, C.; Melton, R.; Osbourn, A. A Gene Cluster for Secondary Metabolism in Oat: Implications for the Evolution of Metabolic Diversity in Plants. *Proc Natl Acad Sci U S A* **2004**, *101* (21), 8233. https://doi.org/10.1073/pnas.0401301101.

(60) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. Drugs for Bad Bugs: Confronting the Challenges of Antibacterial Discovery. *Nat Rev Drug Discov* **2007**, *6* (1), 29–40. https://doi.org/10.1038/nrd2201.

(61) Hajduk, P. J.; Greer, J. A Decade of Fragment-Based Drug Design: Strategic Advances and Lessons Learned. *Nat Rev Drug Discov* **2007**, *6* (3), 211–219. https://doi.org/10.1038/nrd2220.

(62) Rutledge, P. J.; Challis, G. L. Discovery of Microbial Natural Products by Activation of Silent Biosynthetic Gene Clusters. *Nat Rev Microbiol* **2015**, *13* (8), 509–523. https://doi.org/10.1038/nrmicro3496.

(63) Blin, K.; Wolf, T.; Chevrette, M. G.; Lu, X.; Schwalen, C. J.; Kautsar, S. A.; Suarez Duran, H. G.; de los Santos, E. L. C.; Kim, H. U.; Nave, M.; Dickschat, J. S.; Mitchell, D. A.; Shelest, E.; Breitling, R.; Takano, E.; Lee, S. Y.; Weber, T.; Medema, M. H. AntiSMASH 4.0—Improvements in Chemistry Prediction and Gene Cluster Boundary Identification. *Nucleic Acids Research* **2017**, *45* (W1), W36–W41. https://doi.org/10.1093/nar/gkx319.

(64) Skinnider, M. A.; Merwin, N. J.; Johnston, C. W.; Magarvey, N. A. PRISM 3: Expanded Prediction of Natural Product Chemical Structures from Microbial Genomes. *Nucleic Acids Research* **2017**, *45* (W1), W49–W54. https://doi.org/10.1093/nar/gkx320.

(65) Khaldi, N.; Seifuddin, F. T.; Turner, G.; Haft, D.; Nierman, W. C.; Wolfe, K. H.; Fedorova, N. D. SMURF: Genomic Mapping of Fungal Secondary Metabolite Clusters. *Fungal Genetics and Biology* **2010**, *47* (9), 736–741. https://doi.org/10.1016/j.fgb.2010.06.003.

(66) van Heel, A. J.; de Jong, A.; Montalbán-López, M.; Kok, J.; Kuipers, O. P. BAGEL3: Automated Identification of Genes Encoding Bacteriocins and (Non-)Bactericidal Posttranslationally Modified Peptides. *Nucleic Acids Research* **2013**, *41* (W1), W448–W453. https://doi.org/10.1093/nar/gkt391.

(67) Scherlach, K.; Hertweck, C. Discovery of Aspoquinolones A–D, Prenylated Quinoline-2-One Alkaloids from Aspergillus Nidulans, Motivated by Genome Mining. *Org. Biomol. Chem.* **2006**, *4* (18), 3517–3520. https://doi.org/10.1039/B607011F.

(68) Williams, R. B.; Henrikson, J. C.; Hoover, A. R.; Lee, A. E.; Cichewicz, R. H. Epigenetic Remodeling of the Fungal Secondary Metabolome. *Org. Biomol. Chem.* **2008**, *6* (11), 1895. https://doi.org/10.1039/b804701d.

(69) Chiang, Y.-M.; Szewczyk, E.; Davidson, A. D.; Keller, N.; Oakley, B. R.; Wang, C. C. C. A Gene Cluster Containing Two Fungal Polyketide Synthases Encodes the Biosynthetic Pathway for a Polyketide, Asperfuranone, in Aspergillus Nidulans. *J. Am. Chem. Soc.* **2009**, *131* (8), 2965–2970. https://doi.org/10.1021/ja8088185.

(70) Guo, F.; Xiang, S.; Li, L.; Wang, B.; Rajasärkkä, J.; Gröndahl-Yli-Hannuksela, K.; Ai, G.; Metsä-Ketelä, M.; Yang, K. Targeted Activation of Silent Natural Product Biosynthesis Pathways by Reporter-Guided Mutant Selection. *Metabolic Engineering* **2015**, *28*, 134–142. https://doi.org/10.1016/j.ymben.2014.12.006.

(71) Franke, J.; Ishida, K.; Hertweck, C. Genomics-Driven Discovery of Burkholderic Acid, a Noncanonical, Cryptic Polyketide from Human Pathogenic Burkholderia Species. *Angewandte Chemie International Edition* **2012**, *51* (46), 11611–11615. https://doi.org/10.1002/anie.201205566.

(72) Chou, W. K. W.; Fanizza, I.; Uchiyama, T.; Komatsu, M.; Ikeda, H.; Cane, D. E. Genome Mining in Streptomyces Avermitilis: Cloning and Characterization of SAV_76, the Synthase for a New Sesquiterpene, Avermitilol. *J. Am. Chem. Soc.* **2010**, *132* (26), 8850–8851. https://doi.org/10.1021/ja103087w.

(73) Mukherjee, P. K.; Horwitz, B. A.; Kenerley, C. M. Secondary Metabolism in Trichoderma – a Genomic Perspective. *Microbiology* **2012**, *158* (1), 35–45. https://doi.org/10.1099/mic.0.053629-0.

(74) Cram, D. J.; Tishler, M. Mold Metabolites. I. Isolation of Several Compounds from Clinical Penicillin. *J. Am. Chem. Soc.* **1948**, *70* (12), 4238–4239. https://doi.org/10.1021/ja01192a076.

(75) Cram, D. J. Mold Metabolites. II. The Structure of Sorbicillin, a Pigment Produced by the Mold Penicillium Notatum. *J. Am. Chem. Soc.* **1948**, *70* (12), 4240–4243. https://doi.org/10.1021/ja01192a077.

(76) ABE, N.; YAMAMOTO, K.; HIROTA, A. Novel Fungal Metabolites, Demethylsorbicillin and Oxosorbicillinol, Isolated from Trichoderma Sp. USF-2690. *Bioscience, Biotechnology, and Biochemistry* **2000**, *64* (3), 620–622. https://doi.org/10.1271/bbb.64.620.

(77) Derntl, C.; Rassinger, A.; Srebotnik, E.; Mach, R. L.; Mach-Aigner, A. R. Identification of the Main Regulator Responsible for Synthesis of the Typical Yellow Pigment Produced by T. Reesei. *Appl. Environ. Microbiol.* **2016**, *82* (20), 6247. https://doi.org/10.1128/AEM.01408-16.

(78) Derntl, C.; Guzmán-Chávez, F.; Mello-de-Sousa, T. M.; Busse, H.-J.; Driessen, A. J. M.; Mach, R. L.; Mach-Aigner, A. R. In Vivo Study of the Sorbicillinoid Gene Cluster in Trichoderma Reesei. *Frontiers in Microbiology* **2017**, *8*, 2037. https://doi.org/10.3389/fmicb.2017.02037.

(79) Salo Oleksandr; Guzmán-Chávez Fernando; Ries Marco I.; Lankhorst Peter P.; Bovenberg Roel A. L.; Vreeken Rob J.; Driessen Arnold J. M.; Brakhage A. A. Identification of a Polyketide Synthase Involved in Sorbicillin Biosynthesis by Penicillium Chrysogenum. *Applied and Environmental Microbiology* **2016**, *82* (13), 3971–3978. https://doi.org/10.1128/AEM.00350-16.

(80) Reeves Wendy M.; Hahn Steven. Targets of the Gal4 Transcription Activator in Functional Transcription Complexes. *Molecular and Cellular Biology* **2005**, *25* (20), 9092–9102. https://doi.org/10.1128/MCB.25.20.9092-9102.2005.

(81) Lettow, J.; Aref, R.; Schüller, H.-J. Transcriptional Repressor Gal80 Recruits Corepressor Complex Cyc8–Tup1 to Structural Genes of the Saccharomyces Cerevisiae GAL Regulon. *Current Genetics* **2021**. https://doi.org/10.1007/s00294-021-01215-x.

(82) Berg, J. M.; Tymoczko, J. L.; Stryer, L.; Stryer, L. 31.3. The Greater Complexity of Eukaryotic Genomes Requires Elaborate Mechanisms for Gene Regulation. In *Biochemistry*; W.H. Freeman: New York, 2007; pp 902–903.

(83) Berg, J. M.; Tymoczko, J. L.; Stryer, L.; Stryer, L. 31.1. Many DNA-Binding Proteins Recognize Specific DNA Sequences. In *Biochemistry*; W.H. Freeman: New York, 2007; pp 895–896.

(84) MacPherson Sarah; Larochelle Marc; Turcotte Bernard. A Fungal Family of Transcriptional Regulators: The Zinc Cluster Proteins. *Microbiology and Molecular Biology Reviews* **2006**, *70* (3), 583–604. https://doi.org/10.1128/MMBR.00015-06.

(85) Wolfe, S. A.; Nekludova, L.; Pabo, C. O. DNA Recognition by Cys2His2 Zinc Finger Proteins. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29* (1), 183–212. https://doi.org/10.1146/annurev.biophys.29.1.183.

(86) Klar, A. J. S.; Halvorson, H. O. Studies on the Positive Regulatory Gene, GAL4, in Regulation of Galactose Catabolic Enzymes in Saccharomyces Cerevisiae. *Molec. Gen. Genet.* **1974**, *135* (3), 203–212. https://doi.org/10.1007/BF00268616.

(87) Strich, R.; Surosky, R. T.; Steber, C.; Dubois, E.; Messenguy, F.; Esposito, R. E. UME6 !S a Key Regulator of Nitrogen Repression and Meiotic Development. 16.

(88) Johnston, M. A Model Fungal Gene Regulatory Mechanism: The GAL Genes of Saccharomyces Cerevisiae. *MICROBIOL. REV.* **1987**, *51*, 19.

(89) Schjerling, P.; Holmberg, S. Comparative Amino Acid Sequence Analysis of the C6 Zinc Cluster Family of Transcriptional Regulators. *Nucleic Acids Research* **1996**, *24* (23), 4599–4607. https://doi.org/10.1093/nar/24.23.4599.

(90) Marmorstein, R.; Carey, M.; Ptashne, M.; Harrison, S. C. DNA Recognition by GAL4: Structure of a Protein-DNA Complex. *Nature* **1992**, *356* (6368), 408–414. https://doi.org/10.1038/356408a0.

(91) Delaveau, T.; Delahodde, A.; Carvajal, E.; Subik, J.; Jacq, C. PDR3, a New Yeast Regulatory Gene, Is Homologous ToPDR1 and Controls the Multidrug Resistance Phenomenon. *Molecular and General Genetics MGG* **1994**, *244* (5), 501–511. https://doi.org/10.1007/BF00583901.

(92) Kolaczkowska, A.; Goffeau, A. Regulation of Pleiotropic Drug Resistance in Yeast. *Drug Resistance Updates* **1999**, *2* (6), 403–414. https://doi.org/10.1054/drup.1999.0113.

(93) Fields, S.; Song, O. A Novel Genetic System to Detect Protein–Protein Interactions. *Nature* **1989**, *340* (6230), 245–246. https://doi.org/10.1038/340245a0.

(94) Brent, R.; Ptashne, M. A Eukaryotic Transcriptional Activator Bearing the DNA Specificity of a Prokaryotic Repressor. *Cell* **1985**, *43* (3, Part 2), 729–736. https://doi.org/10.1016/0092-8674(85)90246-6.

(95) Scheer, N.; Campos-Ortega, J. A. Use of the Gal4-UAS Technique for Targeted Gene Expression in the Zebrafish. *Mechanisms of Development* **1999**, *80* (2), 153–158. https://doi.org/10.1016/S0925-4773(98)00209-3.

(96) Klueg, K. M.; Alvarado, D.; Muskavitch, M. A. T.; Duffy, J. B. Creation of a GAL4/UAS-Coupled Inducible Gene Expression System for Use Indrosophila Cultured Cell Lines. *genesis* **2002**, *34* (1–2), 119–122. https://doi.org/10.1002/gene.10148.

(97) Imamura, M.; Nakai, J.; Inoue, S.; Quan, G. X.; Kanda, T.; Tamura, T. Targeted Gene Expression Using the GAL4/UAS System in the Silkworm Bombyx Mori. *Genetics* **2003**, *165* (3), 1329–1340. https://doi.org/10.1093/genetics/165.3.1329.

(98) Grau, M. F.; Entwistle, R.; Chiang, Y.-M.; Ahuja, M.; Oakley, C. E.; Akashi, T.; Wang, C. C. C.; Todd, R. B.; Oakley, B. R. Hybrid Transcription Factor Engineering Activates the Silent Secondary Metabolite Gene Cluster for (+)-Asperlin in Aspergillus Nidulans. *ACS Chem. Biol.* **2018**, *13* (11), 3193–3205. https://doi.org/10.1021/acschembio.8b00679.

(99) Chiang, Y.-M.; Ahuja, M.; Oakley, C. E.; Entwistle, R.; Asokan, A.; Zutz, C.; Wang, C. C. C.; Oakley, B. R. Development of Genetic Dereplication Strains in Aspergillus Nidulans Results in the Discovery of Aspercryptin. *Angewandte Chemie International Edition* **2016**, *55* (5), 1662–1665. https://doi.org/10.1002/anie.201507097.

(100) Szewczyk, E.; Nayak, T.; Oakley, C. E.; Edgerton, H.; Xiong, Y.; Taheri-Talesh, N.; Osmani, S. A.; Oakley, B. R. Fusion PCR and Gene Targeting in Aspergillus Nidulans. *Nat Protoc* **2006**, *1* (6), 3111–3120. https://doi.org/10.1038/nprot.2006.405.

(101) Laureti, L.; Song, L.; Huang, S.; Corre, C.; Leblond, P.; Challis, G. L.; Aigle, B. Identification of a Bioactive 51-Membered Macrolide Complex by Activation of a Silent Polyketide Synthase in Streptomyces Ambofaciens. *Proceedings of the National Academy of Sciences* **2011**, *108* (15), 6258–6263. https://doi.org/10.1073/pnas.1019077108.

(102) Derntl, C.; Mach, R. L.; Mach-Aigner, A. R. Fusion Transcription Factors for Strong, Constitutive Expression of Cellulases and Xylanases in Trichoderma Reesei. *Biotechnol Biofuels* **2019**, *12* (1), 231. https://doi.org/10.1186/s13068-019-1575-8.

(103) Derntl, C.; Mach, R.; Mach-Aigner, A. Application of the Human Estrogen Receptor within a Synthetic Transcription Factor in Trichoderma Reesei. *Fungal Biology and Biotechnology* **2020**, *7* (1), 12. https://doi.org/10.1186/s40694-020-00102-3.

(104) Zhang, G.; Liu, P.; Wei, W.; Wang, X.; Wei, D.; Wang, W. A Light-Switchable Bidirectional Expression System in Filamentous Fungus Trichoderma Reesei. *Journal of Biotechnology* **2016**, *240*, 85–93. https://doi.org/10.1016/j.jbiotec.2016.11.003.

(105) Martinez, D.; Berka, R. M.; Henrissat, B.; Saloheimo, M.; Arvas, M.; Baker, S. E.; Chapman, J.; Chertkov, O.; Coutinho, P. M.; Cullen, D.; Danchin, E. G. J.; Grigoriev, I. V.; Harris, P.; Jackson, M.; Kubicek, C. P.; Han, C. S.; Ho, I.; Larrondo, L. F.; de Leon, A. L.; Magnuson, J. K.; Merino, S.; Misra, M.; Nelson, B.; Putnam, N.; Robbertse, B.; Salamov, A. A.; Schmoll, M.; Terry, A.; Thayer, N.; Westerholm-Parvinen, A.; Schoch, C. L.; Yao, J.; Barabote, R.; Nelson, M. A.; Detter, C.; Bruce, D.; Kuske, C. R.; Xie, G.; Richardson, P.; Rokhsar, D. S.; Lucas, S. M.; Rubin, E. M.; Dunn-Coleman, N.; Ward, M.; Brettin, T. S. Genome Sequencing and Analysis of the Biomass-Degrading Fungus Trichoderma Reesei (Syn. Hypocrea Jecorina). *Nature Biotechnology* **2008**, *26* (5), 553–560. https://doi.org/10.1038/nbt1403.

(106) Joint Genome Institute. **2021**. https://mycocosm.jgi.doe.gov/Trire2/Trire2.home.html.

(107) Cimermancic, P.; Medema, M. H.; Claesen, J.; Kurita, K.; Wieland Brown, L. C.; Mavrommatis, K.; Pati, A.; Godfrey, P. A.; Koehrsen, M.; Clardy, J.; Birren, B. W.; Takano, E.; Sali, A.; Linington, R. G.; Fischbach, M. A. Insights into Secondary Metabolism from a Global Analysis of Prokaryotic Biosynthetic Gene Clusters. *Cell* **2014**, *158* (2), 412–421. https://doi.org/10.1016/j.cell.2014.06.034.

(108) Törönen, P.; Medlar, A.; Holm, L. PANNZER2: A Rapid Functional Annotation Web Server. *Nucleic Acids Research* **2018**, *46* (W1), W84–W88. https://doi.org/10.1093/nar/gky350.

(109) Somervuo, P.; Holm, L. SANSparallel: Interactive Homology Search against Uniprot. *Nucleic Acids Res* **2015**, *43* (W1), W24–W29. https://doi.org/10.1093/nar/gkv317.

(110) Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. *Nucleic Acids Research* **1997**, *25* (17), 3389–3402. https://doi.org/10.1093/nar/25.17.3389.

(111) Katoh, K.; Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* **2013**, *30* (4), 772–780. https://doi.org/10.1093/molbev/mst010.

(112) *Benchling [Biology Software], (2021). Retrieved from Https://Benchling.Com*; Benchling, 2021.

(113) Steiger, M. G.; Vitikainen, M.; Uskonen, P.; Brunner, K.; Adam, G.; Pakula, T.; Penttilä, M.; Saloheimo, M.; Mach, R. L.; Mach-Aigner, A. R. Transformation System for Trichoderma Reesei That Favors Homologous Integration and Employs Reusable Bidirectionally Selectable Markers. *Appl. Environ. Microbiol.* **2011**, *77* (1), 114. https://doi.org/10.1128/AEM.02100-10.

(114) Smith, J. L.; Bayliss, F. T.; Ward, M. Sequence of the Cloned Pyr4 Gene of Trichoderma Reesei and Its Use as a Homologous Selectable Marker for Transformation. *Curr Genet* **1991**, *19* (1), 27–33. https://doi.org/10.1007/BF00362084.

(115) Vignolle, G. A.; Schaffer, D.; Zehetner, L.; Mach, R. L.; Mach-Aigner, A. R.; Derntl, C. FunOrder: A Robust and Semi-Automated Method for the Identification of Essential Biosynthetic Genes through Computational Molecular Co-Evolution. *PLoS Comput Biol* **2021**, *17* (9), e1009372. https://doi.org/10.1371/journal.pcbi.1009372.

(116) Marcet-Houben, M.; Gabaldón, T. TreeKO: A Duplication-Aware Algorithm for the Comparison of Phylogenetic Trees. *Nucleic Acids Research* **2011**, *39* (10), e66–e66. https://doi.org/10.1093/nar/gkr087.

(117) Kulkarni, R. D.; Thon, M. R.; Pan, H.; Dean, R. A. Novel G-Protein-Coupled Receptor-like Proteins in the Plant Pathogenic Fungus Magnaporthe Grisea. *Genome Biology* **2005**, *6* (3), R24. https://doi.org/10.1186/gb-2005-6-3-r24.

(118) Rogers, L. A. THE INHIBITING EFFECT OF STREPTOCOCCUS LACTIS ON LACTOBACILLUS BULGARICUS. *J Bacteriol* **1928**, *16* (5), 321–325. https://doi.org/10.1128/jb.16.5.321-325.1928.

(119) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K.-D.; Fischbach, M. A.; Garavelli, J. S.; Göransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Müller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J. T.; Rebuffat, S.; Ross, R. P.; Sahl, H.-G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süssmuth, R. D.; Tagg, J. R.; Tang, G.-L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A. Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Products: Overview and Recommendations for a Universal Nomenclature. *Nat. Prod. Rep.* **2013**, *30* (1), 108–160. https://doi.org/10.1039/C2NP20085F.

(120) Umemura, M. Peptides Derived from Kex2-Processed Repeat Proteins Are Widely Distributed and Highly Diverse in the Fungi Kingdom. *Fungal Biology and Biotechnology* **2020**, *7* (1), 11. https://doi.org/10.1186/s40694-020-00100-5.

(121) Vignolle, G. A.; Schaffer, D.; Mach, R. L.; Mach-Aigner, A. R.; Derntl, C. The Functional Order (FunOrder) Tool – Identification of Essential Biosynthetic Genes through Computational Molecular Co-Evolution. *bioRxiv* **2021**, 2021.01.29.428829. https://doi.org/10.1101/2021.01.29.428829.

(122) Lux, S. E.; John, K. M.; Bennett, V. Analysis of CDNA for Human Erythrocyte Ankyrin Indicates a Repeated Structure with Homology to Tissue-Differentiation and Cell-Cycle Control Proteins. *Nature* **1990**, *344* (6261), 36–42. https://doi.org/10.1038/344036a0.

(123) Derntl, C.; Gudynaite-Savitch, L.; Calixte, S.; White, T.; Mach, R. L.; Mach-Aigner, A. R. Mutation of the Xylanase Regulator 1 Causes a Glucose Blind Hydrolase Expressing Phenotype in Industrially Used Trichoderma Strains. *Biotechnol Biofuels* **2013**, *6* (1), 62. https://doi.org/10.1186/1754-6834-6-62.

(124) Stricker, A. R.; Mach, R. L.; de Graaff, L. H. Regulation of Transcription of Cellulases- and Hemicellulases-Encoding Genes in Aspergillus Niger and Hypocrea Jecorina (Trichoderma Reesei). *Appl Microbiol Biotechnol* **2008**, *78* (2), 211–220. https://doi.org/10.1007/s00253-007-1322-0.

(125) Mach-Aigner Astrid R.; Pucher Marion E.; Mach Robert L. D-Xylose as a Repressor or Inducer of Xylanase Expression in Hypocrea Jecorina (Trichoderma Reesei). *Applied and Environmental Microbiology* **2010**, *76* (6), 1770–1776. https://doi.org/10.1128/AEM.02746-09.

(126) Wang, L.; Zhang, W.; Cao, Y.; Zheng, F.; Zhao, G.; Lv, X.; Meng, X.; Liu, W. Interdependent Recruitment of CYC8/TUP1 and the Transcriptional Activator XYR1 at Target Promoters Is Required for Induced Cellulase Gene Expression in Trichoderma Reesei. *PLoS Genet* **2021**, *17* (2), e1009351. https://doi.org/10.1371/journal.pgen.1009351.

(127) Portnoy Thomas; Margeot Antoine; Seidl-Seiboth Verena; Le Crom Stéphane; Ben Chaabane Fadhel; Linke Rita; Seiboth Bernhard; Kubicek Christian P. Differential Regulation of the Cellulase Transcription Factors XYR1, ACE2, and ACE1 in Trichoderma Reesei Strains Producing High and Low Levels of Cellulase. *Eukaryotic Cell* **2011**, *10* (2), 262–271. https://doi.org/10.1128/EC.00208-10.

(128) Southern, E. Southern Blotting. *Nat Protoc* **2006**, *1* (2), 518–525. https://doi.org/10.1038/nprot.2006.73.

(129) Kautsar, S. A.; Blin, K.; Shaw, S.; Navarro-Muñoz, J. C.; Terlouw, B. R.; van der Hooft, J. J. J.; van Santen, J. A.; Tracanna, V.; Suarez Duran, H. G.; Pascal Andreu, V.; Selem-Mojica, N.; Alanjary, M.; Robinson, S. L.; Lund, G.; Epstein, S. C.; Sisto, A. C.; Charkoudian, L. K.; Collemare, J.; Linington, R. G.; Weber, T.; Medema, M. H. MIBiG 2.0: A Repository for Biosynthetic Gene Clusters of Known Function. *Nucleic Acids Research* **2020**, *48* (D1), D454–D458. https://doi.org/10.1093/nar/gkz882.

(130) Wat, C.-K.; Mcinnes, A. G.; Smith, D. G.; Wright, J. L. C.; Vining, L. C. The Yellow Pigments of *Beauveria* Species. Structures of Tenellin and Bassianin. *Can. J. Chem.* **1977**, *55* (23), 4090–4098. https://doi.org/10.1139/v77-580.

(131) Fisch, K. M.; Bakeer, W.; Yakasai, A. A.; Song, Z.; Pedrick, J.; Wasil, Z.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J.; Cox, R. J. Rational Domain Swaps Decipher Programming in Fungal Highly Reducing Polyketide Synthases and Resurrect an Extinct Metabolite. *J. Am. Chem. Soc.* **2011**, *133* (41), 16635–16641. https://doi.org/10.1021/ja206914q.

(132) Eley, K. L.; Halo, L. M.; Song, Z.; Powles, H.; Cox, R. J.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J. Biosynthesis of the 2-Pyridone Tenellin in the Insect Pathogenic FungusBeauveria Bassiana. *ChemBioChem* **2007**, *8* (3), 289–297. https://doi.org/10.1002/cbic.200600398.

(133) Liu, L.; Zhang, J.; Chen, C.; Teng, J.; Wang, C.; Luo, D. Structure and Biosynthesis of Fumosorinone, a New Protein Tyrosine Phosphatase 1B Inhibitor Firstly Isolated from the Entomogenous Fungus Isaria Fumosorosea. *Fungal Genetics and Biology* **2015**, *81*, 191–200. https://doi.org/10.1016/j.fgb.2015.03.009.

(134) Bergmann, S.; Schümann, J.; Scherlach, K.; Lange, C.; Brakhage, A. A.; Hertweck, C. Genomics-Driven Discovery of PKS-NRPS Hybrid Metabolites from Aspergillus Nidulans. *Nature Chemical Biology* **2007**, *3* (4), 213–217. https://doi.org/10.1038/nchembio869.

(135) Schümann, J.; Hertweck, C. Molecular Basis of Cytochalasan Biosynthesis in Fungi: Gene Cluster Analysis and Evidence for the Involvement of a PKS-NRPS Hybrid Synthase by RNA Silencing. *J. Am. Chem. Soc.* **2007**, *129* (31), 9564–9565. https://doi.org/10.1021/ja072884t.

(136) Ellestad, G. A.; Evans, R. H.; Kunstmann, M. P. Some New Terpenoid Metabolites from an Unidentified Fusarium Species. *Tetrahedron* **1969**, *25* (6), 1323–1334. https://doi.org/10.1016/S0040-4020(01)82703-4.

(137) Li, C.; Matsuda, Y.; Gao, H.; Hu, D.; Yao, X. S.; Abe, I. Biosynthesis of LL-Z1272β: Discovery of a New Member of NRPS-like Enzymes for Aryl-Aldehyde Formation. *ChemBioChem* **2016**, *17* (10), 904–907. https://doi.org/10.1002/cbic.201600087.

(138) Sugawara, T; Shinonaga, H.; Yamamoto, K. Polyene-Based Compounds. 319289, 1996.

(139) Krasnoff, S. B.; Sommers, C. H.; Moon, Y.-S.; Donzelli, B. G. G.; Vandenberg, J. D.; Churchill, A. C. L.; Gibson, D. M. Production of Mutagenic Metabolites by Metarhizium Anisopliae. *J. Agric. Food Chem.* **2006**, *54* (19), 7083–7088. https://doi.org/10.1021/jf061405r.

(140) Donzelli, B. G. G.; Krasnoff, S. B.; Churchill, A. C. L.; Vandenberg, J. D.; Gibson, D. M. Identification of a Hybrid PKS–NRPS Required for the Biosynthesis of NG-391 in Metarhizium Robertsii. *Current Genetics* **2010**, *56* (2), 151–162. https://doi.org/10.1007/s00294-010-0288-0.

(141) Ahuja, M.; Chiang, Y.-M.; Chang, S.-L.; Praseuth, M. B.; Entwistle, R.; Sanchez, J. F.; Lo, H.-C.; Yeh, H.-H.; Oakley, B. R.; Wang, C. C. C. Illuminating the Diversity of Aromatic Polyketide Synthases in Aspergillus Nidulans. *J. Am. Chem. Soc.* **2012**, *134* (19), 8212–8221. https://doi.org/10.1021/ja3016395.

(142) BERTANI, G. Studies on Lysogenesis. I. The Mode of Phage Liberation by Lysogenic Escherichia Coli. *J Bacteriol* **1951**, *62* (3), 293–300. https://doi.org/10.1128/jb.62.3.293-300.1951.

(143) Derntl, C.; Kiesenhofer, D. P.; Mach, R. L.; Mach-Aigner, A. R. Novel Strategies for Genomic Manipulation of Trichoderma Reesei with the Purpose of Strain Engineering. *Appl. Environ. Microbiol.* **2015**, *81* (18), 6314. https://doi.org/10.1128/AEM.01545-15.

(144) MANDELS, M.; PARRISH, F. W.; REESE, E. T. Sophorose as an Inducer of Cellulase in Trichoderma Viride. *J Bacteriol* **1962**, *83* (2), 400–408. https://doi.org/10.1128/jb.83.2.400-408.1962.

(145) Pfaffl, M. W. A New Mathematical Model for Relative Quantification in Real-Time RT-PCR. *Nucleic Acids Research* **2001**, *29* (9), 45e–445. https://doi.org/10.1093/nar/29.9.e45.

(146) Steiger, M. G.; Mach, R. L.; Mach-Aigner, A. R. An Accurate Normalization Strategy for RT-QPCR in Hypocrea Jecorina (Trichoderma Reesei). *Journal of Biotechnology* **2010**, *145* (1), 30–37. https://doi.org/10.1016/j.jbiotec.2009.10.012.

(147) Kluger, B.; Bueschl, C.; Neumann, N.; Stückler, R.; Doppler, M.; Chassy, A. W.; Waterhouse, A. L.; Rechthaler, J.; Kampleitner, N.; Thallinger, G. G.; Adam, G.; Krska, R.; Schuhmacher, R. Untargeted Profiling of Tracer-Derived Metabolites Using Stable Isotopic Labeling and Fast Polarity-Switching LC–ESI-HRMS. *Anal. Chem.* **2014**, *86* (23), 11533–11537. https://doi.org/10.1021/ac503290j.

(148) Klein, S.; Heinzle, E. Isotope Labeling Experiments in Metabolomics and Fluxomics. *WIREs Systems Biology and Medicine* **2012**, *4* (3), 261–272. https://doi.org/10.1002/wsbm.1167.