University of Thessaly

Faculty of Health Sciences

Department of Biochemistry and Biotechnology

Master's Program "**Advanced Experimental and Computational Biosciences**"

Genome editing of endophytic *Fusarium solani* strain K using a modular CRISPR/Cas9 toolkit

Nikolaos Delkis

June 2022

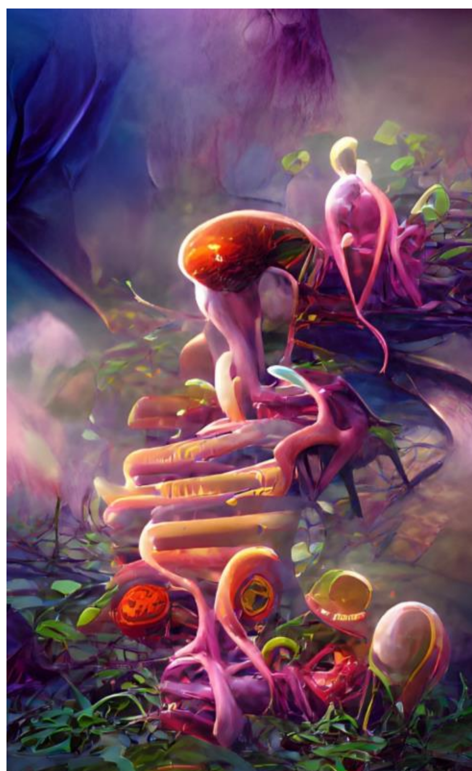
Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

Πανεπιστήμιο Θεσσαλίας

Σχολή Επιστημών Υγείας

Τμήμα Βιοχημείας και Βιοτεχνολογίας

Μεταπτυχιακό Πρόγραμμα Σπουδών «Προηγμένες Πειραματικές και Υπολογιστικές Βιοεπιστήμες»



Γονιδιωματική τροποποίηση του ενδόφυτου *Fusarium solani* στέλεχος K χρησιμοποιώντας ένα αρθρωτό σύστημα CRISPR/Cas9

Νικόλαος Ντελκής

Ιούνιος 2022

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

Supervisor

Kalliope K. Papadopoulou, Professor in Plant Biotechnology, Department of Biochemistry and Biotechnology, University of Thessaly

Examination Committee

Kostas Mathiopoulos, Professor in Molecular Biology

Antonios Giakountis, Assistant Professor in Molecular Biology - Genomics

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

Acknowledgements

As the journey through advanced experimental and computational biosciences comes to an end, I would like to acknowledge all those who contributed directly or indirectly to my personal development, and eventually to this master's thesis project.

A huge thank you goes to all the professors who helped us sharpen our critical thinking and work efficiently under pressure during the master's courses. Coping with everything was nearly impossible at the time, but in hindsight I am thankful of the effort you put in our proper scientific education and hopeful that you will continue at this pace.

I am also so glad to be working in the Plant and Environmental Biotechnology lab, where everyone made me feel like I was "at home". I thank all of you for the pleasant collaboration for all this time. Specifically, I would like to thank my supervisor Kalliope Papadopoulou for coping with my endless excitement and (often) inability to focus on what mattered most and provided the time and resources for me to do what I enjoy most. Also, all this work would not be completed without the relentless work ethic and effort from three brilliant young scientists, Antonia Argyri, Evanthia Georgiadou and Efthimia Zisopoulou, who conducted their Bachelor thesis with me in this and other projects this year. I hope our paths will cross again in the future.

Lastly, I would like to thank my family, for indiscriminately supporting my journey for so many years in all ways possible. I would not be here writing these words without you.

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

Table of Contents

Abstract 6
Introduction 8
Endophytic Fungi 8
Fusarium solani strain K (FsK) 8
Fungal RNAi machinery 8
Synthetic Biology 8
Engineering biology principles 9
Fungal Synthetic Biology 10
Type IIS modular assembly – FungalBraid 10
CRISPR functional genetics 11
Gene knockouts with CRISPR/Cas9 11
Gene regulation with catalytically inactive Cas9 12
CRISPR in fungal systems 12
Aims of the thesis 13
Materials and Methods 13
Results 15
FsK Type IIS basic part toolkit 15
Assembly and *Agrobacterium*-mediated FsK transformation of CRISPR vectors for RNAi genes 17
CRISPR system functionality on FsK RNAi machinery genes 23
Discussion 23
CRISPR system validation on FsK 23
CRISPR applications *Fusarium* and other filamentous fungi 24
Limitations in part assembly using FungalBraid 25
Future prospects for toolkit expansion 26
Conclusions 27
References 27

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

Abstract

Endophytic filamentous fungi play an important role in agriculture by colonizing plants and eliciting beneficial responses to their host. *Fusarium solani* strain K (FsK) is a beneficial endophyte able to colonize a diverse range of plant species and provide biotic and abiotic stress tolerance to tomato (*Solanum lycopersicum*). Very little is known on the fungal genes that are involved in the establishment of beneficial associations with plants and in the manipulation of the host responses. In addition, tools to silence or knock out fungal genes of interest for functional analysis in a high-throughput manner do not exist. Here, we develop a genome editing toolkit for the non-model endophyte FsK using CRISPR/Cas9 technology. To achieve this, we utilize Synthetic Biology principles and the FungalBraid modular cloning platform to generate the desired vectors for fungal transformation. We create a toolkit containing several endogenous and heterologous promoters, terminators, reporters and *Cas9* genes to maximize flexibility and *Cas9* expression. For a proof-of-concept experiment, we aim to disrupt GFP, from a transgenic FsK strain that has the reporter integrated into its genome. With this toolkit, we aim to streamline the creation of gene knockout or silencing FsK strains to help answer long-standing questions on the mode of action of beneficial endophytic fungi, and possibly apply it to different non-model filamentous fungi, of which the genome is known and possibly develop new fungal strains with improved traits tailored to specific applications in industry and agriculture.

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

Introduction

Endophytic Fungi

Plants, like humans and other species, closely interact with and rely on a vast microbiome to perform essential and complementary functions, to ultimately thrive in an ever-changing environment. Endophytic fungi are a class of such organisms and can colonize plant tissues without eliciting a defense response, while providing beneficial traits to the host[1], in exchange of carbon or other resources to survive. Such traits could be against abiotic stress[2], such as salt or drought tolerance, or biotic stress[3] , such as disease suppression from pathogens or other animals. Endophytic fungi are distinct to but often synergize with Arbuscular Mycorrhizal Fungi (AMF)[4], which also colonize plant tissues and elicit beneficial traits, but cannot survive outside of a host and cannot be genetically transformed, in contrast to endophytic fungi. Despite the increasing knowledge of plant–endophyte interactions, plant immune responses induced by endophytic fungal colonization against biotic and abiotic stresses are still not fully understood[1]. This is in part because the vast diversity, biology and functions of endophytic fungi cannot be captured by one model organism, as it is an umbrella term that covers thousands of species. Moreover, the ability to functionally study and dissect new non-model organisms is bottlenecked by the genetic tools available, which are often organism-specific.

Fusarium solani strain K (FsK)

Fusarium solani strain K (FsK) is an endophytic fungus that was discovered in tomato roots and provides protection against biotic[5,6] and abiotic[7] stresses, likely through the ethylene pathway[8], while having the ability to alter the microbial community in the plant rhizosphere[9]. FsK can colonize a broad range of hosts without eliciting pathogenicity[10] and uses the Common Symbiotic Signaling Pathway (CSSP) to enter its host, similar to AMF colonization, although the exact colonization mechanism is still to be fully understood[11]. Furthermore, FsK produces small RNA molecules which can be transferred to plants (Dalakouras et al., under peer review). Overall, FsK is a promising candidate organism to study further, but sufficient genetic tools to perform reverse genetics and elucidate single gene functions are not available.

Fungal RNAi machinery

Small RNA production and signaling, mainly through RNA interference (RNAi), is ubiquitous across all kingdoms of life[12]. RNAi plays an important role in fungi, not only for genome protection against viruses or transposable elements, but also for endogenous physiological mechanisms, such as gene regulation[13]. For example, the RNAi machinery of *Trichoderma atroviride* regulates growth and development and affects conidization[14], as shown indeletion mutants of the Dicer proteins of the species. However, little is known about the RNAi machinery of *Fusarium* species or other endophytic fungi. Phylogenetic analysis from our group revealed orthologs of Argonaute proteins 1 and 2 and Dicer-like proteins 1 and 2 in FsK's genome (Dalakouras et al., under peer review). Small RNA production and predicted RNAi machinery genes indicate that the RNAi pathway in FsK is not only functional, but integral to plant-host interactions. To elucidate their role and functionally characterize them, mutants and/or silencing lines of core RNAi machinery components are required.

Synthetic Biology

Synthetic biology or engineering biology is a multidisciplinary area of research that seeks to create new biological parts, devices, and systems, or to redesign systems that are already found in nature[15,16]. It is a

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

scientific intersection of many disciplines, including molecular biology, biotechnology, genetic engineering, systems biology, computer science, mathematics, and electrical engineering to name a few. A general notion of engineering biology is to employ methods and best practices from the engineering field, such as systems design, bottom-up approaches and the Design-Build-Test-Learn (DBTL) cycle, to the design of biological systems. This is inherently difficult, as all engineering systems have been designed from the human mind, and therefore all system components, elements and interactions are known, whereas a biologist usually follows a *reductio ad absurdum* approach to elucidate the components and interactions of a fraction of a biological system, due to high biological complexity. Nevertheless, biological engineers aim to adopt engineering principles with the ultimate goals of system predictability for existing and new-to-nature systems, lower cost for design, building and testing of the systems, automation and the creation of community-driven resources, such as collections of interchangeable DNA parts, to advance the field further[17,18].

Engineering biology principles

Engineering biology is greatly limited mainly by the enormous biological complexity, the often unpredictable and spontaneous variation of biological behavior that stems from this complexity, and evolution. A few foundational pillars or principles could aid synthetic biologists to navigate the murky path of biological complexity. These principles are standardization, decoupling, abstraction hierarchy and modularity[19]. Standardization helps form basis of human collaboration by providing common ground for everyone to contribute and exchange ideas, information or physical objects and has accelerated the pace of discovery in the last two centuries. In biology, standardized methods, protocols, and eventually physical biological parts, could guide researchers to share and advance their knowledge with higher reproducibility, lower cost and in a common trajectory with other people[20,21]. Decoupling is the idea that a complex problem can be solved by splitting it in smaller, easier, and manageable problems and solving those instead. It is a practice widely used in the engineering field since it allows different people to work on different aspects and problems of a system. A similar approach can be used in engineering biology, where different engineers work on different aspects of a systems. For example, one engineer would build a robust chassis that can host heterologous expression of complex devices, while a second engineer would create a range of different devices for a third engineer to test inside the chassis built by the first engineer. The idea of decoupling is closely correlated with modularity and abstraction hierarchy. Modularity is the principle of designing devices and systems in modules, where each module can be plugged or unplugged without interfering with other modules[22,23]. An example of modularity for biology could be modular basic parts, that carry information for regulatory elements or coding sequences, that once verified and characterized can be re-used by others. Based on these modular parts, a combinatorial metabolic pathway could be built, to increase the production yield of commercially important metabolites[24]. Lastly, abstraction hierarchy employs the idea of the “black box” to navigate through biological complexity, by creating abstraction levels or hierarchies between parts, devices and systems and allowing only certain information to pass through each level. This requires several engineers to work on different levels of abstraction, but ensures that in each level, the information required is adapted to the biological question at hand. For example, an engineer designing a system composed of several biological devices, would not benefit from knowing all the DNA base pairs that create the biological parts used. This information is useful in the lower levels, where parts need to be characterized, and devices need to be built and tested.

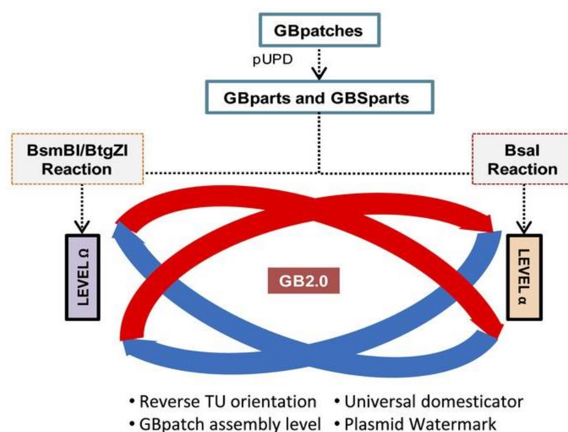
Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

Fungal Synthetic Biology

Fungal synthetic biology is mainly focused on metabolic engineering and heterologous expression of valuable metabolites or proteins[25–28], as alternative platforms to yeast and bacteria. An advantage of endophytic fungi is particular, is that because of their intricate interaction with plants, they usually develop unique specialized metabolite gene clusters through co-evolution and are native overproducers of certain specialized metabolite classes. An example of an effort to provide tools for endophyte engineering is TrichoGate, which includes unitary parts and vectors for functional analysis of *Trichoderma* species[29]. Nonetheless, the application of synthetic biology to study plant-fungal interactions from the perspective of the endophyte is under-represented. Development of robust and modular toolkits able to dissect mechanisms and ultimately modify the interactions between endophytes and their hosts, could point toward optimized agriculture and increase crop productivity[30]. Crop engineering is already advanced, but microbiome engineering could also provide a complementary solution for sustainable agriculture[31].

Type IIS modular assembly – FungalBraid

Standardization is of paramount importance to create robust, accessible and future-proof tools to engineer fungi. FungalBraid[32,33]is a robust Type IIS part toolkit for fungal biology, including coding sequences, peptides and regulatory elements for heterologous expression in filamentous fungi. Based on the Phytobricks standard[23], it is cross-compatible with vast collections of universal basic parts, such as fluorescent reporters, signal peptides and CRISPR components. This system comprises so-called alpha and omega vectors. Both alpha and omega vectors have two positions, 1 (one) and 2 (two). To achieve multi-gene constructs, alpha vectors are combined two at a time to create an omega vector. For the assembly to work, the alpha vectors need to have different positions, i.e. one alpha1 and one alpha 2 vector. The assembly can be performed either in an omega1 or an omega2 vector, depending on the desired position in the final multi-gene construct. For example, for a four transcriptional-unit construct, the cassette in Position 3 should be cloned in alpha1, and then in omega2 when combined with an appropriate alpha2 (final position 4) vector. The omega2 vector is combined with the desired omega 1 (position 1 + 2) vector, to yield a final alpha vector, that contains the desired cassette in Position 3.



Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

Figure 1. Depiction of GoldenBraid/FungalBraid Type IIS assembly method. Level 0 (GBparts) parts are domesticated in pUPD vectors. Higher level transcriptional units are assembled in alpha or omega vectors, depending on the restriction enzyme used. Combination of two alpha vectors can lead to an omega vector, while combination of two omega vectors can lead to the assembly of an alpha vector. Adapted from[34].

CRISPR functional genetics

CRISPR (Clustered Regularly Interspaced Palindromic Repeats) is a naturally occurring microbial adaptive immune system, found in many prokaryotes. In this system, CRISPR-associated (Cas) nucleases detect and cleave incoming DNA (e.g. from phages) to prevent its integration in the genome, and store parts of the sequence in CRISPR repeats, thus obtaining adaptive immunity. The most studied CRISPR system is that of *Streptococcus pyogenes*, a human pathogen. The biotechnological prowess of CRISPR stems from its ability to perform RNA-programmed cleavage from a 20-bp guide sequence, allowing it to target virtually any genomic location of choice, by changing the short RNA guide.[35]

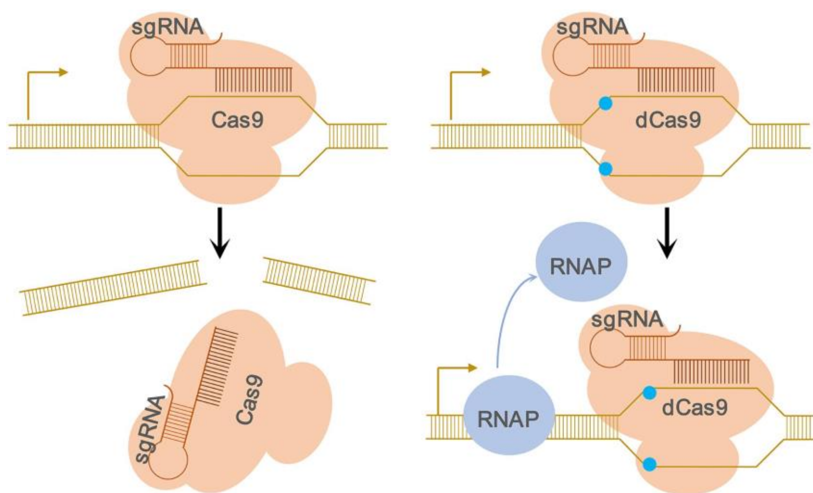


Figure 2. Simplistic depiction of CRISPR and CRISPRi systems. Both systems require and bind to a synthetic guide RNA (sgRNA), that targets a genomic region. The sgRNA provides the specificity required for the system to work only in the targeted region. When Cas9 is present (CRISPR), binding of the Cas9-sgRNA complex leads to a Double Stranded Break (DSB). When Cas9 is catalytically inactive (dCas9, mutations D10A, H840A), the dCas9-sgRNA complex still binds to DNA, but is unable to perform cleavage, thus inhibiting RNA polymerase from initiating transcription, causing downregulation of the targeted gene. Adapted from[36].

Gene knockouts with CRISPR/Cas9

For a CRISPR/Cas9 system to work, two heterologous components need to be expressed; the *Cas9* RNA-programmable nuclease, and the guide RNA (gRNA), which is a synthetic analog of the tracr-crRNA duplex. Upon expression of the two components in the cell, gRNA binds to Cas9 and renders the complex specific to the 20 bp spacer sequence in the gRNA, which usually targets the coding sequence of a gene. For successful binding to the target DNA, a Protospacer Adjacent Motif (PAM) sequence is required to flank

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

the spacer sequence and is recognized by the nuclease. This sequence depends on the Cas enzyme used. In the case of Cas9, the PAM site is 5' NGG 3' and flanks the 3' end of the target site[35]. Binding in the target DNA leads to a Double Stranded Break (DSB), caused by the two nuclease domains. Each domain targets a different DNA strand. For efficient mutagenesis to be observed, the system relies on endogenous DNA repair pathways, non-homologous end joining (NHEJ) and Homology-Directed Repair (HDR). NHEJ is an error-prone process, which uses Ku heterodimers at the end of the DSB to recruit other repair proteins and ligate the two ends together. During this process, small insertions and deletions (indels) are produced, often leading to loss-of-function mutations from early stop codons or frameshifting. During HDR, homology proteins such as Rad51 are recruited to the DSB and facilitate homology repair based on an exogenous repair template. The template can have homology arms ranging from 35-bp to 1.5kb, based on the system, organism and conditions performed. Usually, a marker is introduced in the DNA repair template, such as antibiotic resistance or fluorescence, to facilitate selection of mutated clones. gRNAs are designed within the CDS of the gene and preferentially on or close to the first exon, to increase chances of complete loss of function[37].Moreover, the choice of gRNA is important, as its base composition and starting nucleotide influence gene expression from Pol III promoters, and therefore affect overall system efficiency. Most Pol III promoters used for gRNA expression prefer a G at the start of the gRNA[38].

Gene regulation with catalytically inactive Cas9

Cas9-mediated gene regulation can be achieved by mutating the endonuclease active sites, leading to a catalytically inactive or dead Cas9 (dCas9). dCas9 retains the ability to bind to DNA but cannot produce a DSB. This system can be used for gene silencing or overexpression. To perform gene silencing or CRISPRi (interference), dCas9 alone or fused with transcriptional repressors (e.g. KRAB) are targeted close the gene's Transcriptional Starting Site (TSS) to block RNA polymerase access and repress expression. Similarly, in CRISPRa (activator), dCas9 fused with transcriptional activators, such as VP16, is used to activate or overexpress genes by recruiting transcriptional elements and increase expression. gRNA spacer design is similar to CRISPR knockout systems in terms of length and composition, but the target region is near the TSS instead of the exons. Generally, -/+ 50 bp from TSS produces best results in terms of repression/activation efficiency in mammalian systems[39]. A limitation of this system is that TSS prediction and elucidation is required for an efficient gene regulation strategy to be developed. Overall, CRISPRi/a are useful modular tools, as only gRNA spacer needs to change each time, with CRISPRi being handy when dealing with lethal genes in functional genetics, as it is an orthogonal silencing system, compared to e.g. RNAi.

CRISPR in fungal systems

There are two main approaches used to produce CRISPR mutations in fungi; RNP-based (*in vitro*) and plasmid-based (*in vivo*) mutagenesis[40,41]. In RNP-based mutagenesis, Cas9 protein is mixed or assembled with gRNAs (crRNA and tracrRNA) and delivered through protoplast transformation (Polyethylene glycol gradient transformation). This circumvents the need to transform the organism with foreign DNA and express all CRISPR components to perform mutagenesis, which may not be feasible depending on the organism. RNP-based methods often rely on the Homology Directed Repair (HDR) mechanism, where a Donor DNA containing flanking regions and an antibiotic resistance cassette is delivered together with the RNP complexes during transformation to increase efficiency and lower screening costs. However, this method usually yields low transformation efficiency, while also hurdling with batch-to-batch variation for protoplast preparation and high cost to purchase protein and synthesize gRNAs. Cost can be significantly lowered if gRNAs are synthesized in-house using *in vitro* transcription, purification, and validation, but

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

Cas9 expression in bacterial or other expression systems and further purification to be fit for transformation requires relevant expertise and a long optimization process.

In plasmid-based (*in* vivo) mutagenesis, *Agrobacterium tumefaciens* strains are employed to stably transfer and integrate T-DNA in a random location in the genome in one or more copies. This system is based on plasmid assembly of the desired transcriptional units (TUs) in a T-DNA shuttle vector, where the TUs are contained with the LB and RB sites that will be integrated in the genome. Such a system allows the flexibility to include additional markers and transcriptional units and fine-tune expression by changing regulatory elements. Also, constitutive expression of CRISPR components ensures high mutagenesis efficiency after transformation, which can circumvent the potential high fidelity of DNA DSB repair with NHEJ for some fungi[42].The size of T-DNA and number of TUs is determined by the cloning process of assembling all parts together, which largely depends on the assembly standard that is followed. The downside of the system is that heterologous expression of several components (antibiotic gene, Cas protein, gRNA) is required for the system to function, while Cas and gRNA expression are continuous, which may lead to off-target effects.

These two approaches have been extensively used in several filamentous fungi model organisms, such as *Penicillium chrysogenum*[43], *Trichoderma reesi*[44], *Neurospora crassa*[45], several *Aspergillus* species[46,47]. Previous work in Fusaria has showed successful gene editing using RNP delivery in fungal protoplasts[48,49], but the reliance on high-purity homemade proteins or commercial solutions, as well as the high batch-to-batch variability of protoplast preparation reduces the chances of result reproducibility and increases cost. Moreover, high repair fidelity after DNA Double Strand breaks via the NHEJ pathway, renders mutagenesis using transient RNP delivery inefficient, as mutations often occur as a function of nuclei division[42].

Aims of the thesis

Based on this knowledge, our goal is to create standardized tools for *Fusarium solani* strain K and Fusaria in general, using a low-cost and re-usable plasmid expression system. To this end, we developed a genetic part toolkit for FsK functional genetics, which includes components for CRISPR knockouts and CRISPRi knockdowns, in case of gene lethality.

First aim of the thesis is the development of a genetic part toolkit for FsK, containing both fungal-specific and universal basic parts for gene expression in fungi. This involved several domestication and cloning steps to onboard all desired parts.

Second aim of the thesis in the assembly and transformation of *Agrobacterium tumefaciens* shuttle binary vectors based on the CRISPR/Cas9 system to knockout and knockdown target genes in FsK genome, by using RNAi machinery genes as a proof-of-concept target.

Third aim includes the validation of the CRISPR system, by assessing mutagenesis events, and altered gene expression because of indel mutations or dCas9-based silencing.

Materials and Methods

gRNA design

gRNA design was performed using EuPAaGDT[50], a web platform that allows the usage of custom genomes as databases for high-throughput gRNA prediction and provides several scoring parameters and warnings.

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

gRNA spacers that contained less than four consecutive thymine bases (TTTT) and started with G were chosen, or a G was added to increase transcriptional efficiency[38,51]. For CRISPRi experiments, four gRNAs up to 300 bp upstream of the gene's predicted CDS were chosen. For CRISPR experiments, four to six gRNAs targeting the two first exons were chosen for each gene. Assembly of gRNA cassettes was performed using the FungalBraid system, by ordering the gRNA strands as oligonucleotides with 4 bp overhangs that imitate BsaI digestion. 5' overhang is TTTC, 3' overhang is GTTT.

Assembly of CRISPR transformation vectors

Assembly of transformation vectors was carried out using the FungalBraid Type IIS standard and one-pot digestion ligation (dig-lig) reactions[32]. In brief, unitary parts were domesticated by PCR mutagenesis to remove incompatible cutting sites (BsaI, BsmBI) and cloned in pUPD2. Single transcriptional units were cloned in appropriate alpha vectors using BsaI-HFv2 (NEB), which were then assembled in pairs in omega vectors using Esp3I (NEB). The final vectors (see Results section) were assembled back in alpha vectors, using BsaI-HFv2 (NEB). Dig-lig reactions were optimized for large DNA assemblies, by tuning the time of digestion and ligation (5 minutes for each) and cycle number (50-60), to increase efficiency. Dig-lig reactions were transformed in chemically competent *E. coli* DH5-alpha cells and plated in LB Agar containing appropriate antibiotic and X-gal, to enable blue-white screening. Two to four white colonies from each plate were picked for each vector for growth in 5mL LB liquid cultures. Plasmid isolation was performed using Nucleospin Plasmid miniprep kit (MN, Germany), according to manufacturer's instructions with slight modifications. Correct assembly was validated by restriction enzyme digestion, where GB identifiers enzymes (alpha1: EcoRI, alpha2: HindIII, omega1: BamHI, omega2: EcoRV) were mainly used. Digestion reactions were visualized in 1% agarose gels by Agarose Gel Electrophoresis. Validated clones were Sanger sequenced with primers on the LB, RB sites as well as promoter regions of each transcriptional unit.

***Agrobacterium* transformation of CRISPR vectors**

Validated CRISPR vectors were transformed into *Agrobacterium tumefaciens* AGL1 cells, using electroporation. Competent cells were prepared as follows: AGL1 cells were streaked in LB Agar containing rifampicin 20 ug / mL from glycerol stock and incubated at 28oC for 2 days. Single colonies were picked to inoculate 5mL LB pre-cultures with rifampicin 20 ug / mL and incubated at 28oC/160rpm for 2 days. Cell density was determined using OD_{600} spectrophotometer measurement. Flasks with 100 mL LB were inoculated from the pre-cultures with final OD_{600} = 0.02, and incubated 28oC/160rpm for 16h, or until OD_{600} = 0.5. Cells were centrifuged at 4000 x g for 10 min and washed with ice-cold deionized water two times, followed by washing with ice-cold 10% glycerol solution. 100 mL of LB culture yields 500-1000 uL of competent cells, which can then be aliquoted and used for electroporation. 10-20 ng of plasmid DNA was mixed with 40 uL of competent cells and electroporated with one pulse at 2.2 kV. Cells were immediately recovered with 960 uL LB and incubating at 28oC/160rpm for 2-3 h. 100-250 uL of the transformation reaction was plated in LB Agar with appropriate antibiotics and incubated at 28oC for 2 days. Single colonies from the plates were used for FsK transformation and snap-freezed with liquid nitrogen for future use.

Agrobacterium-mediated FsK transformation

FsK transformation was performed as described in Sesma and Osbourn[52] and Skiada et al[10]. Antibiotic selection with G418 was carried out using 40 ug/mL in PDA plates, after performing a kill curve in FsK WT

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

(data not shown). Transformed colonies were shown within one week of FsK-Agrobacterium co-cultivation and ~30-50 per construct were transferred in PDA with antibiotic. Antibiotic resistance was validated by transferring the colonies 2-3 times. GFP fluorescence was assessed using a blue-light lamp, before proceeding to DNA extraction and single conidia isolation for 5-10 colonies per construct. PCR for the antibiotic resistance gene (*nptII* or *hph*) was performed to verify T-DNA insertion. Four conidia per colony were assessed for growth and antibiotic resistance.

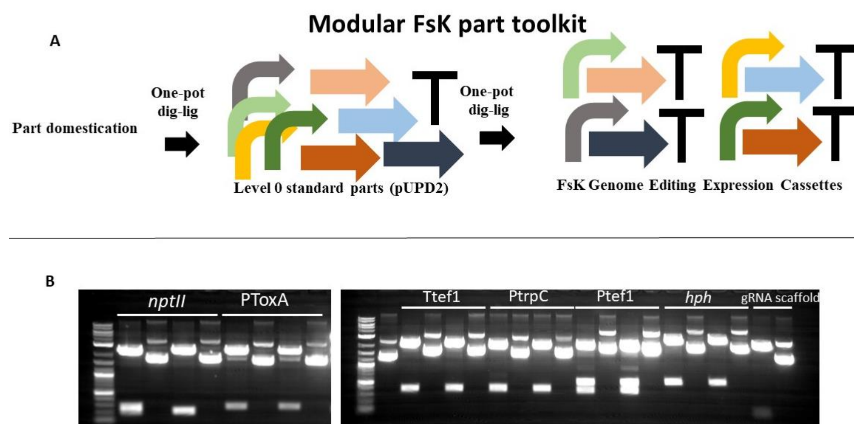
Transformant molecular analysis

Validated FsK clones were further analyzed for mutagenesis in target sites and gene expression levels of target genes. For CRISPR experiments, DNA was extracted using the CTAB[53] method and PCR was performed targeting the regions surrounding the gRNA spacers used for each gene. PCR products were purified (MN Nucleospin Gel) and sequenced using Sanger, to assess indel presence. For CRISPRi experiments, RNA was extracted using the Tritidy RNA extraction buffer and precipitation. DNase treatment was performed, and samples were quantified using Qubit RNA BR. Gene expression analysis was carried out using One-step RT-qPCR with the Luna Mix (NEB) per manufacturer's instructions, for all target genes, using *tef1a* as a reference gene relative expression analysis. ANOVA analysis was performed to determine the variance between groups, followed by a paired T-test for each combination of groups to identify significantly decreased mRNA levels between conditions using Excel Data Analysis Toolpak.

Results

FsK Type IIS basic part toolkit

First aim of the thesis was the development of a genetic part toolkit for FsK. This should contain all required elements to perform fungal transformation, thus *Agrobacterium* shuttle vectors and DNA elements that allow expression of transcriptional units (TUs). We chose the Type IIS assembly method FungalBraid as the heart of the toolkit, as it offers interchangeability and modularity of parts between different expression units, vectors and organisms and has been previously used in filamentous fungi (Figure 3).



Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

Advanced Experimental and Computational Biosciences

DEPARTMENT OF Biochemistry & Biotechnology UNIVERSITY OF THESSALY

Figure 3. FsK Type IIS modular part toolkit. A) Workflow of part toolkit for FsK. Parts are domesticated using PCR mutagenesis and cloned in pUPD2 entry vector. Using iterative one-pot dig-lig reactions, modular vectors containing 1, 2 or more transcriptional units are created, that can be expressed in FsK. B) Indicative electropherogram of several basic parts from the toolkit cloned in pUPD2. Two clones per construct were analyzed, and both the cut (first) and uncut (second) plasmid were used in the electropherogram. 100% cloning efficiency was observed.

We domesticated all basic parts using the FungalBraid[32] assembly (Figure 3B). We removed illegal cutting sites for BsaI and BsmBI from the part sequence using PCR mutagenesis and subsequent cloning in the pUPD2 entry vector. A list of all parts created during the thesis can be found in Table 1. The parts domesticated were either available in the lab, tested before in other FsK expression vectors or were found in the literature.

Part name	Description	Part Type	4-bp overhangs
PtrpC	*trpC* promoter from *A. nidulans*	Promoter + 5 UTR	GGAG - AATG
Ptef1	*tef1a* promoter from *A. nidulans*	Promoter + 5 UTR	GGAG – AATG
PToxA	*ToxA* promoter from *P. tritici-repentis*	Promoter + 5 UTR	GGAG – AATG
PtrpC-5f	*trpC* promoter from *A. nidulans*	Promoter + 5 UTR (f)	GGAG - CCAT
Ptef1-5f	*tef1a* promoter from *A. nidulans*	Promoter + 5 UTR (f)	GGAG – CCAT
PToxA-5f	*ToxA* promoter from *P. tritici-repentis*	Promoter + 5 UTR (f)	GGAG – CCAT
FsKU6p	U6 snRNA promoter from FsK	Promoter (CUSTOM)	GGAG – TTTC
H2B_NLS	Nuclear Localization Signal (NLS) from Histone 2B from FsK	Signal Peptide (NTAG)	CCAT - AATG
***sfGFP* NTAG**	*Superfolder GFP*, ready for fusion at 5' end of CDS	NTAG	CCAT - AATG
***mCherry* NTAG**	*mCherry,* ready for fusion at 5' end of CDS	NTAG	CCAT - AATG
***mCerulean3* NTAG**	*mCerulean3* (blue), ready for fusion at 5' end of CDS	NTAG	CCAT - AATG
hph	*Hygromycin phosphotransferase*, confers resistance to hygromycin	CDS	AATG - GCTT
nptII	*Neomycin phosphotransferase*, confers resistance to G418, gentamycin, kanamycin, neomycin	CDS	AATG – GCTT
AtCas9	*SpCas9* codon-optimized for dicot plants, SV40 NLS at C-terminal	CDS	AATG – GCTT
dAtCas9	Dead SpCas9 codon-optimized for dicot plants, D10A, H840A point mutations, SV40 NLS at C-terminal	CDS	AATG – GCTT
hCas9	*SpCas9* codon-optimized for humans, SV40 NLS at C-terminal	CDS	AATG – GCTT
AnCas9	*SpCas9* codon-optimized for *Aspergillus*, SV40 NLS at C-terminal	CDS	AATG – GCTT
flp	Recombinase for FLP-FRT marker recycling system	CDS	AATG – GCTT

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

AcrIIA4	Anti-CRISPR protein IIA4, inactivates SpCas9, inhibits DNA binding	CDS	AATG – GCTT
Csy4	Sequence-specific RNA endonuclease, used for CRISPR multiplexing	CDS	AATG – GCTT
tRFP	*Turbo RFP*, emits red fluorescence	CDS	AATG – GCTT
sfGFP	*Superfolder* GFP, emits green fluorescence	CDS	AATG – GCTT
***sfGFP* CTAG**	*Superfolder* GFP, ready for fusion at 3' end of CDS	CTAG	TTCG-GCTT
***mCherry* CTAG**	*mCherry*, ready for fusion at 3' end of CDS	CTAG	TTCG-GCTT
***mCerulean3* CTAG**	*mCerulean3* (blue), ready for fusion at 3' end of CDS	CTAG	TTCG-GCTT
gRNA scaffold	gRNA scaffold that connects to gRNA spacer, binds to Cas9, includes terminator	sgRNA	GTTT - CGCT
Ttef1	*tef1a* terminator from *A. nidulans*	Terminator	GCTT - CGCT

Table 1. List of domesticated unitary parts created in this thesis. Parts are cloned in pUPD2 (Chloramphenicol resistance) using BsmBI and are flanked by BsaI sites, making them compatible with the GoldenBraid, FungalBraid and Golden Gate assembly methods.

As seen in Table 1, currently the toolkit contains several promoters, one terminator, one native NLS, several coding sequences encoding useful functions, and the basic components required to build single-plex, multiplex and anti CRISPR-based CRISPR circuits, as well as three codon-optimized versions of SpCas9 and a catalytically inactive (dead) SpCas9. This collection of parts, along with the binary vector set from GoldenBraid v3.0 (GB3.0)[33], ensures that a variety of transcriptional units can be created, with different regulatory elements to fine-tune gene expression. Moreover, the constitutive Pol II promoters were converted in two Part types, to both an AATG and CCAT as the 3' overhang. This allows connection of the promoter with any N-terminal tag, such as the *H2B* NLS, a MYC-tag, SV40 NLS, 6X-his tag, or fluorescent proteins. Similarly, tags and fusion proteins can be added in C-terminus of the CDS, by using the CTAG part type, with the sole requirement that re-modification of the desired CDS, to add compatible 3' overhang bases, in this case TTCG and remove the stop codon to be satisfied.

Hence, the DNA parts included in the toolkit allow for robust assembly of several transcriptional units with different functionalities.

Assembly and *Agrobacterium*-mediated FsK transformation of CRISPR vectors for RNAi genes

We sought to test the functionality of the FsK part toolkit by targeting the core RNAi machinery of the fungus. As mentioned in the introduction, FsK has two predicted Argonaute (*AGO1*/*2*) and two predicted Dicer-like (*DCL1*/*2*) proteins, which supposedly compose its RNAi pathway. We sought to produce silencing by CRISPRi on *AGO1* and *AGO2*, and loss-of-function mutations by CRISPR on all four genes, by designing gRNA spacers close to the start codon of each gene, as seen in Figure 4A. Abolishing functionality of DCL and AGO proteins could lead to complete disruption of the RNAi pathway, thus enabling the

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

elucidation of small RNA molecules in plant-microbe interactions (Figure 4B). A full list of gRNA spacers used to target the FsK genes, as well as their relative position to the start codon, can be found in Table 2.

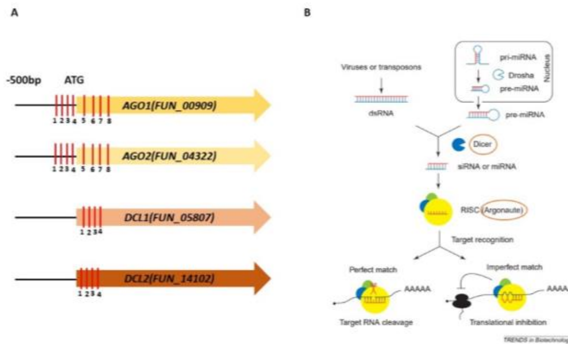


Figure 4. Application of CRISPR system in FsK RNAi genes. A) Graphical representation of target genes and gRNA spacers for the experiment. For *AGO1* and *AGO2*, 8 gRNA spacers were designed, of which four were used in CRISPRi vectors, and four in CRISPR vectors. For *DCL1* and *DCL2*, four spacers used in CRISPR vectors were included. B) Simplistic RNAi pathway. Targeted genes (Argonaute and Dicer proteins) are circled in orange. Dicer-like knockouts could lead to disruption of siRNA/miRNA production, while *AGO* knockouts should abolish target recognition. Adapted from[54] doi.org/10.1016/j.tibtech.2006.02.006

Commented [P1]: Ayto mallon einai apo mammalian-yparxei kati for fungi? A more suitable graphic?

Commented [NN2R1]: Den brika kati kalytero, g ayto afisa ayto. Found this, but RNAi system is different in FsK https://sci-hub.se/https://doi.org/10.3852/15-246

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

Gene Target	gRNA name	Position (relative to ATG)	Cloned in vector
AGO1(FUN_00909)	909.1	-65	CRISPRi
AGO1(FUN_00909)	909.2	-41	CRISPRi
AGO1(FUN_00909)	909.3	-15	CRISPRi
AGO1(FUN_00909)	909.4	-9	CRISPRi
AGO1(FUN_00909)	909.5	+75	CRISPR
AGO1(FUN_00909)	909.6	+240	CRISPR
AGO1(FUN_00909)	909.7	+290	CRISPR
AGO1(FUN_00909)	909.8	+350	CRISPR
AGO2(FUN_04322)	4322.1	-153	CRISPRi
AGO2(FUN_04322)	4322.2	-98	CRISPRi
AGO2(FUN_04322)	4322.3	-41	CRISPRi
AGO2(FUN_04322)	4322.4	-19	CRISPRi
AGO2(FUN_04322)	4322.5	+21	CRISPR
AGO2(FUN_04322)	4322.6	+234	CRISPR
AGO2(FUN_04322)	4322.7	+276	CRISPR
AGO2(FUN_04322)	4322.8	+358	CRISPR
DCL1(FUN_05807)	5807.1	+201	CRISPR
DCL1(FUN_05807)	5807.2	+299	CRISPR
DCL1(FUN_05807)	5807.3	+567	CRISPR
DCL1(FUN_05807)	5807.4	+693	CRISPR
DCL2(FUN_14102)	14102.1	+107	CRISPR
DCL2(FUN_14102)	14102.2	+221	CRISPR
DCL2(FUN_14102)	14102.3	+535	CRISPR
DCL2(FUN_14102)	14102.4	+619	CRISPR

Table 2. gRNA spacers designed and cloned during the thesis.

In order to produce the desired silencing and mutations, binary vectors containing all required components need to be assembled. Since we achieve construct delivery using *Agrobacterium tumefaciens* transformation, we need to create stable transgenic lines, and thus require marker genes, as well as the CRISPR components. The transcriptional units used in both CRISPR and CRISPRi vectors, albeit in different order because of cloning inefficiencies (see Discussion), were four. An antibiotic resistance cassette, encoding the *nptII* gene which confers resistance to G418, was driven by the *trpC* promoter. A Cas9 cassette, encoding the *A. thaliana* codon-optimized *SpCas9* or *dSpCas9* (for CRISPRi), was driven by the *tef1a* promoter. A reporter cassette, encoding for the fluorescent protein *superfolder GFP*, was driven by the *trpC* promoter. Finally, the gRNA cassette, encoding the gRNA spacer and the gRNA scaffold, was driven by endogenous RNA polIII FsK U6 promoter. The final layout of each system can be seen in Figure 5B.

Initially, the design included the strong *ToxA* promoter to drive the reporter cassette, and the fungal codon-optimized *Cas9* (Table 1). However, the use of the two parts was incompatible with the subsequent cloning of two or more transcriptional units. Switching the parts with *trpC* and *AtCas9*, respectively, enabled us to efficiently clone all four transcriptional units together. We speculate that the DNA sequence of those parts in combination with the adjacent parts, surfaced issues that may be

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

attributed to *E. coli* toxicity. Another possible technical issue may arise from the repeated use of the *tef1a* terminator, which was the only one available for domestication.

To create the final vectors, several steps have to be taken. First, all one transcriptional-unit vectors were assembled in alpha vectors, according to their position in the final layout. For example, the antibiotic resistance cassette was cloned in alpha1, and then in omega1 along with either the Cas9 cassette (CRISPR) or the gRNA cassette (CRISPRi). All intermediate vectors were constructed once, except for the vectors containing the gRNA spacers, which need to be constructed each time to target the specific gene. A list of all such vectors created can be found in Table 3, while electropherograms for intermediate and final vectors can be seen in Figure 6.

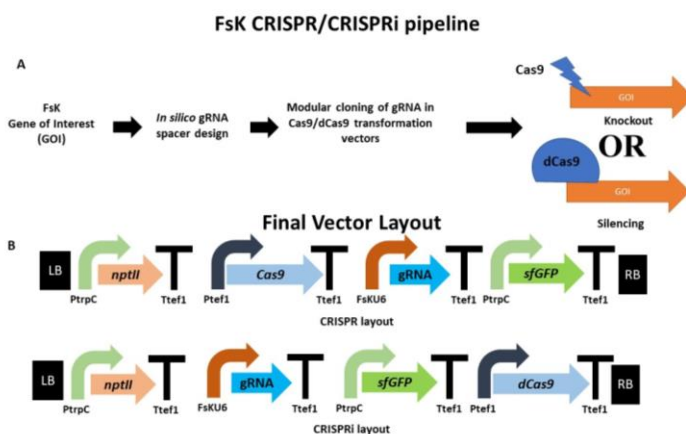


Figure 5. A) Proposed pipeline to perform CRISPR or CRISPRi on FsK. First, genes of interest (GOI) are identified in the genome and gRNA spacers are designed that specifically target either the CDS or 5' UTR region, depending on whether knockout (CRISPR) or silencing (CRISPRi) are of interest. gRNAs are ordered as ssDNA oligonucleuotides and are incorporated in modular vectors using the FungalBraid Type IIS assembly method, to create the final vector. B) Final Vector Layout for FsK. The FungalBraid vectors are binary vectors and thus carry LB and RB sites to perform T-DNA integration through *Agrobacterium* transformation. The multi-gene construct layout comprises the antibiotic resistance cassette, the Cas9 (or dCas9) cassette, the gRNA cassette, and the sfGFP (reporter) cassette.

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

Vector Name	Cloned in	Antibiotic Resistance
Alpha1-trpC-*nptII*-Ttef1	Omega1-nptII-AtCas9, Omega1-nptII-gRNA	Kanamycin
Alpha2-Ptef1-*AtCas9*-Ttef1	Omega1-nptII-AtCas9	Kanamycin
Alpha2-FsKU6-gRNA-scaffold	Omega1-nptII-gRNA	Kanamycin
Alpha1-FsKU6-gRNA-scaffold	Omega2-gRNA-sfGFP	Kanamycin
Alpha2-trpC-*sfGFP*-Ttef1	Omega2-gRNA-sfGFP	Kanamycin
Alpha1-trpC-*sfGFP*-Ttef1	Omega2-sfGFP-dAtCas9	Kanamycin
Alpha2-Ptef1-*dAtCas9*-Ttef1	Omega2-sfGFP-dAtCas9	Kanamycin
Omega1-*nptII-AtCas9*	Alpha2-nptII-AtCas9-gRNA-sfGFP	Spectinomycin
Omega2-gRNA-*sfGFP*	Alpha2-nptII-AtCas9-gRNA-sfGFP	Spectinomycin
Omega1-*nptII*-gRNA	Alpha2-nptII-gRNA-sfGFP-dAtCas9	Spectinomycin
Omega2-*sfGFP-dAtCas9*	Alpha2-nptII-gRNA-sfGFP-dAtCas9	Spectinomycin
Alpha2-nptII-AtCas9-gRNA-sfGFP	-	Kanamycin
Alpha2-nptII-gRNA-sfGFP-dAtCas9	-	Kanamycin

Table 3. List of intermediate alpha/omega and final vectors required for assembly of CRISPR and CRISPRi multi-gene constructs for this thesis.

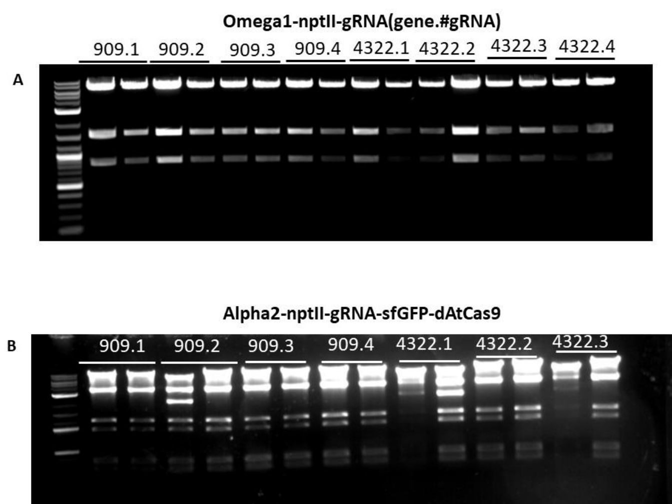


Figure 6. Electropherograms of successful intermediate and final CRISPR vectors. A) Indicative electropherogram of intermediate omega1 vector, containing the antibiotic resistance cassette and the gRNA cassette. Cut with BamHI, 100% assembly efficiency. B) Indicative electropherogram of final CRISPRi vector. 909.2 #1, 4322.1 #1 and 4322.3 #1 are incorrect. Cut with EcoRV.

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

Once all the pre-constructed vectors were available, the assembly of the final vectors from gRNA design to final vector validation took around three weeks, as the gRNAs need to pass through the alpha and omega intermediate vectors, before assembly in the final alpha vectors. In the future, this pipeline will be improved by creating CRISPR acceptors, which will be assembled with gRNA spacers in one step.

One construct per gRNA spacer in the validated final vectors was inserted into *Agrobacterium tumefaciens* AGL1 cells using electroporation. Transformed AGL1 strains were co-cultivated with FsK conidia.The process of FsK transformation, requires approximately two weeks to perform. Transformed FsK clones were identified by antibiotic resistance. The validation of clones requires both phenotypic and molecular evaluation, by reporter and antibiotic screening and T-DNA PCR, respectively. A graphical depiction of the workflow can be found in Figure 7. Data about the efficiency of the FsK transformation workflow for the CRISPR vectors is shown in Table 4, for constructs 909.6 and 4322.6. This table will be completed, as most constructs are still in the validation phase. Notably, all transformed clones show retarded growth, as recorded by radial hyphal growth on rich media. This may be either due to RNAi gene disruption or because of the burden of the four cassettes being expressed. The protein size of Cas9, i.e. 1200aa may also be restrictive for the fungal growth. The effect on the fungal vegetative and reproductive physiology as well as the possible underlying reasons of it need to be further investigated.

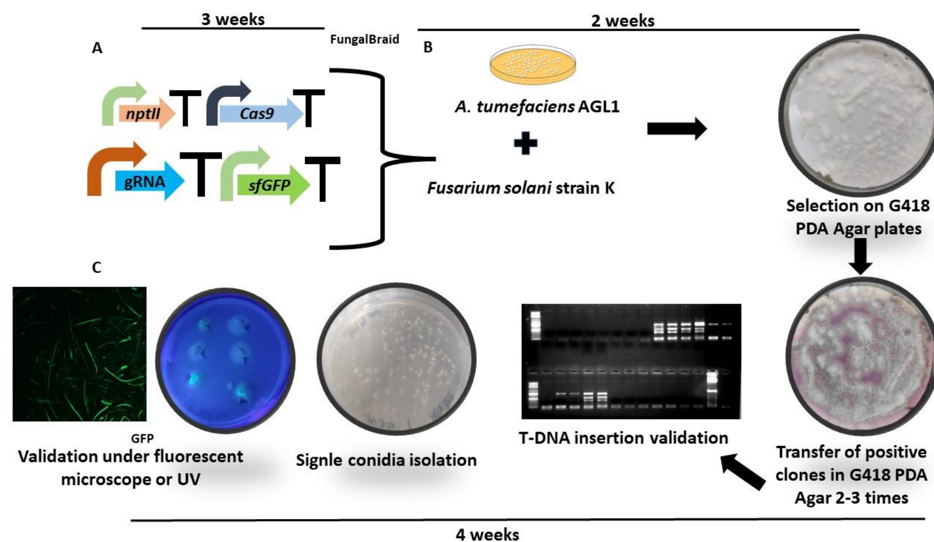


Figure 7. FsK transformation and clone validation workflow. A) Assembly of final CRISPR vector. Pre-validated transcriptional units are combined with the desired gRNA spacer to create final CRISPR vectors, using FungalBraid. After validation, final vectors are transformed in *Agrobacterium tumefaciens* AGL1. B) FsK Transformation. *A. tumefaciens* AGL1 transformed with CRISPR vectors are co-cultivated with FsK conidia and plated on Cellulose Nitrate membranes until colonies are formed. Antibiotic selection is applied to identify transformed clones. C) Validation of transformed clones. Transformed clones are analyzed for T-DNA insertion after DNA extraction and PCR, by amplifying a 1000 bp region of the T-DNA (antibiotic resistance cassette). Positive clones are cultured in 10 mL liquid cultures, conidia are harvested, and 50 conidia are spread in a PDA plate. Single conidia are re-cultured in PDA with

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

antibiotic. Through the whole process, screening under black light (excitation ~480nm) or fluorescent microscope is performed, to choose the clones that fluoresce.

Construct	# of clones picked	Survived in PDA G418	Transformed (Green)	Efficiency
909.6	36	34/36	22/34	64.7%
4322.6	36	33/36	23/33	69.7%

Table 4. Efficiency of FsK transformation based on fluorescent screening of resistant clones under black light lamp for sfGFP expression. Efficiency column is the percentage of the division of transformed clones by survived clones.

CRISPR system functionality on FsK RNAi machinery genes

The final step to validate the CRISPR system functionality in FsK is to observe phenotypic or molecular differences caused by the specific activity of the CRISPR system on the target genes. We first investigated whether transformed lines with different gRNA spacers can affect expression of the target gene.

Four validated constructs, targeting AGO1 and AGO2, were used to extract RNA and perform qPCR analysis. As shown in Figure 8, gRNA 909.6 is able to almost abolish AGO1 expression, although this result comes from one line and has to be repeated. For AGO2, mRNA levels are decreased significantly in three gRNAs tested when compared with the control FsK wild type strain. In both AGO1 and AGO2 levels, we observe that the Empty Vector condition, which contains a random non-specific gRNA instead of a target gRNA, shows relatively lowered mRNA levels compared with WT, albeit not statistically significant.

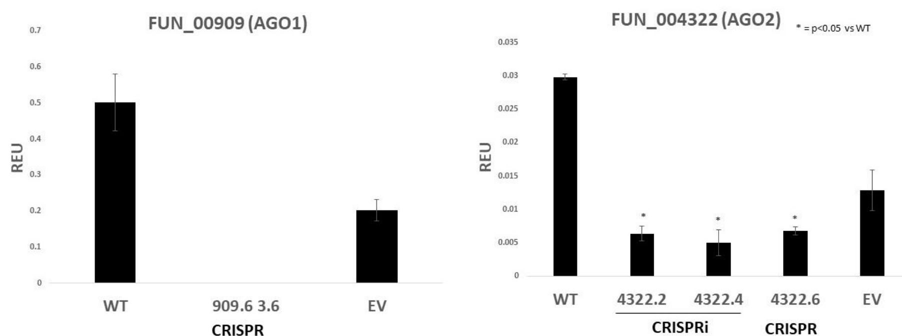


Figure 8. Gene expression analysis of AGO1 and AGO2 for CRISPR and CRISPRi. EV = Empty Vector. ANOVA was performed with n=3 samples for AGO2, followed by t-test of all conditions against the WT control to assess statistical significance. AGO1 mutant (909.6) could not be statistically checked, as only one transformant was available for RNA extraction.

Discussion

CRISPR system validation on FsK

The results presented above provide a first indication of CRISPR system functionality in FsK, both for silencing and mutagenesis purposes. To complement these preliminary results, more clones for each gRNA construct will be isolated, in order to have better sample size and be confident about the results. To complement the transcript levels results, for the CRISPR lines, PCR amplification of the gRNA regions will be performed, to identify potential indels caused the CRISPR system's activity.

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

The reduction in the expression levels of the target genes in the empty vector control condition compared with WT, though not significant, is of interest. This may be caused by the constitutive expression of the CRISPR system, which entails the 1200aa Cas9 protein and may be toxic or burdensome in the system. Since we observe the same reduction in both the two tested target genes, *AGO1* and *AGO2*, this result is most probably not specific. Fine-tuning of the expression with a weaker constitutive or inducible promoter could be a plausible solution to the problem. An alternative is to include more endogenous reference genes in the qPCR analysis. The tef1a gene, encoding for the translation elongation factor, may not be ideal when the system is loaded with a 17kb transgene. Housekeeping genes that will be used in future analyses include beta-tubulin, ubiquitin and actin, which have been used in the literature for *Fusarium* species[55,56].

To further characterize the system, CRISPR components expressed in FsK should be quantified. gRNA expression can be assessed using RT-qPCR, while Cas9 levels can be determined using both RT-qPCR and Western Blot. For the latter, transformation of FsK with an appropriate vector, assembled to include a Cas9 tagged with an antibody epitope, such as the 4X MYC-tag, will be needed.

It has been shown that the Histone 2B NLS from *F. oxysporum* localizes GFP in the nucleus more efficiently than the classical viral SV40 NLS, which is widely applied in mammalian systems[57]. We have verified that the H2B NLS sequence in FsK is the same as in *F. oxysporum* and FsK. Thus, to improve the efficiency of the mutagenesis and silencing system, the native Histone 2B NLS will be fused in the 5' end of the *Cas9* CDS.

CRISPR applications *Fusarium* and other filamentous fungi

Applications of CRISPR technology in fungi has been steadily expanding to more species and different systems[40]. Several groups mainly work with RNP-based CRISPR systems, supposedly because plasmid-based expression is inefficient due to large Cas9 protein size. However, such systems have been shown to work in all kingdoms of life, and it is likely that under-representation of plasmid-based expression systems for CRISPR exists due to the lack of tools, modular parts and transformation vectors that work in a broad range of organisms. The genetic toolkit developed in this thesis will hopefully be a step toward a more systematic and standardized approach to performing CRISPR experiments and heterologous expression in general in *Fusarium* species. However, a transition to other genera fungal systems would only require compatible regulatory genetic elements, such as promoters and terminators, and a feasible *A. tumefaciens* transformation protocol. In this way, rapid vector assembly, transformation and selection could pave the way for functional screening of many fungal isolates, especially endophytes, to enable larger comparative studies, e.g. between beneficial endophytes that confer a similar trait, such as abiotic stress tolerance. This could allow the identification of genes involved in induction of plant immunity, or genes that determine the lack of pathogenicity. FsK is an organism that provides many benefits to its host plant, while its beneficial role in other plant hosts is also established. Thus, FsK is an ideal platform for studying both plant-microbe interactions in the molecular level, as well as the relationship between phenotype and genotype, which will be enabled by this toolkit.

CRISPRi systems provide an easy platform for silencing genes of interest by only changing 20 base pairs for the target gRNA spacer. However, silencing efficiency in such systems is closely correlated with the proximity of the *dCas9* binding to the transcriptional starting site (TSS) of the target gene. While in most model organisms, the TSS sequences have been determined, either by long-range sequencing or CAP-seq[58], such information is not available for FsK's genome, or any close *Fusarium* genome. This makes

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

identifying a gRNA with high repressing efficiency much harder and requires screening more spacers upstream of the gene of interest, which increases labor and cost. However, this CRISPRi system was set up as an alternative to the CRISPR knockout system, and with the belief that TSS information will be generated by the lab or will be available by other groups in *Fusarium* species, so then it can be utilized properly.

Limitations in part assembly using FungalBraid

Shortcomings of the FungalBraid system mainly fall into the problems that one experiences with cloning (large) constructs in general: it is limited by biology. Some parts or part combinations likely contain DNA sequences that are toxic to *E. coli* and, thus, it is almost impossible to clone them. The more assemblies and combinations one performs, more such problems are probable to occur. Circumventing this requires flexibility in the number of available parts; ideally regulatory elements should be used once per vector, and "spare parts" that could replace a problematic promoter or terminator should be available. For example, we experienced problems when assembling higher level vectors (2-4 TUs) that contained the ToxA promoter for sfGFP or RFP expression. This promoter is taken from a toxin gene from *Pyrenophora tritici-repentis* and normally produces a toxin. Taking regulatory elements from a toxin production system is not ideal, but it is remarkable that the assembly problems did not occur at lower levels, i.e. cloning the promoter in entry vectors or assembling the fluorescent reporter (ToxA:*sfGFP*:Ttef1) transcriptional unit, which surely merits further investigation on TypeIIS assembly fidelity. Moreover, having enough parts to avoid multiple times the same one, might help in this direction. In our assembly problems, we were in possession of only one terminator, Ttef1, which terminated transcription in three out of four TUs. Since transcriptional terminators produce secondary DNA structures to effectively stop transcription, having different parts in different TUs could increase assembly efficiency. In this case, domestication of new parts and assembly of different cassettes will be required, although changing only the specific regulatory elements.

Furthermore, because we aimed to transform a multi-gene construct, the vector and T-DNA size also determine the assembly and transformation efficiency of the system. In this work, we used *E. coli* DH5-alpha for all assemblies, with the final constructs exceeding 17 kb in size. With exception to the obstacles described in the previous paragraph, we did not experience other issues regarding cloning. However, correct white colony selection for the final vectors was not always 100% (which was the case for 1-2 TU vectors), in our heavily optimized workflow. This could be a potential issue, as more colonies are required for screening, which increases cost and labor. We suggest two avenues to tackle this potential issue. First, we propose the use of other competent cell lines, like *E. coli* DH10-beta, which are optimized for larger plasmids/cosmids/BACs. The use of DH10-beta could increase transformation efficiency and ideally produce less recombination events during plasmid propagation. However, this cell line has *SpecR* resistance gene in its genome, which renders them incapable of selecting omega plasmids, which also carry the same *SpecR* gene. Despite that, lower-level assemblies could be performed in DH5-alpha, and only the final vectors, which are kanamycin-resistant, could be assembled in DH10-beta. Second, we propose the use of minimal transformation vectors, to decrease the burden in the cells during transformation, and increase T-DNA size that can be transformed. The current FungalBraid system vectors are derived from the gold-standard pCambia and pGreen vectors, which have worked wonderfully in many plant and fungal species. However, their backbone exceeds 7 kb in size, which limits the T-DNA size and also the assembly efficiency. Minimal T-DNA transformation vectors have been developed to tackle this problem, while retaining the Type IIS assembly standard of FungalBraid vectors[59]. The latter system also

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

has plasmids with different replications of origin for bacterial transformation, which renders feasible co-transformation attempts (e.g. CRISPR-gRNA dual vector system).

Future prospects for toolkit expansion

This version of our toolkit serves as a proof of concept for both CRISPR applications and modular cloning on FsK, which will render many projects feasible in the future. These could be centered around gene function to elucidate other biological questions beside the RNAi machinery genes presented here, such as colonization mechanisms, secondary metabolite production, miRNA signalling, cazyme-mediated cell wall penetration, and effector protein signaling, where knockout or knockdown lines can be easily produced for reverse genetics experiments. Assigning a physiological role for many fungal genes is a major contribution to an existing gap in fungal biology. It could accelerate significantly the targetter heterologous production of proteins, metabolites or signaling components towards novel identifying mechanisms that mediate the mode of action of endophytes, e.g. secondary metabolites, or hormones that can modulate plant responses.

We aim to expand the toolkit features in the future, with a focus on **multiplexing; other Cas enzymes; system inducibility; and marker recycling.**

With the current toolkit version, gRNA expression units can be stacked to target different genes, and create double or higher-level mutants. This, however, will increase the burden both during vector assembly as well as expression in the host, as each gRNA will be driven by a Pol III promoter, which will likely become inefficient after two or three gRNA units. An elegant addition to the toolkit would be a multiplexing system, where gRNAs are transcribed driven by one promoter (Pol II or Pol III) in a long mRNA strand and are then processed to release functional gRNAs by an endogenous or heterologous mechanism. Two such mechanisms have been shown to work efficiently in CRISPR systems; tRNA splicing and Csy4-mediated cleavage. In tRNA splicing, tRNA tertiary loops are included between gRNAs during the design phase. In Csy-4 mediated cleavage, a heterologous RNA processing enzyme found in *Pseudomonas aeruginosa* that specifically recognizes and binds to a 15-bp RNA sequence[60] is expressed in the host, . By introducing this specific sequence between gRNAs, gRNA arrays can be created, which are then processed to yield functional gRNAs. This system has been shown to work for up to 6 gRNAs per array (expression unit, with one promoter). For our project, we have designed unitary parts and transcriptional units for Csy4-mediated processing, as it is more orthogonal and does not rely on host machinery to process gRNAs. An immediate application of this multiplexing approach could be the creation of double or quadruple mutants of RNAi components, where either Argonaute or Dicer proteins, or all of them, would be knocked out to gauge gene essentiality and yield a truly RNAi null FsK line. Alternatively, multiplexing could be used to excise specific exons from the genes, to investigate their potential function and determine if the RNAi mechanism is conserved across kingdoms.

Furthermore, incorporation of different Cas enzymes, such as *Cas12*, *Cas13* and chimeric versions of *Cas9* such as base editors, could increase the flexibility and usability of the toolkit. Depending on the biological target, *Cas12* could be used to increase efficiency and availability of gRNAs, as it requires a different PAM sequence (TTTN), which allows targeting more A-T rich regions. *Cas13* targets RNA instead of DNA, so it could be used for RNA editing purposes, or for silencing of target genes, where the TSS is not known. Finally, chimeric versions of Cas9, such as base or primer editors, could allow very precise modifications

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

in genes or regulatory regions, for example to increase promoter strength in a native gene or produce an inactivating mutation in the catalytic site of an enzyme, to investigate its role and phenotype. All these will be made possible by domesticating the Cas genes in the Type IIS assembly standard and assembling TUs that contain the different Cas genes instead of *Cas9* or *dCas9*. An already available resource for this is the iGEM Distribution 2022, a collection of characterized unitary parts available to the community. Among other part collections, it contains a CRISPR-Cas collection of Type IIS – compatible parts, ranging from *Cas9*, *Cas12*, *Cas13*, to base and primer editors. This resource is available to the department and the parts are already cloned and verified as unitary parts (Level 0).

Feedback loops are essential in biology and having the ability to build such systems in an artificial and orthogonal manner could provide additional tools for genetic engineering and Synbio approaches. A useful extension based on CRISPR editing could be anti-CRISPR feedback regulation, where anti-CRISPR proteins (aCr) are expressed in an inducible manner (by heat, cold or chemical ligand) and inactive Cas9 enzymes, hindering their activity. This system could be more utilized in a dCas9-based system, where oscillations in temperature or chemical (e.g. doxycycline) concentration could lead to oscillations in gene expression of target genes. This would require domestication of inducible promoters or other elements, such as transcriptional regulators that are modulated by chemical ligands, such as lacI or TetR.

Finally, including a marker recycling system would promote biosafety and biosecurity, by excising the antibiotic resistance gene from the genome after transformation and validation is complete. This could be achieved through an inducible homologous recombination system, such as *flp*-FRT or *cre*-LoxP. For this approach, specific recombination sites are introduced in the T-DNA surrounding the antibiotic resistance cassette and the recombination cassette. Upon induction by temperature or a chemical ligand, depending on the induction system, recombinase is expressed to excise the fragments from the genome, removing both cassettes, creating a marker-free transformant. This technology could prove highly useful in potential downstream applications, where marker-free lines are required.

Conclusions

Plant-microbe interactions have played a pivotal role in agriculture. Harnessing this potential will enable humans to achieve sustainable agricultural practices and overcome the growing food and energy demands we face. For this to be realized, however, we need to the right molecular tools to answer the right questions. This thesis provides a toolkit for functional reverse genetics in fungi, using the CRISPR system. Our results, although preliminary, show that this toolkit is functional and robust, enabling rapid assembly and fungal transformation to produce mutations and gene regulation of fungal target genes and make us hopeful for what lies ahead.

References

1 Yan, L. *et al. Appl. Microbiol. Biotechnol.* 103, 3327–3340 (2019)

2 Khan, A. L. *et al. Crit. Rev. Biotechnol.* 35, 62–74 (2015)

3 Latz, M. A. C. *et al. Plant Ecol. Divers.* 11, 555–567 (2018)

4 Zhou, Y. *et al. Functional Ecology* 32, (2018)

5 Pappas, M. L. *et al. Front. Plant Sci.* 9, 1–17 (2018)

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

6 Garantonakis, N. *et al. Front. Ecol. Evol.* 6, 1–7 (2018)

7 Kavroulakis, N. *et al. Rhizosphere* 6, 77–85 (2018)

8 Kavroulakis, N. *et al. J. Exp. Bot.* 58, 3853–3864 (2007)

9 Karpouzas, D. G. *et al. Eur. J. Soil Biol.* 47, 400–408 (2011)

10 Skiada, V. *et al. Fungal Genet. Biol.* 127, 60–74 (2019)

11 Skiada, V. *et al. New Phytol.* 226, 1429–1444 (2020)

12 Gutbrod, M. J. *et al. Nat. Rev. Genet.* 21, 311–331 (2020)

13 Lax, C. *et al. Int. J. Mol. Sci.* 21, 1–22 (2020)

14 Carreras-Villaseñor, N. *et al. Mol. Microbiol.* 89, 96–112 (2013)

15 Ellis, T. *et al. Integr. Biol.* 3, 109–118 (2011)

16 Lazebnik, Y. 69, (2004)

17 de Lorenzo, V. *Bioeng. Bugs* 2, 3–7 (2011)

18 Chao, R. *et al. Metab. Eng.* 42, 98–108 (2017)

19 Endy, D. *Nature* 438, 449–453 (2005)

20 Beal, J. *et al. EMBO Rep.* 21, 1–5 (2020)

21 Jessop-Fabre, M. M. *et al. Front. Bioeng. Biotechnol.* 7, (2019)

22 Del Vecchio, D. *Trends Biotechnol.* 33, 111–119 (2015)

23 Vazquez-Vilar, M. *et al. Plant Sci.* 273, 33–41 (2018)

24 Biggs, B. W. *et al. Curr. Opin. Biotechnol.* 29, 156–162 (2014)

25 Cairns, T. C. *et al. Biotechnol. Biofuels* 12, 1–18 (2019)

26 Mattern, D. J. *et al. Front. Microbiol.* 6, 1–7 (2015)

27 Deng, H. *et al. Crit. Rev. Biotechnol.* 40, 180–198 (2020)

28 Xiao, H. *et al. Trends Biotechnol.* 34, 242–255 (2016)

29 Nogueira-López, G. *et al. Front. Microbiol.* 10, 1–18 (2019)

30 Roell, M. S. *et al. Curr. Opin. Biotechnol.* 61, 102–109 (2020)

31 Ke, J. *et al. Trends Biotechnol.* 39, 244–261 (2021)

32 Hernanz-Koers, M. *et al. Fungal Genet. Biol.* 116, 51–61 (2018)

33 Vazquez-Vilar, M. *et al. Nucleic Acids Res.* 45, 2196–2209 (2017)

34 Sarrion-Perdigones, A. *et al. Plant Physiol.* 162, 1618–1631 (2013)

35 Hsu, P. D. *et al. Cell* 157, 1262–1278 (2014)

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229

36 Zhang, R. *et al. Front. Microbiol.* 12, 1–8 (2021)

37 Hanna, R. E. *et al. Nat. Biotechnol.* 38, 813–823 (2020)

38 Gao, Z. *et al. Transcription* 8, 275–287 (2017)

39 Radzisheuskaya, A. *et al. Nucleic Acids Res.* 44, (2016)

40 Schuster, M. *et al. Fungal Genet. Biol.* 130, 43–53 (2019)

41 Song, R. *et al. Appl. Microbiol. Biotechnol.* 103, 6919–6932 (2019)

42 Nødvig, C. S. *et al. PLoS One* 10, 1–18 (2015)

43 Pohl, C. *et al. ACS Synth. Biol.* 5, 754–764 (2016)

44 Liu, R. *et al. Cell Discov.* 1, 1–11 (2015)

45 Matsu-ura, T. *et al. Fungal Biol. Biotechnol.* 2, 1–7 (2015)

46 Katayama, T. *et al. Appl. Environ. Microbiol.* 85, (2019)

47 Zhang, C. *et al. Fungal Genet. Biol.* 86, 47–57 (2016)

48 Ferrara, M. *et al. Sci. Rep.* 9, 1–9 (2019)

49 Lightfoot, J. D. *et al. Microorganisms* 7, (2019)

50 Peng, D. *et al. Microb. Genomics* 1, 1–7 (2015)

51 Gao, Z. *et al. Mol. Ther. - Nucleic Acids* 10, 36–44 (2018)

52 Sesma, A. *et al. Nature* 431, 582–586 (2004)

53 Lee, S. B. *et al. Fungal Genet. Rep.* 35, 23 (1988)

54 van Rij, R. P. *et al. Trends Biotechnol.* 24, 186–193 (2006)

55 Lee, T. *et al. Plant Pathol. J.* 30, 33–42 (2014)

56 Chen, F. *et al. J. Proteome Res.* 10, 2273–2285 (2011)

57 Wang, Q. *et al. Fungal Genet. Biol.* 117, 21–29 (2018)

58 Jiang, F. *et al. RNA Biol.* 16, 950–959 (2019)

59 Pasin, F. *et al. ACS Synth. Biol.* 6, 1962–1968 (2017)

60 Borchardt, E. K. *et al. Rna* 21, 1921–1930 (2015)

Institutional Repository - Library & Information Centre - University of Thessaly
05/07/2024 09:20:54 EEST - 3.147.46.229