UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
DEPARTAMENTO DE BIOLOGIA CELULAR E MOLECULAR

Development of synthetic biology tools applied to fungi of medical and industrial importance

LUÍSA CZAMANSKI NORA

Ribeirão Preto
Brasil
2018

LUÍSA CZAMANSKI NORA

Development of synthetic biology tools applied to fungi of medical and industrial importance

Dissertação apresentada à faculdade de Medicina de Ribeirão Preto na Universidade de São Paulo para a obtenção do título de Mestre em Ciências – Área de concentração: Biologia Celular e Molecular

Orientador: Prof. Dr. Rafael Silva Rocha

Ribeirão Preto
Brasil
2018

NORA, Luísa Czamanski

Desenvolvimento de ferramentas de biologia sintética aplicadas a fungos de importância médica e industrial, 2018.

77 f. : il. ; 30 cm

Dissertação de Mestrado, apresentada à Faculdade de Medicina de Ribeirão Preto/USP. Área de concentração: Biologia Celular e Molecular.

Orientador: Prof. Dr. Rafael Silva-Rocha.

1. Biologia Sintética 2. Plasmídeo 3. Vetor binário 4. Promotores Constitutivos 5. Promotores Bidirecionais, 6. *Rhodosporidium* 7. *Paracoccidioides* 8. *Agrobacterium* 9. Engenharia Metabólica 10. Fungos

Aluna: **Luísa Czamanski Nora**

Título: **Development of synthetic biology tools applied to fungi of medical and industrial importance**

Dissertação apresentada à Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo para obtenção do título de Mestre em Ciências. Área de concentração: Biologia Celular e Molecular. Orientador: Prof. Dr. Rafael Silva Rocha

Aprovado em: ___/___/___

Banca Examinadora

Prof(a).Dr(a).:_____
Instituição:_____
Assinatura:_____

Prof(a).Dr(a).:_____
Instituição:_____
Assinatura:_____

Prof(a).Dr(a).:_____
Instituição:_____
Assinatura:_____

AGRADECIMENTOS

Primeiramente gostaria de agradecer às duas pessoas mais importantes da minha vida, meus pais: André e Valesca (ou melhor, Pingo e Val). Vocês são a razão da pessoa que me tornei e de tudo o que conquistei até hoje. Por todos os valores ensinados, e por todo o tempo, amor e sacrifícios dedicados a mim durante minha vida inteira, essa dissertação é pra vocês.

Agradeço ao Abel, pela paciência, compreensão, por ouvir minhas reclamações e meus ensaios, muitas vezes sem nem entender, e por nunca medir esforços pra me ver feliz.

À vocês três, mãe, pai e Abel: obrigada por sempre se esforçarem para que os bem mais de mil quilômetros de distância entre nós sempre parecessem muito menores.

Ao meu irmão, Tomás, e minha cunhada, Cláudia, por serem meus maiores exemplos. Por me mostrarem que temos que ir atrás dos nossos sonhos, mas nunca deixarmos de prestar atenção nas pequenas coisas que realmente importam na vida. Obrigada por terem me ensinado a ser mindfull (ou pelo menos tentado). E também por terem me dado o presente mais lindo da minha vida, meu afilhado, Caio. É por ele que eu tenho forças pra continuar nessa busca por um futuro mais sustentável pro nosso planeta.

Agradeço ao meu orientador, e também amigo, professor Rafael Silva Rocha. Rafa, apesar de sempre te chamar de chefe, a verdade é que nunca te considerei um chefe, mas sim um verdadeiro líder. Por sempre nos acompanhar de perto (e agora mais do que nunca eu sei o quanto isso é importante), entender nossas limitações, nos motivar, nos inspirar, nos envolver nas decisões do Lab e por estar mais preocupado com nossa formação do que com os resultados em si (os resultados vêm por consequência). Obrigada por sempre acreditar na minha capacidade mesmo quando nem eu acreditava mais. E obrigada por sempre criar novos desafios e nunca nos deixar estagnar. Graças a ti, eu voltei a acreditar na ciência brasileira. Obrigada.

Aos amigos do Laboratório de Proteases e Biologia do Câncer (Bruno, Marina, Rodrigo, Márcio, Gabriel, Márcia, Elaine e a agregada Carol) pelo companheirismo, pelas risadas, por tantas idas no bandex, por todos os cafés e bolos divididos. Em especial à Marina, por ser sempre meu ombro amigo e praticamente me adotar como parte da família.

Aos amigos do Laboratório das abelhas (Denyse, Giovana, Mário, Douglas, Gustavo, Matheus e os agregados Murilo, Frank, Camila e Felipe) por todos os churrascos, replays de churrascos, Marcão, Silvão, Bodeguita e por aí vai. Obrigada por sempre serem tão animados e receberem a vizinha de lab de vocês tão bem.

Aos amigos do Laboratório de Metagenômica Funcional (Luana, Guilherme, Tiago, Lucas, Euge), pela amizade, pelo empréstimo de reagentes, pelas aulas de cerveja e de queijo. Vocês fazem muita falta aqui no nosso dia-a-dia, mas fico muito feliz que agora vocês tem um cantinho próprio para aprimorar a ciência de vocês (e, claro, para fazermos festinhas). Em especial à professora Maria Eugenia Guazzaroni, por sempre

ter alguma palavra de motivação e nos inspirar a sermos não só cientistas, mas pessoas melhores.

Aos amigos do meu laboratório, o SSB Lab, obrigada por tudo. Ananda, Cauã, Lummy, Leonardo, Murilo, Bianca, Greicy, Kauan, Juliana, Maria Clara, Adriano, Marina, Gabriel (G2) e Claudinha, obrigada por terem se tornado uma família pra mim. São tantas histórias que eu não saberia nem por onde começar. Passamos por tanta coisas juntos: deadlines, reviews, festas juninas, halloweens, congressos, dias de maldade... A verdade é que, muito mais do que uma dissertação, eu vou sair desse mestrado com um acervo de memórias e histórias para contar, e devo isso a vocês. Obrigada.

Ao Bruno, Cauã e Gabs: obrigada por terem sido os melhores colegas de mestrado. Vocês transformaram o que era pra ter sido a parte mais tensa da pós-graduação nas minhas melhores lembranças. Obrigada por todas as madrugadas afora que passamos discutindo artigos, e por me fazerem rir até minha barriga doer. Por causa de vocês, todos os dias eu me pergunto "mas como é que volta?"

À Vanessa, minha roommate, por todas as conversas e desabafos, por todos os almoços, jantas e receitas compartilhadas, e por sempre cuidar de mim (principalmente quando eu ficava doente). Agradeço muito por ter te conhecido e por tudo o que tu me ensinou nesse tempo morando juntas.
À Iuly, minha nova roommate, pelo companheirismo e paciência nessa reta final do mestrado que enfrentamos ao mesmo tempo (e sobrevivemos!).

À todos do Laboratório de Biotecnologia Molecular e ao professor Roberto Silva, por terem me emprestado o tempo, o espaço, os reagentes, o fluxo mais concorrido da história e o fungo de vocês para meus tantos experimentos – que espero que um dia ainda nos tragam bons frutos.

Ao Relber, por ter levado o pLUO até Portugal para provar que ele funciona (e conseguido!). E ao professor Fernando Rodrigues e à Beatriz Ferreira pela nossa colaboração de sucesso.

Ao Joint BioEnergy Institute, na Califórnia, e todas as pessoas que trabalham lá, por terem me acolhido durante 6 meses e me ajudado a desenvolver parte da pesquisa a ser apresentada nessa dissertação.

Agradeço à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo auxílio financeiro tanto na Califórnia Brasileira (Processo FAPESP 2016/03763-3) quanto na Califórnia Americana (Processo FAPESP 2018/02227-6).

Agradeço também ao Departamento de Biologia Celular e Molecular da FMRP-USP e a todos seus os alunos, funcionários e professores por todo apoio. E pela oportunidade de organizar o Curso de Verão e o Simpósio do departamento, que foram grande parte da minha formação científica. Um agradecimento em especial à Gabi por toda a ajuda, principalmente na organização desses eventos.

Enfim, agradeço a todos que, de alguma forma, fizeram parte dessa caminhada. Afinal, ela é muito mais importante do que a linha de chegada.

“Around here, however, we don’t look backwards for very long. We keep moving forward, opening up new doors and doing new things, because we're curious… and curiosity keeps leading us down new paths.”

– Walt Disney

GENERAL INDEX

RESUMO

NORA, Luísa Czamanski. Desenvolvimento de ferramentas de biologia sintética aplicadas a fungos de importância médica e industrial. [dissertação]. Ribeirão Preto: Universidade de São Paulo, Faculdade de Medicina de Ribeirão Preto; 2018.

Conforme novas tecnologias e metodologias estão surgindo, e pesquisadores estão sedentos por ferramentas moleculares mas rápidas, mais eficientes e fáceis de usar, dominar os princípios e tecnologias do design de vetores e padronização de partes biológicas tornaram-se desafios fundamentais. Isso está abrindo espaço para o surgimento de uma disciplina inteiramente nova chamada Biologia Sintética. Esta área de estudo inovadora combina partes e módulos biológicos para criar sistemas mais confiáveis e robustos. Linhagens fúngicas são comumente alvo desses estudos, não apenas porque muitos achados fundamentais em relação à clonagem molecular surgiram das lições dadas por elas, mas também devido a um imenso e inexplorado potencial desses organismos em uma ampla gama de aplicações – desde biocombustíveis e produção de químicos finos até terapias biomédicas. Neste contexto, a presente dissertação é dividida em duas partes: a primeira diz respeito ao design e construção de uma ferramenta modular e versátil para ser aplicada em várias linhagens de fungos. Essa ferramenta é um plasmídeo binário para transformação mediada pro *Agrobacterium tumefaciens*, que foi construído em quatro diferentes versões contendo GFP ou mCherry como proteínas repórter e um gene sintético de resistência à higromicina como marcador de seleção. O vetor foi validado em *Paracoccidioides lutzii*, um patógeno oportunista humano dimórfico que é muito importante para medicina, mas ainda carecia de ferramentas genéticas eficientes. A segunda parte consiste na criação de uma biblioteca de promotores para a levedura oleaginosa *Rhodosporidium toruloides*, um promissor hospedeiro para a produção de bioprodutos a partir de biomassa, uma vez que pode eficientemente consumir açúcares C5 e C6 e aromáticos derivados da lignina. Vinte e nove promotores foram testados em um *cassete* de duplo-repórter – compreendendo ambas as proteínas fluorescentes GFP e mRuby – utilizando citometria de fluxo para análise de células únicas. A coleção de promotores apresentados neste trabalho é a maior disponível para *R. toruloides* até o momento e foi um avanço indispensável para superar a

escassez de ferramentas para este organismo. Notavelmente, também apresentamos os primeiros promotores bidirecionais descritos para essa levedura e otimizamos o protocolo de transformação. Portanto, a Biologia Sintética foi eficientemente aplicada para expandir a coleção de partes biológicas padronizadas e otimizar vetores para transformação e manipulação genética de fungos. Estas ferramentas são de valor imediato e são aplicáveis a desafios muito distintos, mas igualmente importantes: a busca de novas soluções para a saúde humana e para uma economia bio-sustentável.

Palavras-chave: biologia sintética, plasmídeo, vetor binário, fungos, promotores constitutivos, promotores bidirecionais, *Rhodosporidium, Paracoccidioides, Agrobacterium,* engenharia metabólica.

ABSTRACT

NORA, Luísa Czamanski. Development of synthetic biology tools applied to fungi of medical and industrial importance. [dissertação]. Ribeirão Preto: Universidade de São Paulo, Faculdade de Medicina de Ribeirão Preto; 2018.

As new technologies and methodologies are surfacing, and researchers are now eager for fast, enhanced and easy-to-use molecular tools, mastering the principles and technologies of vector design and standardization of biological parts have become fundamental challenges. This is making room for the rise of an entirely novel discipline called Synthetic Biology. This innovative field of study combines biological parts and modules to create more reliable and robust systems. Fungal strains are commonly the target of these studies, not only because several fundamental findings regarding molecular cloning arose from lessons given by them, but also due to an immense and much unexplored potential of those organisms in a wide range of applications – ranging from biofuels and fine chemicals production to biomedical therapies. In this context, the present dissertation is divided in two parts: the first one concerns the design and construction of a modular and versatile tool to be applied in several fungal strains. This tool is a plasmid binary vector for *Agrobacterium tumefaciens*-mediated transformation, which was built in four different versions containing either GFP or mCherry as reporter proteins and a synthetic hygromycin resistance gene as selection marker. The vector was validated in *Paracoccidioides lutzii*, a dimorphic human opportunist pathogen that is very important for health care but was still lacking efficient genetic tools. The second part consists in the creation of a promoter library for the oleaginous yeast *Rhodosporidium toruloides*, a promising host for the production of bioproducts from biomass since it can efficiently consume C5 and C6 sugars and lignin-derived aromatics. Twenty-nine promoters were tested with a dual-reporter *cassette* – comprising both GFP and mRuby fluorescent proteins – using flow cytometer for single-cell analysis. The assortment of promoters presented in this work is the largest set available for *R. toruloides* until now and was an imperative advancement to overcome the scarcity of tools for this organism. Remarkably, we also presented the first bidirectional promoters described for this yeast and optimized the transformation protocol. Thus, we efficiently applied Synthetic Biology to expand the collection of standard biological parts and to

optimize vectors for fungal transformation and genetic manipulation. These tools are of immediate value and are applicable for very distinct but equally important challenges: the pursuit of new solutions for human health and for a sustainable bio-based economy.

Keywords: synthetic biology, plasmid, binary vector, fungi, constitutive promoters, bidirectional promoters, *Rhodosporidium, Paracoccidioides, Agrobacterium,* metabolic engineering.

I. INTRODUCTION

Part of this introduction was published or submitted as:

Nora, L. C., Westmann, C. A., Martins‐Santana, L., Alves, L. D. F., Monteiro, L. M. O., Guazzaroni, M. E., & Silva‐Rocha, R. (2018). The art of vector engineering: towards the construction of next‐generation genetic tools. *Microbial Biotechnology*, 1-23.

Martins-Santana, L., **Nora, L. C.**, Sanches-Medeiros, A., Lovate, G. L., Cassiano, M. H., & Silva-Rocha, R. (2018). Systems and synthetic biology approaches to engineer fungi for fine chemical production. *Frontiers in Bioengineering and Biotechnology,* 6 (117), 1-18.

Nora, L. C., Westmann, C. A., Guazzaroni, M. E., Gupta, V. K., Silva-Rocha, R. (2018). Advanced genetic tools to engineer microbes for biotechnological applications – *submitted.*

1. Synthetic Biology and tool development

With the advent of molecular cloning techniques, seven decades ago, our ability to manipulate and engineer living systems has drastically changed. The development of plasmid vectors has revolutionized Molecular Biology, unlocking a completely new universe of possibilities through the transportation, delivery and manipulation of foreign DNA inside a host cell. More recently, the advances in DNA manipulation techniques such as automated synthesis, sequencing and assembly have been combined with Engineering principles to build the Synthetic Biology framework (Andrianantoandro *et al*., 2006; Ellis *et al*., 2011; Purnick and Weiss, 2009; Westmann *et al*., 2018), providing new perspectives on how researchers explore and re-engineer living systems.

Considering this, the study of microorganisms has been extremely relevant not only for building the grounds of molecular cloning studies, but also for revealing the potential of those organisms in a wide range of applications such as biofuels (Kircher, 2015), fine chemicals (Sheldon, 2014; Jullesson *et al.*, 2015), biosensors (Courbet *et al.*, 2015), bioremediation (Gavrilescu *et al.*, 2015), and biomedical therapies (Din *et al.*, 2016). Thus, efforts to understand, manipulate and engineer them are imperative and, as we acquire more knowledge, the classical genetic engineering approaches are no longer sufficient to answer all questions.

Fungi are outstanding microorganisms that are acquiring progressively more notice throughout the world due to not only biomedical concerns but especially due to their biotechnological applications. From all of the species of fungi that have been studied, yeasts are the most common ones to be used as hosts for genomic and metabolic engineering (Borodina and Nielsen, 2014). *Saccharomyces cerevisiae* was long established as the model organism for fungi studies, thus, most of the tools available nowadays were developed based on this species.

Recently, however, the scientific community has realized that model organisms such as *S. cerevisiae*, although highly adapted to laboratory conditions, were not suitable for a wide-range of biotechnological applications, which could be readily achieved by the usage of unconventional organisms. Thus, there is a current need for developing molecular toolboxes for taming novel microbial chassis and exploring their intrinsic metabolic/physiological properties. These toolboxes should rely mostly on the generation of stable vectors with compatible biological parts (such as

promoters, terminators, origin of replication *etc.*) for both protein expression and genome engineering (Kim *et al.*, 2016). Furthermore, for gene expression studies of pathogenic strains, it is also important to improve the versatility and orthogonality of these molecular tools, allowing vectors to be functional and to replicate in a wide range of hosts. Hence, over the last decade, researchers have been facing the great challenge of shifting from model fungal hosts to non-conventional chassis with endogenous capacities for dealing with specific tasks.

2. Evolution of vector engineering for fungi

2.1. Tools for yeast transformation

Circular DNA was already known to be present in prokaryotes over 60 years ago. In eukaryotes, more precisely in *S. cerevisiae*, circular DNA was first discovered in mitochondria, and it took several years for researchers to realize this was too prokaryotic DNA (Margulis and Chapman, 1998). A non-mitochondrial natural plasmid DNA was first described in 1971 (Guerineau *et al.*, 1971). Three years later, researchers found an antibiotic resistance gene present in those plasmids and realized they could behave just like the bacterial ones allowing DNA cloning (Guerineau *et al.*, 1974). Soon after, yeast transformation was described as well as integration of the genes in the yeast genome (Hinnen *et al.*, 1978; Cameron *et al.*, 1983). Guerineau's group noticed that yeast plasmids had a length of about 2 microns, consequently, they were named 2μ. In the years that followed, many researchers began constructing vectors for yeast manipulation, many of which used the 2μ origin (Struhl *et al.*, 1979; Ferguson *et al.*, 1981).

Plasmids used to transform *Saccharomyces* can be divided into three groups: yeast centromeric plasmids (YCps), Yeast Episomal plasmids (YEps), and Yeast Integrative plasmids (Yips). The YCps need autonomously replicating sequences (ARS) and centromeric sequences (CEN) where kinetochore complexes attach, thus behaving like a microchromosome (Clarke and Carbon, 1980; Westermann *et al.*, 2007). The YEps are based on the endogenous 2μ plasmid mentioned above with addition of a bacterial origin of replication and selection marker, yeast selection marker, and the expression cassette. Still, the YIps need to have homology sequences so they can integrate in the yeast genome via homologous recombination (Gnügge and Rudolf,

2017). For all of them, the selection marker is usually auxotrophic, which somewhat restricts their use because there is usually only URA3 (encoding orotidine-5′-phosphate decarboxylase), LEU2 (encoding 3-isopropylmalate dehydrogenase), HIS3 (encoding imidazoleglycerol-phosphate dehydratase), and TRP1 (encoding phosphoribosylanthranilate isomerase) options, besides the need for a strain with the original gene deleted. Generally, all series of vectors present the three types; in a matter of deciding which methodology to use, the biological question being examined should be considered (**Figure 1**).

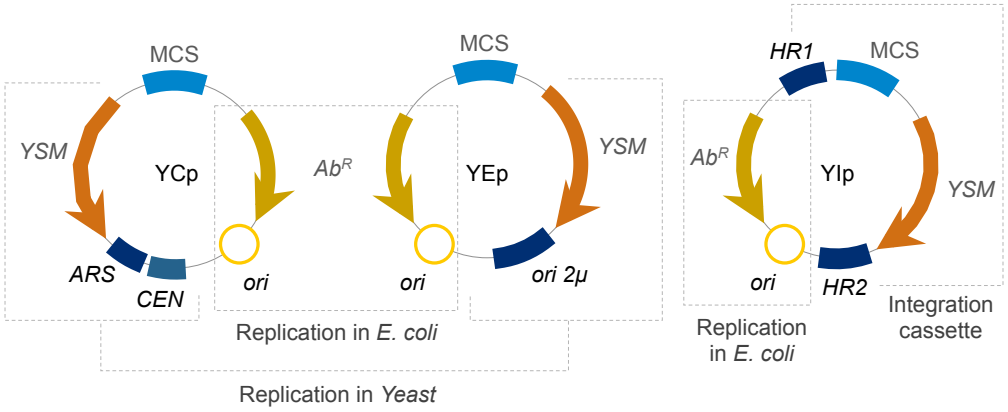

Figure 1. Modular vectors designed for yeast. The yeast centromeric plasmids (YCps) harbor the autonomously replicating sequences (ARS) and centromeric sequences (CEN), which allows the vectors to behave as mini-chromosomes. The Yeast Episomal plasmids (YEps) are endowed with the 2µ origin of replication and are similar to plasmids in bacteria. In the case of Yeast Integrative plasmids (Yips), homologous regions (labeled as HR1 and HR2) to the host chromosome allow the integration of the target region through homologous recombination events. In all vectors, the yeast selection marker (YSM) represents a gene that allows the selection of transformants harboring the vectors. In all cases, specific regions for replication of the bacterial host (usually *E. coli*) and the region required for replication or integration in yeast are highlighted.

Despite the fact that it was Hinnen and colleagues that developed a transformation method for yeast using bacterial vectors in 1978, the vectors containing the ability to replicate in both *Escherichia coli* and *S. cerevisiae* were developed a year later by Struhl *et al.* (Struhl *et al.*, 1979). Shuttle vectors, as they are called now, are still the preferred method for yeast transformation, and they are constantly evolving. Ferguson *et al.* (Ferguson *et al.*, 1981) developed the series of pRC1, pRC2, and pRC3 derived from pKC7 (which was originally derived from pBR322). Gietz & Sugino (Gietz and Sugino, 1988) built the plasmids YCplac,

YEplac, and YIplac based on pUC19 in 1988. Around the same time, Ma *et al*. (Ma *et al.*, 1987) developed a method of easy recombination of plasmid parts and established the series of shuttle vectors YCp400 and YEp400. All of these vectors were considerably large (some more than 10 kb) and had no more than 10 unique restriction sites for cloning. However, in 1989, there was the breakthrough for yeast scientists: the pRS series (Sikorski and Hieter, 1989). These vectors were made small (around 5 kb), with several restriction sites (around 13 unique sites each), four different selection markers to choose from, and since they were based on pBLUESCRIPT, positive transformants could be selected by color in *E. coli*. pRS vectors were such a turning point in this field that new designs for yeast tools were very limited for almost twenty years afterward, with scientists only adding some adaptations such as new resistance markers (Chee and Haase, 2012), and these tools are still frequently used and adapted (Avalos *et al.*, 2013).

Nevertheless, since the pRS series, some other plasmids and tools have been developed. Mikkelsen *et al.* (2012) developed a series of integrative plasmids called pX that allows multiple genomic integrations into specific chromosomic regions (Mikkelsen *et al.*, 2012). New tools have been evolving from that idea. The first example is the EasyCloneMulti series, which is the latest evolution of EasyClone (Jensen *et al.*, 2014) and EasyClone2.0 (Stovicek *et al.*, 2015). This collection of plasmid relies in a method for integration into long terminal repeats (LTR) of Ty retrotransposons sequences, enabling multiple genome integrations. These plasmids contain USER friendly site for cloning (Bitinaite *et al.*, 2007) and also present *loxP*/Cre system to enable the recycling of selection markers (Maury *et al.*, 2016). Almost at the same time, the pRG collection was created, presenting a series of modular vectors that allow multiple integrations into the genome using auxotrophic selection markers (Gnügge *et al.*, 2016). **Table 1** summarizes the most important tools for yeast manipulation from 2007 to 2017.

Additionally, as a breakthrough in the synthetic biology field applied for this model yeast, Lee et al., 2015 developed a multipart and modular toolkit to facilitate and advance yeast assembly studies (Lee *et al.*, 2015). In their work, they characterized several parts of interest: promoters, terminators and protein degradations tags. They also highlighted the importance of integration of expression *cassettes* into the genome, providing less variability to the system, over plasmid-derived expression. They provide information for increasing the success of the

integrations in two distinct ways: using the CRISPR/Cas systems, or using a homing endonuclease I-SceI to generate a double-strand break and stimulate homologous recombination. They rely on MoClo strategy (Weber *et al.*, 2011) to build their devices so one can easily replicate that, making it user friendly and easily accessible as any toolkit should be. Other native promoters were characterized by Partow et al. (2010) and by Lee et al. (2012) whereas Blazeck et al. (2012) built synthetic hybrid promoters to enhance promoter activity (Partow *et al.*, 2010; Blazeck *et al.*, 2012; Lee *et al.*, 2013). Other terminators are also well understood for *S. cerevisiae* as for this work by Yamanishi et al. (2013) where they characterized more than 5 thousand terminator regions (Yamanishi *et al.*, 2013).

One of the greatest advantages of *S. cerevisiae* over other yeast species is the fact that transformation methods are completely standardized (Gietz, 2014) and do not offer any drawback for the researchers – in fact, it can even be performed by robotic platforms (G. Liu *et al.*, 2017). For new emergent strains, however, this still constitutes a problem, since not all fungi behave the same for each transformation method, meaning that protocol optimization is often required (Kawai *et al.*, 2010; Park *et al.*, 2018). Different methods for fungi transformation can be seen in **Figure 2**.

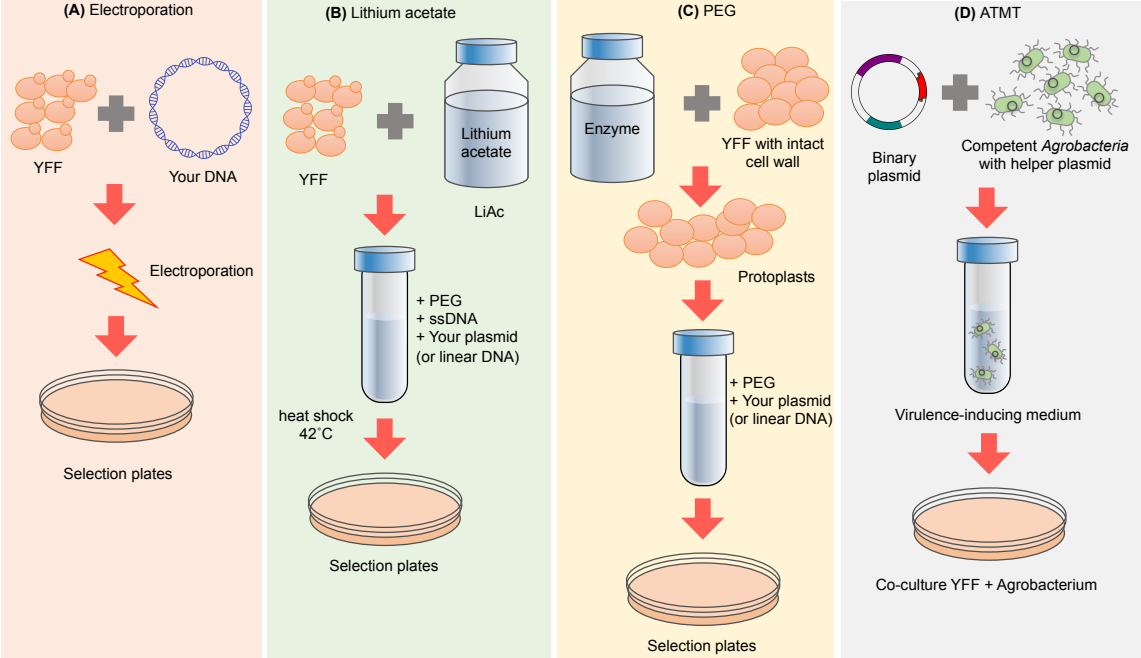

Figure 2. Different methods available for transformation of fungi. **(A)** Electroporation method consists of providing a pulse of electricity to open pores in the cell membrane. It is the most common method for bacteria and occasionally works for some yeast species, but it is difficult to be applied to fungi due to their thick cell wall. **(B)** Lithium acetate (LiAc) method is the most common one for transformation of yeasts since the LiAc can permeabilize their cell walls, making them competent. Single-stranded DNA (ssDNA) is added to favor the entrance of the double-stranded DNA into the nucleus and polyethylene glycol (PEG) is added to facilitate the passage of DNA through the membrane. Heat-shock is also needed to open pores in the membrane. **(C)** Polyethylene glycol (PEG) method consists in using enzymes to lyse the cell wall of the fungal strains. The cells without cell wall are called protoplasts, which are now ready to receive exogenous DNA – that will pass through the membrane with the assistance of PEG. This method is frequently used for filamentous fungi. **(D)** *Agrobacterium tumefaciens*-mediated transformation (ATMT) can be used for both yeast and filamentous fungus species. It consists in: (i) transforming an *Agrobacterium* strain – that already contains a helper plasmid – with the binary plasmid containing the DNA of interest (ii) growing this *Agrobacterium* in a medium containing antibiotics and acetosyringone to induce the activation of the virulence genes that are needed for the transfer of the DNA to the host, and (iii) co-cultivating both the bacterium and the fungal strains in agar plates. After a few days growing in conditions favorable to *A. tumefaciens*, cefotaxime is added to new plates to kill the bacteria and then one can start selecting for their positive fungi transformants. It is the most demanding and time-consuming method of all, but it is also the one with highest transformation frequencies. YYF means Your Favorite Fungus.

2.2. Tools for filamentous fungi

Since genetically manipulating filamentous fungi is considerably more complex than manipulating yeast, much discussion has been raised to improve the molecular tools for genetic transformation of filamentous species (Kawai *et al.*, 2010). *Agrobacterium tumefaciens*-mediated transformations (ATMT) – besides being used sometimes in yeasts as well – was one of the greatest breakthroughs for filamentous

species allowing a broader range of filamentous fungi to be transformed. It is indeed the preferred method for this type of fungus not only for being less complex than the need for protoplasts in the polyethylene glycol (PEG) method, but also leads to a higher frequency of recombination – single-copy or multiple-copy (Hooykaas *et al.*, 2018). ATMT was shown to work in several genera such as *Aspergillus*, *Fusarium*, *Neurospora*, *Penicillium*, *Trichoderma*, among others. A more detailed schematic representation of genetic tools using the ATMT is shown in **Figure 3**.

For the filamentous fungi group, the two genera that are the focus of recent studies due to their significant applications in industry are *Trichoderma* and *Aspergillus*. *Trichoderma*, specifically *Trichoderma reesei,* has been extensively studied due to its capacity to produce highly efficient biomass degrading enzymes (Kubicek *et al.*, 2009). These enzymes are already commercialized as enzyme cocktails such as Accelerase® from DuPont and are being applied in several biofuel plants around the globe (Bischof *et al.*, 2016). *Aspergillus* genus is greatly represented by *Aspergillus niger*, which also produces and secretes important cellulolytic enzymes – some studies even take advantage of enzymes of both organisms together to degrade biomass (Florencio *et al.*, 2016) – but it is mostly known for the production of organic acids. In fact, production of citric acid by this organism generates such high yields that it is being used as a model for other fermentation studies (Baker, 2006). The fact that both strains have their whole genome sequenced also makes them promising targets for genome engineering (Baker, 2006; Martinez *et al.*, 2008). However, genetic manipulation in these organisms is not trivial and, although many have successfully transformed using PEG protocol – for *Trichoderma*: (Steiger *et al.*, 2011; Jørgensen *et al.*, 2014); for *Aspergillus*: (Storms *et al.*, 2005; Arentshorst *et al.*, 2015) – existing tools still largely rely on ATMT as a consistent method due to higher transformation frequencies and the possibility of either multiple or single-copy integrations (Michielse *et al.*, 2008; Hooykaas *et al.*, 2018).

As examples of ATMT tools developed for *Trichoderma* genus are the vectors pCBGW-GFP and pGWB2-GFP that were used for gene expression studies (Zhu *et al.*, 2009), and the ATMT plasmids pWEF31 and pWEF32 developed by Lv and colleagues (2012) both using Hygromycin B as selection markers (Zhu *et al.*, 2009; Lv *et al.*, 2012a). Later, Zhang *et al.,* 2016 created several versatile integration vectors to transform *T. reesei* using *bar* (the phosphinothricin acetyltransferase gene) as a selection marker (L. Zhang *et al.*, 2016). For *Aspergillus*, two versatile vectors

were recently developed for gene expression studies: pEX1 and pEX2 (Nguyen *et al.*, 2016). Promoters and terminators are usually constitutive ones previously characterized in *Aspergillus* such as *trpC* and *gpdA* (Hamer and Timberlake, 1987; Gressler *et al.*, 2015), whereas Yin *et al.*, 2017 reported a promoter P*gas* from *A. niger* to be modulated by low pH (Yin *et al.*, 2017). **Table 1** summarizes some of the most important vectors built for fungal transformation.

Filamentous fungi are, by far, the toughest organisms to transform and manipulate among all microbes (Ruiz-Díez, 2002). Even with the advances towards the state-of-the-art vector creation, fungal plasmids used for genetic engineering are still poorly understood, unreliable, and inefficient, especially when compared to bacterial ones. Still, the increasing development of synthetic biology studies focused on fungi has strongly contributed to the rising and constant necessity of creating new tools for fungal applications in several areas of molecular biology (Amores *et al.*, 2016). New forms of vector design, plus characterization and standardization of their parts, are crucially needed to promote a better understanding of fungi molecular mechanisms.

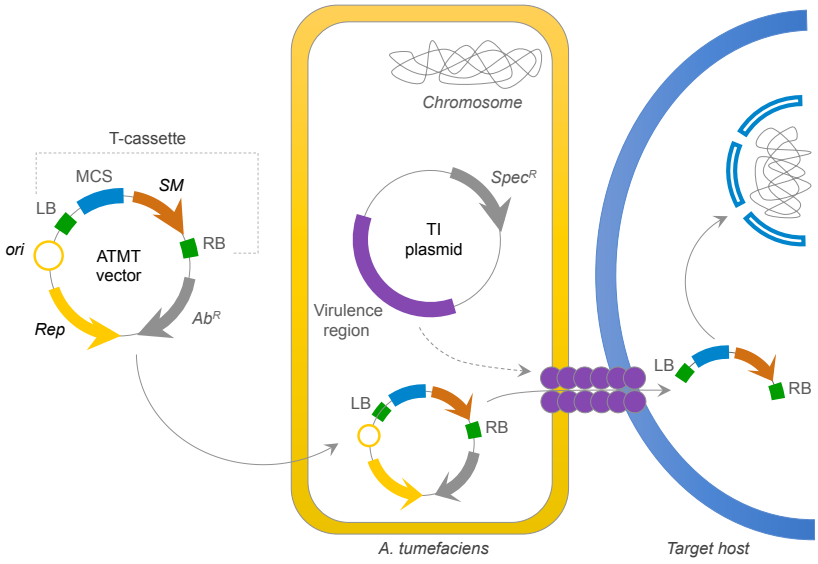

Figure 3. Genetic tools based on *Agrobacterium tumefaciens*-mediated transformation (ATMT). ATMT vectors are based on broad-host range plasmids and harbors a T-cassette, which is composed of a MCS and a selection marker (SM) flanked by the left and right borders (LB and RB) required for the recognition of the *A. tumefaciens* machinery. Once inserted in the proper *A. tumefaciens* strain harboring the TI plasmid (which expresses the components for T-cassette mobilization), this vector can be used to introduce the T-cassette into hosts such as fungi and plants.

Table 1. Plasmid vectors used for fungal transformation.

Plasmid Series Name	Fungal Selection Marker	Approximate Average Vector Size	Features	Fungi Type	References
pAG	HIS3, LEU2, TRP1, URA3	7 kb	More than 200 options; contains fluorescence reporters	Yeast	(Alberti *et al.*, 2007)
pXP	HIS3, LEU2, MET15, TRP1, URA3	5 kb	Recycling of selection markers by loxP/Cre technology	Yeast	(Fang *et al.*, 2011)
EasyClone	HIS3, LEU2, LYS5, URA3	6 kb	Multiple integrations; recycling of markers.	Yeast	(Jensen *et al.*, 2014)
EasyClone2.0	*amds, ble, dsd, hph, kan, nat*	6 kb	Compatible with prototrophic strains; recycling of markers	Yeast	(Stovicek *et al.*, 2015)
EasyCloneMulti	*Kl*.URA3-*degradation signal*	6 kb	Integrates into Ty sequences; recycling of markers.	Yeast	(Maury *et al.*, 2016)
pRG	HIS3, LEU2, LYS2, MET15, URA3	6 kb	Modular design; multiple integrations; recycling of markers	Yeast	(Gnügge *et al.*, 2016)
pWEF	*hph*	12 kb	Binary vector	Filamentous	(Lv *et al.*, 2012b)
pDESTR	*hph*	5 kb	Gene targeting and disruption	Filamentous	(Abe *et al.*, 2006)
pCBGW-GFP	*hph*	8 kb	Expression vector	Filamentous	(Zhu *et al.*, 2009)
pGWB2-GFP	*hph*	Not shown	Binary vector	Filamentous	(Zhu *et al.*, 2009)
pEX1 and pEX2	*pyrG*	10 kb	Binary vector	Filamentous	(Nguyen *et al.*, 2016)
pBI-hph	*hph*	15 kb	Binary vector	Filamentous	(Zhong *et al.*, 2007)
pALS-1	*qa-2+*	13 kb	Tested in *N. crassa.*	Filamentous	(Sthol and Lambowitz, 1983)

3. Fungi applied to biotechnological and biomedical advances

3.1. Non-conventional fungi as industrial hosts

Despite the fact that *S. cerevisiae* is an exceptional model organism for genetic studies, other organisms present extra favorable traits to be considered for production in large scale (e.g.: higher growth rates, complex post-translational modifications, resistance to fluctuating environment and/or extreme temperatures). Organisms that are able to grow on low cost feedstocks being capable of degrading a broad range of carbon sources such as C5 sugars and aromatic compounds present in lignocellulose hydrolysates are also preferable (Vogl *et al.*, 2013; Wagner and Alper, 2016; Park *et al.*, 2018). Some episomal vectors have been developed to study these emergent fungi species, however, integrative vectors are usually preferred in industrial strains as well, since, as already mentioned, the expression of the genes of interest should be as stable as possible (Obst *et al.*, 2017; Vogl, Gebbie, *et al.*, 2018). Some of the fungi that have been mostly studied and/or applied to biotechnology on an industrial scale are cited below.

Yarrowia lipolytica is an oleaginous yeast capable of assimilating alkanes. This yeast can produce higher yields of products derived from fatty acids like omega-3 eicosapentaenoic acid and lycopene and other products such as citric acid and erythritol (Wagner and Alper, 2016; Darvishi *et al.*, 2018). A *loxP*/Cre recombination system was a pioneer study for this oleaginous yeast. De Pourcq *et al.* (2012), while trying to humanize the glycosylation patterns in proteins, developed this method that not only successfully transformed their cassette into the yeast genome but also enabled the recycling of the selection marker (de Pourcq *et al.*, 2012). Yet, in 2017, Bredeweg et al. (2017) developed some strains to facilitate integration where they replaced the NHEJ recombination gene *ku70* for an Hygromycin B resistance gene (Bredeweg *et al.*, 2017). They also created an integrative vector called pYL15 which enabled the study of localization of proteins fusion to GFP. EasyCloneYALI is the most recent toolbox developed to facilitate engineering of this organism through gene integration. From this collection of 16 plasmids, 11 vectors can be employed for integration using the following markers: Uridine auxotrophy and Nourseothricin or Hygromycin B resistance*;* the other 5 vectors can be applied to engineer its genome by the CRISPR/Cas method. Just as the EasyCloneMulti for *S. cerevisiae*, this series

for *Y. lipolytica* is also based on USER method for cloning (Holkenbrink *et al.*, 2018). This non-model organism has many parts characterized, such as promoters, terminators and selection markers (Wagner and Alper, 2016; Darvishi *et al.*, 2018), as well as synthetic hybrid promoters (Blazeck *et al.*, 2011). Also, it has already been successfully transformed by lithium acetate (Bredeweg *et al.*, 2017) and electroporation (Markham *et al.*, 2018) including by high-throughput method in a 96-well plate (Leplat *et al.*, 2015).

Pichia pastoris is frequently applied to heterologous protein production – it is actually the most commonly used eukaryote for single protein expression (Vogl, Gebbie, *et al.*, 2018). This species has been widely used for its capacity of making fully humanized glycosylation. For example, there is a tool called *Pichia* GlycoSwitch® System (https://pichia.com/glycoswitch/) comprising strains and vectors that one can use to have proteins with different forms of glycosylation. *P. pastoris* is also used in the pharmaceutical industry to produce Kalbitor®, a protein used to treat hereditary angioedema (Vogl *et al.*, 2013). The study of this yeast is facilitated by the fact that it has high-quality genome sequences available for strain engineering (De Schutter *et al.*, 2009; Küberl *et al.*, 2011; Vogl *et al.*, 2013) and most of the tools available are based on genomic integration (Vogl, Gebbie, *et al.*, 2018). In 2011, Pan *et al*. developed a tool for making multiple deletions in the *P. pastoris* genome using the *LoxP/Cre* technology and allowing the recycling of the marker – which, in this paper, was Zeomicin® (Pan *et al.*, 2011). Similar work was done later by Naatsaari *et al*., (2012), where they also achieved multiple integrations and recycled the marker by using the FLP recombinase system (Näätsaari *et al.*, 2012). Obst *et al*., (2017) developed a multipart toolkit for which they characterized 124 constructs that can be efficiently applied to protein production and secretion in this yeast (Obst *et al.*, 2017). Besides the antibiotic marker Zeomicin®, auxotrophic markers can also be employed when engineering *Pichia*, and transformation methods vary from electroporation to polyethylene glycol (PEG) (Obst *et al.*, 2017; Vogl, Gebbie, *et al.*, 2018). Regarding characterization of parts, the promoter for AOX1 (alcohol oxidase 1) gene is the most widely used since it is inducible and strong (Vogl and Glieder, 2013; Wagner and Alper, 2016). Throughout the years, this promoter was engineered multiple times to facilitate studies and create synthetic variants with different features, including regulation by sugar depletion (Hartner *et al.*, 2008; Xuan *et al.*, 2009; Vogl and Glieder, 2013). For terminators, this and other yeasts still rely

on the ones that were initially characterized for *S. cerevisiae*, such as CYC1*t* (Wagner and Alper, 2016).

Kluyveromyces is another important genus of yeasts for the biotechnology industry. The species *Kluyveromyces lactis* is used in industry to produce the lactose degrading enzyme commercially called lactase, and is also the source of the marketable milk clotting enzyme bovine chymosin, among several other important proteins (Van Ooyen *et al.*, 2006; Spohner *et al.*, 2016). Whereas *Kluyveromyces marxianus* is a thermo-tolerant microorganism that can be used to produce bioethanol in higher temperatures, reducing not only contamination but also the expenses of refrigeration in bioreactors (Yang *et al.*, 2015; Nambu-Nishida *et al.*, 2017). Genes from both species can usually be expressed using *Saccharomyces* promoters and terminators. However, Yang et al. (2015) have shown the native promoters from *K. marxianus* to be stronger than the ones from *S. cerevisiae* (Yang *et al.*, 2015). The selection markers are usually acetamide (*amdS*), G418 and Hygromycin B; in addition to auxotrophic markers for uridine, leucine and tryptophan as well, and they are generally transformed using the lithium acetate protocol (Colussi and Taron, 2005; Van Ooyen *et al.*, 2006; Wagner and Alper, 2016; Nambu-Nishida *et al.*, 2017). Tools for these two relevant non-conventional yeasts are being the targets of recent studies, since their growing importance in ethanol and food industry. Multiple genomic integrations using the *loxP*/Cre method have been already established for *K. marxianus* (Ribeiro *et al.*, 2007), while Nambu-Nishida *et al*. (2017) developed strains with impaired NHEJ proteins to facilitate homologous integration in this yeast (Nambu-Nishida *et al.*, 2017). Colussi and Taron (2005) engaged *K. lactis* to produce bovine enterokinase, which is toxic to *E. coli*, using a semi-synthetic promoter called pLAC4-PBI. They created an exceptional integrative vector called pKLAC1 containing the aforementioned promoter, that can be used in this organism to express proteins that bacteria are usually unable to (Colussi and Taron, 2005). This worthwhile vector is continuously utilized nowadays for protein expression in this host (Lambertz *et al.*, 2016).

3.1.1. *Rhodosporidium toruloides* as an emergent microbial cell factory

R. toruloides is emerging as an interesting and game-changer host for production of fine chemicals manly because of two important characteristics: (i) it can process complex carbon sources more efficiently, such as C5 and C6 sugars, as well as lignin derived aromatics; (ii) it has a high rate of lipid production that can generate several industrial products. Recent studies are focusing on its central carbon metabolism, since this organism can store huge amounts of Acetyl-CoA, a compound whose metabolism can generate several added value and extremely important products such as fatty acids (for biodiesel), polyketides (for antibiotics), and terpenoids – such as amorphadiene, which is a precursor of the antimalarial drug artemisinin (Yuan and Ching, 2016; Campbell *et al.*, 2017a; Soccol *et al.*, 2017; Yaegashi *et al.*, 2017). Yaegashi *et al.* (2017) have reported a high production of two terpenes: bisabolene, a precursor of biodiesel and amorphadiene, precursor of artemisinin, by *R. toruloides*. Remarkably, the production of these chemicals was increased when the organism grew using lignocellulose hydrolysates when compared to purified substrates. Additionally, *R. toruloides* is usually called "pink yeast" since it can produce high amounts of carotenoids. The carotenoids are also valuable to industry since they are precursors of vitamin A or can be used as colorants and antioxidants (Park *et al.*, 2018). Hence, a microbial cell factory that presents these noteworthy characteristics must be further explored to generate advances in biorefinary.

For tool development in this organism, some promoters have already been characterized, including some constitutive (Liu *et al.*, 2013; Wang *et al.*, 2016), inducible (Johns *et al.*, 2016), and some strong intronic promoters (Liu *et al.*, 2016). The selection markers most commonly used for this species are antimicrobial such as Hygromicin B, Nourseothricin, Bleomycin and G418 (Lin *et al.*, 2014; Johns *et al.*, 2016), while (Yang *et al.*, 2008) identified the gene URA3 in *Rhodosporidium* to be used as an auxotrophic selection marker. Additionally, the most common method used for transformation is the *Agrobacterium tumefaciens*-mediated transformation (ATMT) for single or multiple integrations (Liu *et al.*, 2013; Lin *et al.*, 2014), and just recently there were reports of transformation by electroporation (H. Liu *et al.*, 2017) and lithium acetate methods (Tsai *et al.*, 2017a). Nonetheless, more parts need to be better characterized to overcome the scarcity of tools available for this organism, as well as the transformation methods should be enhanced for a high-throughput

screening of mutants. Hence, synthetic biology focused in *R. toruloides* is promising area since there is still space for improvement.

3.2. Fungi of biomedical importance

The increase of life-threatening systemic fungal infections has been a major concern in recent years, especially duo to antifungal drug resistance. The expansion of those infections in the immunocompromised patient population (e.g. HIV-infected patients, patients with cancer who have chemotherapy-induced neutropenia and transplant recipients) is also gaining more attention of health care entities (Perea and Patterson, 2002). The majority of fungal infections worldwide are caused by *Candida* species, with *Candida albicans* being the most common etiological agent of fungal bloodstream infections. Together with *Candida* infections, there are other major opportunistic fungal pathogens such as *Aspergillus fumigatus* and *Cryptococcus neoformans;* and other less common fungi, many of which are intrinsically resistant to the available antifungal drugs, such as Zygomycetes *(Rhizopus arrhizus, Absidia corymbifera*, and *Rhizomucor pusillus), Fusarium* species, *Trichosporon beigelii, Blastoschizomyces capitatus,* among others (Perea and Patterson, 2002). Additionally, in Latin America, the fungus *Paracoccidioides brasiliensis* causes one of the most prevalent systemic mycoses called paracoccidioidomycosis (Almeida *et al.*, 2010).

Altogether, these aforementioned pathogenic fungi have many complex mechanisms not only to infect human cells but also to resist to antifungal drugs. Information about the clinical, cellular, and molecular factors contributing to fungal infection and drug resistance are much needed to understand and, therefore, overcome these difficulties (Kontoyiannis and Lewis, 2002). Thus, studies of gene expression are of extreme value in the pursuit of dealing with major fungal infections.

3.2.1. *Paracoccidiodes lutzii* as a case of study

Paracoccidioidomycosis (PCM) is a chronic pulmonary disease that affects both immunocompetent and immunodepressed individuals, with high incidences in Brazil, Argentina, Colombia, and Venezuela (Teixeira *et al.*, 2013). Acquisition of the disease occurs through inhalation of fungal propagules resulting in several clinical manifestations, starting from asymptomatic pulmonary lesions that can lead to systemic generalized infections (Almeida *et al.*, 2010). The disease was discovered in

1908 by Lutz and Splendore and since then it has been attributed to the thermodimorphic fungus *Paracoccidioides brasiliensis* (Marques-da-Silva *et al.*, 2012).

More recently, it was shown that the monotypic taxon created to accommodate the etiological agent of PCM, *P. brasiliensis*, was a complex of cryptic species, comprising at least four different phylogenetic, potentially human-pathogenic lineages. Three of these lineages differing were identified as S1, PS2, and PS3 by Matute *et al.* (2006), and a fourth group phylogenetically distinct from those was classified as the "Pb01-like group," duo to the similarity of the strains with isolate Pb01 (Matute *et al.*, 2004; Carrero *et al.*, 2008; Theodoro *et al.*, 2008). Teixeira *et al.*, (2009), however, showed that 'Pb01-like' strains have more elongated conidia with a more extended size range than the other phylogenetic groups by morphological analysis. Yet, in 2013, Teixeira *et al*. (2013) identified, named and described this strain as a new *Paracoccidiodes* species, *Paracoccidioides lutzii*. In Brazil, a higher incidence of *P. lutzii* is found in the western-central region (Scorzoni *et al.*, 2015), as can be seen in the map in **Figure 4**. Concerns over this species are increasing since more cases have been described (Marques-da-Silva *et al.*, 2012). Also, a report was made of an atypical case of PCM caused by *P. lutzii* where the disease was fatal and progressed rapidly, even though the patient did not present any evidence of immunosuppression, and there was no response to classical drug administration (Hahn *et al.*, 2014). This shows that is extremely important to have a better understanding of the fungal molecular functioning to circumvent this illness.

The development of efficient genetic transformation systems for fungi has been crucial for establishing a link between molecular mechanisms and *in vivo* function (Magee *et al.*, 2003). As can be seen in **Figure 2**, transformation methods using electroporation, protoplasting, and lithium acetate were initially developed for the transformation of several non-pathogenic fungi, but *Agrobacterium tumefaciens*-mediated transformation (ATMT) led to a further expansion of the range of fungal species that could be transformed, including the *Paracoccidioides* genus (Almeida *et al.*, 2007). Still, the major drawback in current molecular studies of both pathogenic and non-pathogenic species is the dependence on the use of enormous plasmid vectors such as pUR5750 (Menino *et al.*, 2012) and pAD1625 (Abuodeh *et al.*; Leal *et al.*, 2004), all of which consist in more than 10 kilobases. These gigantic vectors usually cause several complications for molecular cloning, emphasizing the need for smaller

and more robust tools. Additionally, a versatile tool that could work in several distinct species would also be an ideal facilitator for undercurrent fungal studies.

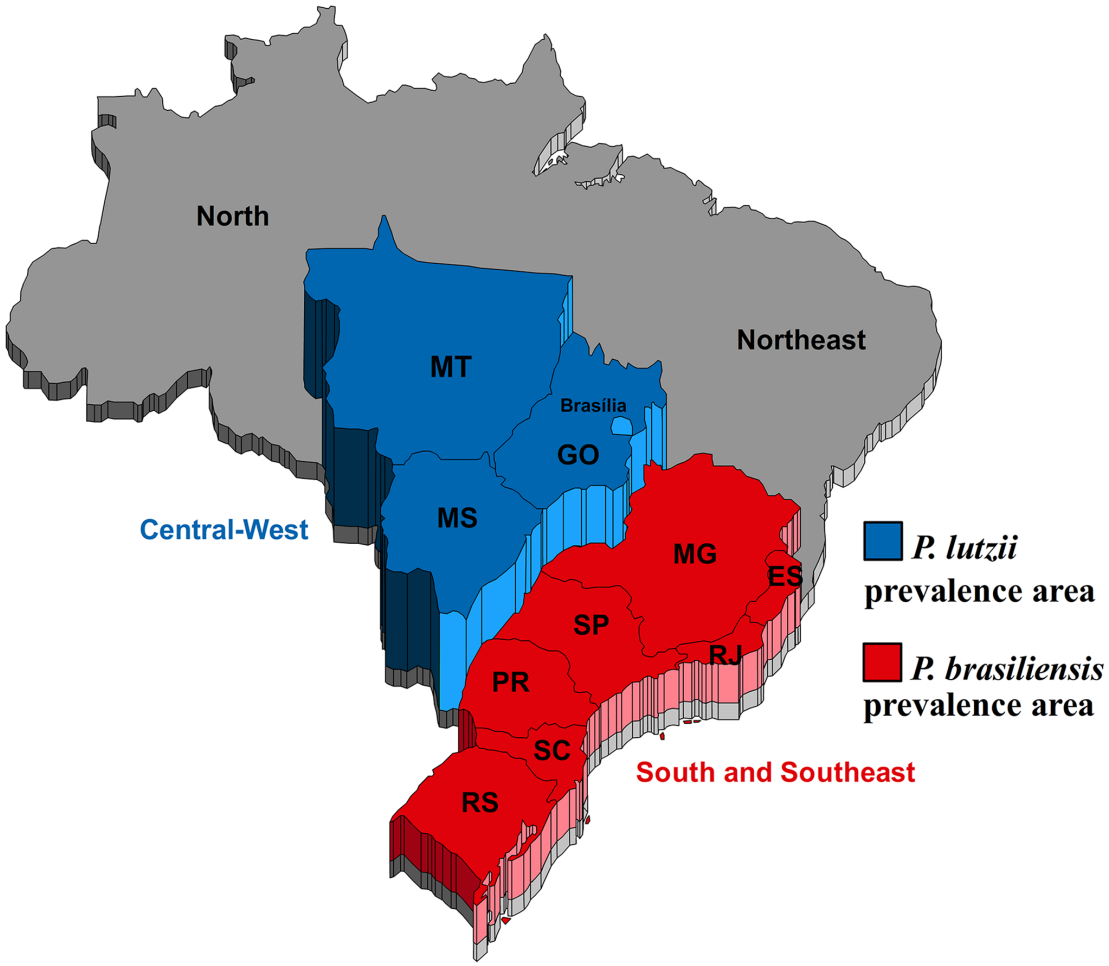

Figure 4. Map of Brazil, highlighting the regions where *Paracoccidiodes* strains are more prevalent. *P. lutzii* predominates in the Central - West region while *P. brasiliensis* predominates in the the South and Southeast region. Figure extracted from Gegembauer *et al.*, (2014).

II. OBJECTIVES

General Objective

To use Synthetic Biology-based approaches to design and develop tools to satisfy the increasing demand for molecular engineering in important fungal strains.

Specific Objectives:

A) To design and construct minimal and synthetic plasmid vectors for *Agrobacterium tumefaciens*-mediated transformation to facilitate gene expression studies in both yeast and filamentous fungi species.

B) To select, build and analyze a set of distinct promoter constructs enabling the creation of a promoter library for the emergent oleaginous yeast *Rhodosporidium toruloides*.

III. RESULTS

CHAPTER ONE

Synthetic and minimalist vectors for *Agrobacterium tumefaciens*-mediated transformation of Fungi

This chapter was published as:

Nora, L. C., Gonçales, R. A., Martins-Santana, L., Ferreira, B. H., Rodrigues, F., Silva-Rocha, R. (2018). Synthetic and minimalist vectors for *Agrobacterium tumefaciens*-mediated transformation of Fungi. *Genetics and Molecular Biology*.

1. Specific Background

Fungi are organisms comprising a universe that has not been fully explored by mankind (Leigh *et al.*, 2003), but have been extensively studied because of their huge impact in everyday life and their endless applications in industry such as production of biofuels (Glass *et al.*, 2013), foods and feedstock (Bhat, 2000), human therapeutics (Ward, 2012), among many others. Likewise, even greater efforts are being engaged in studying their pathogenicity (Almeida *et al.*, 2007; Teixeira *et al.*, 2013). Tools that can provide a better understanding of the molecular mechanisms that control gene expression in those organisms are useful, not only for shedding light on their functioning, but also because it can be used for genetic engineering and delivery of products. Synthetic Biology is an ever-growing field responsible for building new genetic circuits with known biological parts, and a great amount of the challenge in this area is in finding minimal synthetic vectors that provide a desirable setting for this cycle of re-designing parts (Silva-Rocha *et al.*, 2013; Pasin *et al.*, 2017). Fungi and synthetic biology are a promising combination that is opening brand-new doors for science, however, there is still plentiful of work to be done (Amores *et al.*, 2016). In the pursuit of supplying the lack of tools like this for fungal studies, we developed the pLUO vectors, a collection of minimal and versatile binary plasmid vectors for *A. tumefaciens*-mediated transformation (ATMT).

2. Materials and Methods

2.1. Construction of vectors

For the construction of the plasmids, the eGFP protein was amplified from pMCB17-apx (Fernandez-Abalos *et al.*, 1998) and its terminator was from *cyc1* gene from *S. cerevisae* genome. These fragments were fused using overlapping PCR with Phusion® High Fidelity DNA Polymerase (NEB) adding restrictions sites for HindIII, SpeI and BamHI. The fragment was cloned into the high-copy number vector pUC19 (Yanisch-Perron *et al.*, 1985) using HindIII and BamHI for the digestion reaction, and then transformed into chemocompetent *E. coli* DH10B. The *hph* gene is from pGL4.14 (Promega®) and its terminator was amplified from the *adh1* gene from *S.*

cerevisae. The variations of promoters modulating the *hph* gene were the following: the *ura3* promoter (*Pura3*) for yeast was amplified from pRS426 shuttle vector (Christianson *et al.*, 1992), and *Prp2* for filamentous fungi was amplified from *T. reesei* genome (He *et al.*, 2013). For construction of the yeast *cassette*, the three fragments – *Pura3*, *hph*, *Tadh1* – were fusion by overlapping PCR and the restriction sites for XbaI, SpeI and EcoRI were added by primers. Then, this fragment was digested with XbaI and EcoRI and inserted into pUC19 containing the reporter module eGFP_*Tcyc*, digested with the same enzymes. The entire expression *cassette* was amplified from pUC19 using primers to include the borders and, then, cloned into the pGLR2 vector using *PacI* and *SpeI.*

The variations in the *cassette* were all built by a few reactions of overlapping PCR using the first one as template. All enzymes used in this work were from New England Biolabs and all primers are shown in **Table 2**. The strain of *A. tumefaciens* was LBA1100 with a disarmed octopine-type pTiB6 plasmid (Menino *et al.*, 2012) and was transformed with the vectors by electroporation.

2.2. Transformation of *P. lutzii*

P. lutzii transformation through ATMT was done as described in (Menino *et al.*, 2012). Colonies of *P. lutzii* were randomly selected and plated into solid BHI media containing 75 μg/mL Hygromicin B three consecutive times. Subsequently, they were serially transferred to media with or without the selection marker for three times each, totalizing nine rounds of selection, growing for 15-20 days between each round.

Table 2. Primers used in this work. Sites for restriction enzymes are underlined.

Name	Sequence (5'- 3')	Target DNA
5'_GFP_*HindIII*	GCGCAAGCTTCGGTATCGATCATGAGTAAAG	pMCB17-apx
3'_GFP	TTAAGCCGGCGCGCC	pMCB17-apx
5'_*Tcyc*_GFP	GGCGCGCCGGCTTAACTCCTCCCACATCCGC	*S. cerevisae*
3'_*Tcyc_SpeI_BamHI*	GCGCGGATCCACTAGTAAGCCTTCGAGCGTCCC	*S. cerevisae*
5'_*hph*	CTGACACTAGCGCCACC	pGL4.14
3'_*hph*	GTTTAAACTCGACCTACCTCC	pGL4.14
5'_*Pura3_XbaI*	GCGCTCTAGAGTGCACCATACCACAGC	pRS426
3'_*Pura3_hph*	TGGCGCTAGTGTCAGTGAGATTTATCTTCGTTTCCTGC	pRS426
5'_*Tadh1_hph*	AGGTCGAGTTTAAACGGTAGATACGTTGTTGACAC	*S. cerevisae*
3'_*Tadh1_SpeI_EcoRI*	GCGCGAATTCACTAGTGTGGTCAATAAGAGCGACC	*S. cerevisae*
5'_LB_GFP_*PacI*	GCGCTTAATTAATGGCAGGATATATTGTGGTGTAAACATAAC AATTTCACACAGGACCTAGG	pLUO
3'_RB_*Tadh1_SpeI*	GCGCACTAGTGTTTACCCGCCAATATATCCTGTCAGTGGTCA ATAAGAGCGACC	pLUO
5'_*Tcyc*	CTCCTCCCACATCCGC	pLUO
3'_*Tcyc*	AAGCCTTCGAGCGTCC	pLUO
5'_*HindIII*_mCherry	GCGCAAGCTTGGTATGGTGAGCAAGGGC	pMR1
3'_mCherry_*Tcyc*	GGTTAGAGCGGATGTGGGAGGAGTTACTTGTACAGCTCGTCC	pMR1
5'_*Tcyc_Prp2*	GGTTTTGGGACGCTCGAAGGCTTCGGCTGCGTGAACAGACG	*T. reesei*
3'_*Prp2_hph*	GGTGGCGCTAGTGTCAGGTGGTTTGAGTTGGGTTGAGATAGG	*T. reesei*

3. Results e Discussion

The pLUO vectors were constructed using minimal essential parts so that they could be reduced in size still keeping its functionality. This was achieved by employing the pGLR2 plasmid as vector backbone (Benedetti *et al.*, 2012) that is also minimal and presents a broad-host range RK2 origin of replication, so it replicates in *E. coli* and in *A. tumefaciens*. pLUO vectors present a multiple cloning site (MCS) with eleven different restriction sites for several cloning options, so any given promoter can be placed modulating a red (mCherry) or a green (eGFP) reporter protein. The selection marker is a synthetic, codon-optimized and free from restriction

sites gene allowing resistance to Hygromicin B (*hph*) flanked by two different optimized promoters – so one can choose to transform it into yeast using *Pura3* or into filamentous fungi using *Prp2* – and a terminator (*Tadh1*). Two regions of 25 direct imperfect repeats were added in both ends of this *cassette*, the left and right borders, so that *A. tumefaciens* can recognize, nick and transfer the DNA from the binary vector to the host (**Figure 5**). The method of ATMT for fungi has been widely used for a long time due to its high yield of positive transformants (Michielse *et al.*, 2008).

The validation of pLUO vectors was performed in the fungus *P. lutzii*, a dimorphic human opportunist pathogen. Most of the vectors used for its molecular studies contained more than 15,000 base pairs (Almeida *et al*., 2007), making pLUO a desirable substitute, since it comprises only 6 kb. Six rounds of selection were performed as established to reach mitotic stability. Three more round were done to verify stability (**Figure 6A**) with the expression *cassette* (**Figure 6B**), for a total of nine rounds. The electrophoresis gel shows that the transformants were positive for *hph* gene, which proves that the transformation was successful (**Figure 6C-D**). Henceforward, this vector would be applicable as an efficient method to study gene expression in this pathogenic organism. The expression *cassette* tested contains a eGFP reporter and *Pura3* modulating *hph* – the following versions were all adapted by overlapping PCR reactions. Thus, expansion of the modules can be performed by using the primers provided (**Table 2**), or variations, to build additional versions that suit their own target organism. *P. lutzii* was used as a model to prove the functionality of this vector, but studies in other strains will be further developed using this collection of vectors.

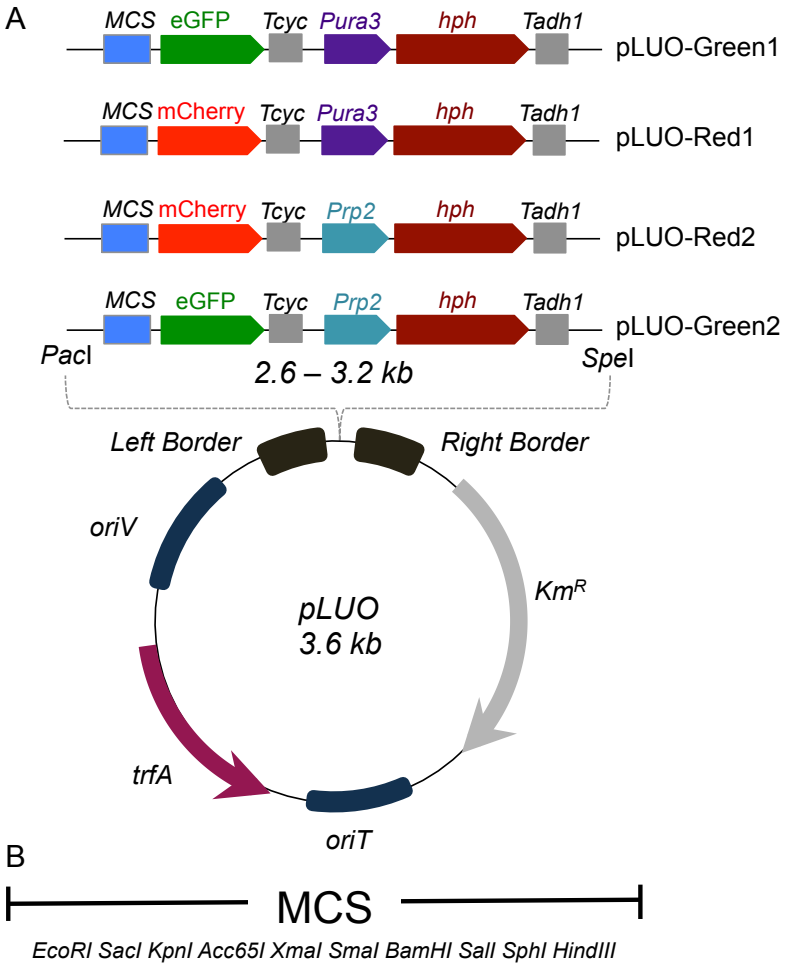

Figure 5. Representative scheme of the minimal binary vector for fungi transformation through ATMT. **A)** Design of the plasmid showing all the minimal modules that compose it, including the four versions of the expression *cassette* that were constructed. pLUO-Green1 (with eGFP) and pLUO-Red1 (with mCherry) have a *Pura3* promoter for *hph* expression, while pLUO-Green2 and pLUO-Red2 uses *Prp2* from *T. reesei*. **B)** Representation of the restriction enzymes available in the MCS for several cloning options.

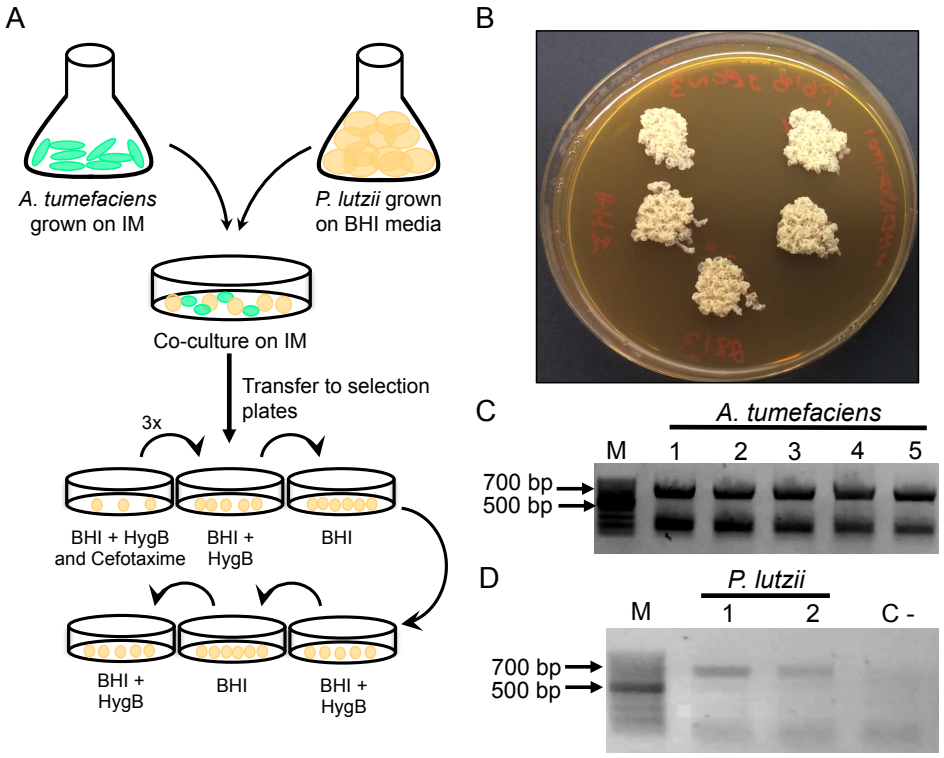

Figure 6. Validation of pLUO vector in *P. lutzii*. A) Representative image of experiment workflow. First selection plate has Cefotaxime 100 µg/mL to kill the remaining *A. tumefaciens* colonies. Afterwards, fungal colonies are selected on BHI containing only Hygromicin B for three rounds, and then alternating media with or without Hygromicyn until it reaches nine rounds of mitotic stability. **B)** *P. lutzii* colonies transformed with pLUO vector after reaching mitotic stability. **C)** Eletrophoresis gel showing the colony PCR reactions from *A. tumefaciens* as a positive control with gene *hph* amplified by its primers (resulting in a band of 705 base pairs; lower band is inespecific one that always appears for *A. tumefaciens*). **D)** Eletrophoresis gel showing the PCR reactions of *P. lutzii* transformants. M stands for molecular ladder and C – is the negative control of the reaction.

4. Conclusion

We presented a collection of minimalist binary vectors for transformation through ATMT applicable to several fungi species, which was validated in *P. lutzii*. pLUO vectors can be applied to several studies of single-cell gene expression because they consist of a reporter module containing either distinct fluorescent protein: mCherry or eGFP. Hence, we applied synthetic biology methods to design and develop an efficient and reliable molecular tool for fungal transformation: minimalist, synthetic, modular, and available in four different versions, which can still be readily modified using a few primers and few cloning steps.

CHAPTER TWO

A toolset of constitutive promoters for engineering *Rhodosporidium toruloides*

This chapter was submitted as:

Nora, L. C., Wehrs, M. Kim, J., Skerker, J.M., Simmons, B., Gladden, J.M., Silva-Rocha, R., Mukhopadhyay, A., Kirby, J. (2018). A toolset of constitutive promoters for engineering *Rhodosporidium toruloides*. *Submitted.*

1. Specific Background

With the aim of replacing many petrochemical-derived fuels and chemicals with renewable alternatives, the field of synthetic biology is branching out from model organisms such as *E. coli* and *S. cerevisiae* as host of choice. The oleaginous, carotenogenic yeast *R. toruloides* is emerging as a promising host for sustainable bioproduction due to its innate capacity to efficiently utilize a larger portion of lignocellulosic biomass. Not only is co-utilization of C5 and C6 sugars more efficient than in *E. coli* or *S. cerevisiae*, but *R. toruloides* also has the capacity to grow on aromatic compounds found in lignin hydrolysates, such as *p*-coumaric acid and ferulic acid (Yaegashi *et al.*, 2017). Much of the work on metabolic engineering or bioproduct formation in *R. toruloides* has focused on lipid accumulation (Zhang *et al.*, 2016; Wang *et al.*, 2018). However, the high lipid content suggests a suitability for applications to any acetyl-CoA-derived product, from fatty alcohols and methyl ketones to polyketides and terpenes, translating to applications ranging from fuels and specialized chemicals to complex drugs (Yuan and Ching, 2016; Campbell *et al.*, 2017b; Yaegashi *et al.*, 2017; Park *et al.*, 2018). The surface has only been scratched as far as metabolic engineering work is concerned, but results so far look promising. For example, Yaegashi *et al.* (2017) have reported moderately high-titer production of two terpenes, bisabolene and amorphadiene, following no modifications other than expression of the exogenous terpene synthases. Furthermore, although lignocellulosic hydrolysates often contain components that are inhibitory to microbial growth and bioproduct formation, terpene production in *R. toruloides* was, remarkably, enhanced in corn stover hydrolysate relative to a mock medium with similar sugar content. Hence, the increased attention that *R. toruloides* has received in recent years appears to be well justified.

In order to improve prospects for more advanced metabolic engineering in this host, the development of genetic tools and resources is an immediate priority. The direct application of tools from other host organisms is unlikely to be successful, or at least not a trivial task, in *R. toruloides*, a basidiomycete with a high chromosomal DNA GC-content that lacks episomal vectors (Johns *et al.*, 2016). Thus far, all genetic modifications and heterologous expression has been achieved by genomic integration, underscoring the need for a knowledge base on stable promoters for expressions of genes of interest. A central goal of synthetic biology is the development of standard

parts that are reliable, orthogonal and robust; and in this case that can be used in non-conventional fungal strains (Martins-Santana *et al.*, 2018; Nora *et al.*, 2018).

A small selection of promoters has already been characterized to modulate expression of heterologous genes in *R. toruloides* but the majority of them are metabolite-responsive promoters rather than constitutive. Liu et al. (2015) examined an amino acid responsive promoter, while Johns *et al*. (2016) characterized a set of four promoters that are modulated by copper, nitrate, acetate, and methionine (Johns *et al.*, 2016), and Liu et al. (2016) studied a strong intronic promoter (Liu *et al.*, 2016). Promoters that are highly responsive to certain metabolites are indeed valuable tools and they would be best complemeted with with a toolset of reasonably well characterized constitutive promoters. Given the number of possible pathways in *R. toruloides* that are attractive targets for metabolic engineering, there is a clear need for more biological parts for this host. In this context, functional bidirectional promoters would also be a valuable addition to the toolset, since they can be convenient for optimization of multi-gene pathways.

The most common method used for transformation of *R. toruloides* is *Agrobacterium tumefaciens*-mediated transformation (Liu *et al.*, 2013; Lin *et al.*, 2014), a useful strategy for genomic integration of transgenes but a laborious and time-consuming process. More recently, transformation by electroporation (Liu *et al.*, 2017) and lithium acetate methods (Tsai *et al.*, 2017a, 2017b) were reported. However, higher transformation rates are needed to facilitate high-throughput screens and site-specific recombination, which is a rare event in *R. toruloides*. In this context, we sought to establish the most convenient and efficient transformation method to streamline our workflow and expand the toolbox for metabolic engineering of *R. toruloides*.

Here, we performed RNA sequencing analysis to select promoters from *R. toruloides* that are likely to result in high- and medium-level constitutive expression. Several of these promoters were selected for bidirectionality. We characterized these promoters by cloning them in a dual-reporter system and monitored expression in different media over a cultivation time of seven days. Concurrently, we established an efficient protocol for the lithium acetate chemical transformation method to facilitate and improve the engineering workflow. Thus, we have expanded the collection of parts for metabolic engineering of the emergent microbial host, *R. toruloides*, and also improved the means by which these tools can be inserted into the genome. We present

a *R. toruloides* toolbox with the potential for application in a range of different studies.

2. Materials and Methods

2.1. Media and growth conditions

Synthetic defined (SD) medium was made following instructions of manufacturer with Difco yeast nitrogen base (YNB) without amino acids (Becton, Dickinson & Co., Sparks, MD) and complete supplemental mixture (CSM; Sunrise Science Products, San Diego, CA). Luria broth (LB) was made using Difco LB (Becton, Dickinson & Co., Sparks, MD) and yeast peptone dextrose (YPD) media were made using 2% peptone and 1% yeast extract (Becton, Dickinson & Co., Sparks, MD) and 2% dextrose. The xylose and dextrose used for the media containing sugars were from Sigma-Aldrich (St. Louis, MO) .

For the promoter studies, all 60 *R. toruloides* IFO0880 Δ*ku70* strains (58 different promoters, the wild-type and the CAR2 deletion white control with no promoter) were inoculated into 24-well plates containing 2mL of LB and were grown overnight at 30 ˚C in a shaker at 200 rpm. These cultures were then used to inoculate the sugar-containing media: SD 1% xylose, SD 1% glucose, SD 1% xylose 1% glucose and YPD in a 1/100 dilution and were grown overnight at 30 ˚C in a shaker at 200 rpm for cells to reach exponential phase. The OD_{600} was measured for these samples and the respective amount of culture was transferred to new media so the starting OD_{600} of all cultures was 0.05. The new inoculum was made in 24-well plates containing 3 mL of the same 4 different sugar-containing media and the cells were grown for 7 days at 30 ˚C in a shaker at 200 rpm with samples being taken every 8, 24, 48, 96 and 168 hours.

2.2. Strains and plasmid construction

R. toruloides strain IFO0880, which is a haploid strain, with a deletion for the *ku70* gene was used for transformations to favor homologous recombination. Strains and plasmids used in this study are available upon request through The Joint BioEnergy Institute Inventory of Composable Elements (JBEI-ICEs) at https://public-

registry.jbei.org/. Gene synthesis and plasmid construction were performed by Joint Genome Institute (JGI) and the constructs were designed so that each promoter was positioned between the fluorescent proteins GFP and mRuby, in both directions. Constructs also contain the Nourseothricin acetyltransferase (*nat1*) for resistance to Nourseothricin antibiotic and homology regions corresponding to the CAR2 locus. A schematic representation of the expression *cassette* built for integration into *R. toruloides* genome can be seen in **Figure 7** and selected promoters and their combined annotations can be seen in **Table 3**.

2.3. RNA-sequencing

RNA-sequencing was used to identify a set of promoters that drive high or medium expression in different media (synthetic defined medium containing 1% (w/v) glucose or xylose). High expression was defined as log2 of the mean FPKM (Fragments Per Kilobase of transcript per Million mapped reads) values above 10, and medium expression was defined as log2 of the mean FPKM values between 8 and 10. In order to select promoters with constitutive expression, 20 genes with the lowest variance across different growth conditions were selected from each high and medium expression gene. Subsequently, promoter regions of high and medium expression genes, spanning up to 1000 bp upstream of CDS when possible, were selected for testing. In some cases, the predicted promoter region of a selected gene was cut short by an upstream CDS on the opposite strand and was annotated as a putative bidirectional promoter.

2.4. *R. toruloides* transformation

Plasmids were retrieved from *E. coli* DH10B strains by plasmid miniprep (Qiagen) and digested with Thermo Scientific™ FastDigest *Pvu*II enzyme to cleave in the regions flanking the expression cassette (with exception of constructs 5, 14, 18, 34, 43 and 47 that were amplified by PCR using Thermo Scientific™ Phusion High Fidelity DNA Polymerase since they had *Pvu*II sites inside their promoter sequences). *R. toruloides* IFO0880 Δ*ku70* strain was inoculated into 10 mL of YPD and incubated overnight at 30 °C with shaking at 200 rpm. Then, the culture was diluted to an OD600 of 0.2 in YPD and incubated at 30 °C on a shaker at 200 rpm until it reached

an OD600 of 0.8. The cultures were harvested, centrifuged at 5000 rpm for 10 min, washed with water and then resuspended in 100 mM LiAc (pH 7.5). The cells were centrifuged again and the LiAc was removed. Next, the transformation mixture was added, consisting of PEG 4000 (50% w/v), 1.0M LiAc (pH 7.5), single-stranded Salmon Sperm DNA (10 mg/mL) and 5 µg of the transforming DNA. The mixture was vortexed and incubated at 30 °C for 30 min. Subsequently, DMSO was added and mixed thoroughly before the heat shock step at 42 °C for 15 minutes. Cells were then centrifuged and the supernatant was removed by pipetting. Cells were resuspended in YPD and were recovered overnight at 30˚C on a shaker at 200 rpm. The cells were plated in YPD agar plates containing 100µg/mL of Nourseothricin Sulfate (G-Biosciences®) and incubated at 30 ˚C for two days. The plates were then transferred to the 4 ˚C fridge for two more days to differentiate the white colonies from the orange ones. The white colonies were re-streaked on YPD agar plates and single-colonies were picked from 3 independent transformants for each promoter construct for the following experiments.

2.5. Flow cytometry

High-throughput flow cytometry experiments were performed using the Accuri C6 flow cytometer equipped with an autosampler (BD Accuri, Model number City, State). All the fluorescence measurements were performed in biological triplicates. A total of 30,000 events were recorded at a flow rate of 35 µL/min and a core size of 16-µm for each sample. Both fluorescent proteins were excited at 488 nm, GFP emission was detected at 530 nm and mRuby emission was detected at 675 nm. Data acquisition was performed as described in the Accuri C6 Sampler User's Guide. The acquired data was analyzed in the FlowJo® software (Becton, Dickinson & Company), where populations were gated and the median of all fluorescence curves were obtained.

2.6. Statistics

For the fluorescence values, the median value of each fluorescence curve of each replicate was extracted using FlowJo®. Then, the mean of the triplicates was calculated from the median values, as well as the standard deviation (STD) and the

coefficient of variation (CV). The CV of all promoters were calculated to be between 0 and 1. The resulting means were then divided by the mean fluorescence of the negative control Δ*car2* to normalize all samples in relation to background fluorescence. The background flourescence for GFP and mRuby were considered the ones measured in the respective channel in the flow cytometer. The resulting normalized values were then transformed to Log2 values represented in the figures as Relative Fluorescence Units (RFU). Heatmaps were done using Multiple Experiment Viewer (MeV) where promoters were clusetered by hierarchical clustering using Euclidean distance as the distance metric selection through average linkage clustering. The remaining graphs were created using GraphPad Prism® (GraphPad Software, San Diego, CA), also used to calculate linear regression when needed.

2.7. Growth rate experiments

For growth rate experiments, strains were grown in the same conditions as the fluorescent measurement experiments. Samples were taken at 0, 8, 24, 48, 96 and 168 hours after inoculation. OD_{600} was monitored using a SpectraMax Plus 384 spectrophotometer from Molecular Devices (San Jose, CA).

2.8. High performance liquid chromatography (HPLC) analysis

Sugar and metabolite concentrations were quantified on a 1200 series HPLC (Agilent Technologies) equipped with an Aminex H column. Samples were filtered through 0.45 μm filters (VWR) to remove cells, and 5 μl of each sample was injected onto the column, preheated to 50◦C. The column was eluted with 4 mM H_2SO_4. Sugars and metabolites were monitored by a refractive index detector, and concentrations were calculated by peak area comparison to known standards.

Table 3. Putative promoter sequences investigated in this study. Promoters were selected from RNA-sequencing data and are listed with their locations in *R. toruloides* genome, the links to their respective constructs in the JBEI registry, and the combined annotations with putative functions of the genes that each promoter modulates. Protein ID and Transcript ID refers to the *Rhodosporidium toruloides* IFO0880 v4.0 genome that can be found in https://genome.jgi.doe.gov/Rhoto_IFO0880_4/Rhoto_IFO0880_4.home.html. *2p* means that the promoter was predicted to be bidirectional with the reverse gene annotation.

Promoter	Location	Registry ID	RV		Protein Id	Transcript Id	Combined Annotations
1	RTO3_9 59308-915227	JBx_081788	JBx_081872	target	10325	10453	K02984: RP-S3Ae, RPS3A; small subunit ribosomal protein S3Ae
				2p	10324	10452	BLAST: 40S ribosomal protein S1 [Exophiala sideris]
2	RTO3_9 08567-908568	JBx_081790	JBx_081874	target	15909	16037	K04564: SOD2; superoxide dismutase, Fe-Mn family
				2p	15908	16036	K01012: bioB; biotin synthase
3	RTO3_9 17545	JBx_081792	JBx_081876	target	11021	11149	K15040: VDAC2; voltage-dependent anion channel protein 2
4	RTO3_8 80855	JBx_081794	JBx_081878	target	14514	14642	K02925: RP-L3e, RPL3; large subunit ribosomal protein L3e
5	RTO3_9 14310	JBx_081796	JBx_081880	target	10055	10183	K02969: RP-S20e, RPS20; small subunit ribosomal protein S20e
6	RTO3_9 06435	JBx_081798	JBx_081882	target	15265	15393	DUF
7	GAPDH	JBx_081800	JBx_081884	target	10613	10741	K00134: GAPDH, gapA; glyceraldehyde 3-phosphate dehydrogenase
8	RTO3_8 93291-893294	JBx_081802	JBx_081886	target	8752	8880	KOG0004: Ubiquitin/40S ribosomal protein S27a fusion
				2p	8751	8879	K02922: RP-L37e, RPL37; large subunit ribosomal protein L37e
9	RTO3_8 94745-894744	JBx_081804	JBx_081888	target	9231	9359	K11254: H4; histone H4

				2p	9232	9360	K11253: H3; histone H3
10	RTO3_899252	JBx_081806	JBx_081890	target	13099	13227	K08770: UBC; ubiquitin C
11	RTO3_903413-903417_short	JBx_081808	JBx_081892	target	14295	14423	K02979: RP-S28e, RPS28; small subunit ribosomal protein S28e
				2p	14294	14422	K02989: RP-S5e, RPS5; small subunit ribosomal protein S5e
12	RTO3_903413-903417_long	JBx_081810	JBx_081894	target	14295	14423	K02979: RP-S28e, RPS28; small subunit ribosomal protein S28e
				2p	14294	14422	K02989: RP-S5e, RPS5; small subunit ribosomal protein S5e
13	RTO3_953517-953518	JBx_081812	JBx_081896	target	9006	9134	K02947: RP-S10e, RPS10; small subunit ribosomal protein S10e
				2p	9007	9135	K18027: PTPN3, PTPH1; tyrosine-protein phosphatase non-receptor type 3
14	RTO3_897758	JBx_081814	JBx_081898	target	12693	12821	K03231: EEF1A; elongation factor 1-alpha
15	RTO3_897781	JBx_081818	JBx_081900	target	12704	12832	K05863: SLC25A4S, ANT; solute carrier family 25 (mitochondrial adenine nucleotide translocator), member 4/5/6/31
16	RTO3_956621	JBx_081826	JBx_081902	target	14937	15065	K02978: RP-S27e, RPS27; small subunit ribosomal protein S27e
17	RTO3_947134	JBx_081832	JBx_081904	target	15825	15953	0
18	RTO3_918680-918676	JBx_081840	JBx_081906	target	11331	11459	K01647: CS, gltA; citrate synthase
19	RTO3_910373-957980	JBx_081846	JBx_081908	target	16419	16547	K02958: RP-S15e, RPS15; small subunit ribosomal protein S15e
				2p	16418	16546	K02943: RP-LP2, RPLP2; large subunit ribosomal

							protein LP2
20	876864_upstream_1kbp	JBx_081852	JBx_081910	target	8813	8941	K03768: PPIB, ppiB; peptidyl-prolyl cis-trans isomerase B (cyclophilin B)
21	894205-894206_intergenic	JBx_081854	JBx_081912	target	9048	9176	K03094: SKP1, CBF3D; S-phase kinase-associated protein 1
				2p	9047	9175	0
22	393806-896133_intergenic	JBx_081856	JBx_081914	target	12216	12344	KOG1817: Ribonuclease
22				2p	12215	12343	0
23	864008_upstream_573bp	JBx_081858	JBx_081916	target	12844	12972	HMMPfam:Domain of unknown function (DUF543):PF04418
24	954791-898916_intergenic	JBx_081860	JBx_081919	target	13007	13135	KOG3469: Cytochrome c oxidase, subunit VIa/COX13
				2p	13006	13134	HMMPfam:GDSL/SGNH-like Acyl-Esterase family found in Pmr5 and Cas1p:PF13839,SUPERFAMILY::SSF52266
25	528078-955376_intergenic	JBx_081862	JBx_081921	target	13608	13736	HMMPfam:Ubiquinol-cytochrome-c reductase complex subunit (QCR10):PF09796
				2p	13607	13735	K10859: ALKBH2; alpha-ketoglutarate-dependent dioxygenase alkB homolog 2
26	901123_upstream_225bp	JBx_081864	JBx_081923	target	13614	13742	K00411: UQCRFS1, RIP1, petA; ubiquinol-cytochrome c reductase iron-sulfur subunit
27	956312_upstream	JBx_081866	JBx_081925	target	14603	14731	K17279: REEP5_6; receptor expression-enhancing protein 5/6

28	907093-957091_intergenic	JBx_081868	JBx_081927	target	15467	15595	K12845: SNU13, NHP2L; U4/U6 small nuclear ribonucleoprotein SNU13
				2p	15466	15594	KOG2633: Hismacro and SEC14 domain-containing proteins
29	939244-882334_intergenic	JBx_081870	JBx_081929	target	15484	15612	K04393: CDC42; cell division control protein 42
				2p	15485	15613	K12868: SYF2; pre-mRNA-splicing factor SYF2

3. Results

3.1. Creation of promoter library

Promoters were selected from RNA-sequencing data based on their level of expression; high and medium expression genes were classified as described in the Methods section. All of the selected promoters are considered to be constitutive, based on variance in the RNA-sequencing data. Each promoter fragment was cloned into the reporter construct in both orientations, with the *forward* orientation driving GFP expression and the *reverse* orientation driving mRuby (**Figure 7**). Some of the promoter fragments comprise a complete intergenic sequence (i.e., between start codons of two neighboring divergent genes) and were therefore identified as putative bidirectional promoters, denoted here as *2p* promoters. For *2p* promoters, fluorescence for both reporter genes was measured in each of the two construct orientations using the BD Accuri Flowcytometer. In total, 29 promoters were selected and, from those, 13 were predicted to be bidirectional. A strain containing the reporter gene expression cassette lacking a promoter fragment was used as negative control, referred to here as Δ*car2*. Therefore, in total, 59 constructs were transformed into *R. toruloides* genome, 29 *forward* constructs, facing GFP, 29 *reverse* constructs, facing mRuby, and one control with no promoter.

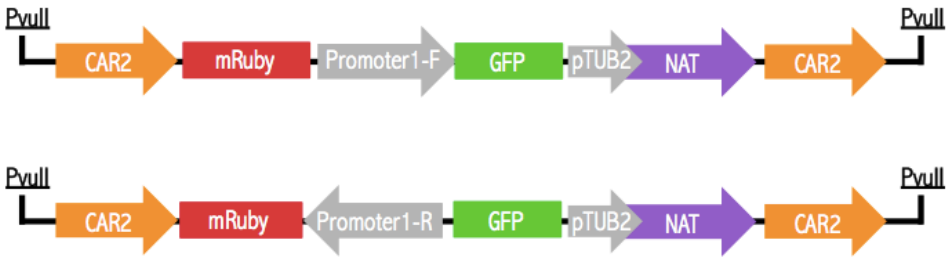

Figure 7. Schematic representation of the expression cassettes to be inserted into the *R. toruloides* genome. Top: representative design of the forward constructs, with promoter of coding sequence facing GFP. Bottom: representative design of the reverse constructs, with promoter of coding sequence facing mRuby. Promoters size range from 200 bp to 1 kb. Bidirectional promoters (*2p* promoters) are supposed to express both fluorescent proteins independently of the direction the promoter region is facing. Sequences homologous to the CAR2 locus are employed to facilitate integration of the cassette into the *R. toruloides* chromosome. *Nat*: resistance to the Nourseothricin antibiotic modulated by the constitutive promoter of the *Tub2* gene. PvuII sites used to enzymatically digest the cassette to remove it from the plasmid are shown. All genes contain terminators although they are not shown in this schematic representation.

The workflow from construction of expression constructs to promoter analysis by flow cytometry is shown in **Figure 8**. The parent strain for all work described here is *R. toruloides* IFO0880 harboring a deletion for the non-homologous end joining (NHEJ) gene, *KU70*. The Δ*ku70* was used in all experiments to favor homologous recombination and single-locus integration. Each construct was flanked by 1 kb sequences homologous to the *CAR2* locus (**Figure 7**), to facilitate site-specific integration. Successful site-specific integration events resulted in white colonies due to deletion of CAR2 (**Figure 8A**), which encodes a bifunctional phytoene/lycopene synthase (Liu *et al.,* 2015). For each construct, several white transformants were grown overnight in YPD and tested for fluorescence by flow cytometry (**Figure 8A**).

Three confirmed transformants for each construct were inoculated into 24-deep-well plates containing LB medium, grown overnight, and then subcultured into the various test media at a starting OD of 0.05. One rich medium (YPD) and three defined media (SD supplemented to 1% (w/v) with glucose, xylose, both) were selected to represent typical conditions employed to culture *R. toruloides*. Samples were collected at 8, 24, 48, 96 and 168 hours for measurement of fluorescence by the flow cytometry. See **Figure 8B** for an example of promoter 9 (a strong bidirectional promoter). Expression of both GFP and mRuby were measured at the same time using different channels in the flow cytometer (see **Figure 8C**, still using promoter 9 as an example).

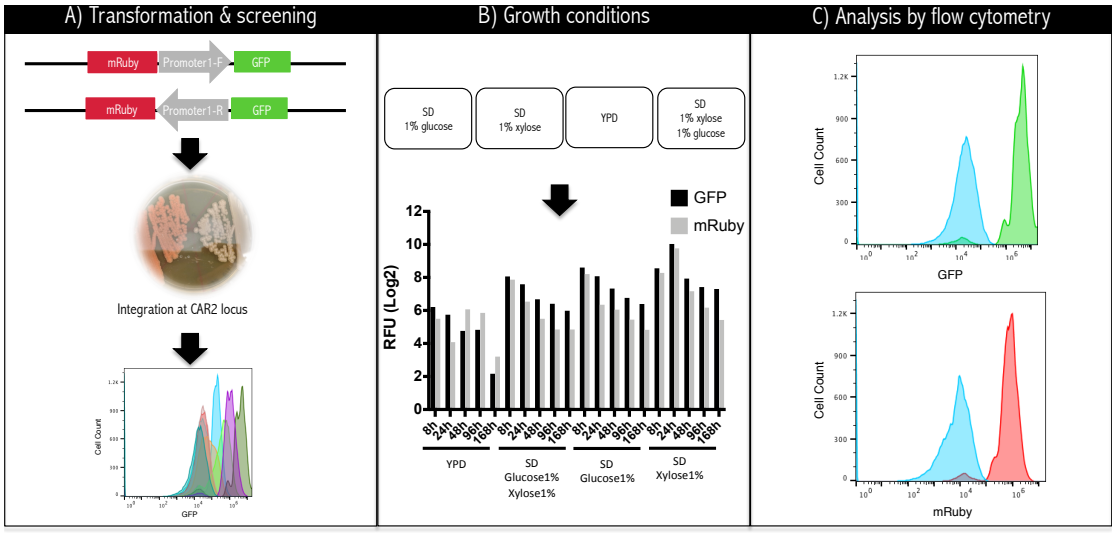

Figure 8. Experimental workflow. (A) Constructs were transformed into *R. toruloides* genome using an optimized Lithium Acetate transformation protocol. Positive integrations were determined visually and confirmed by fluorescence measurements. Histograms show different profiles of fluorescence for several forward constructs grown in SD Glucose 1%. **(B)** The final 59 different strains were cultivated in 24 deep well plates in four different media compositions: SD Glucose 1%, SD Xylose 1%, YPD and SD Glucose 1% Xylose 1%. Cultures were grown during a time course of 7 days. Samples were collected in the following timepoints: 8, 24, 48, 96 and 168 hours for measurement in the flow cytometry. A representative graph for GFP and mRuby expression in RFU (Relative Fluorescence Units) for all the time points for all the media tested is shown. **(C)** Expression of GFP and mRuby were measured at the same time using different channels in the flow cytometer. The upper histrogram represents expression profile of GFP of promoter 9 compared with the negative control Δ*car2*. The lower histrogram represents expression profile of mRuby of promoter 9 compared with the negative control Δ*car2*. These histograms are absolute values from 24 hours of growth in SD Glucose 1% Xylose 1%.

Flow cytometry is a useful method to analyze expression of different fluorescent proteins over time and on a single cell level. For *R. toruloides,* is especially important since the organism grows poorly in 96-well plates (data not shown), and our method allowed us to use 24 deep well plates for all cultures and then transfer a little amount of culture for flow cytometry methods. Still, for all the media and all conditions tested, *R. toruloides* presented a second smaller population in the earlier time points of 8 hours and 24 hours – meaning it is a growth-related condition. However, this population disappears in the later time points.

We performed confocal microscopy with cells of all conditions and several time points and we have not found any contamination (data not shown), which means this is a very specific condition of *R. toruloides*. Therefore, we gated to the right–sized

population for all the conditions and all time points and collected 30.000 events for this gate. All the results shown in the paper are from the gated populations.

Growth and sugar consumption were monitored and representative data (for the parental strain, *R. toruloides* IFO0880 Δ*ku70* and a strain harboring reporter for promoter 14) can be found in **Figure 9** and **10**, respectively.

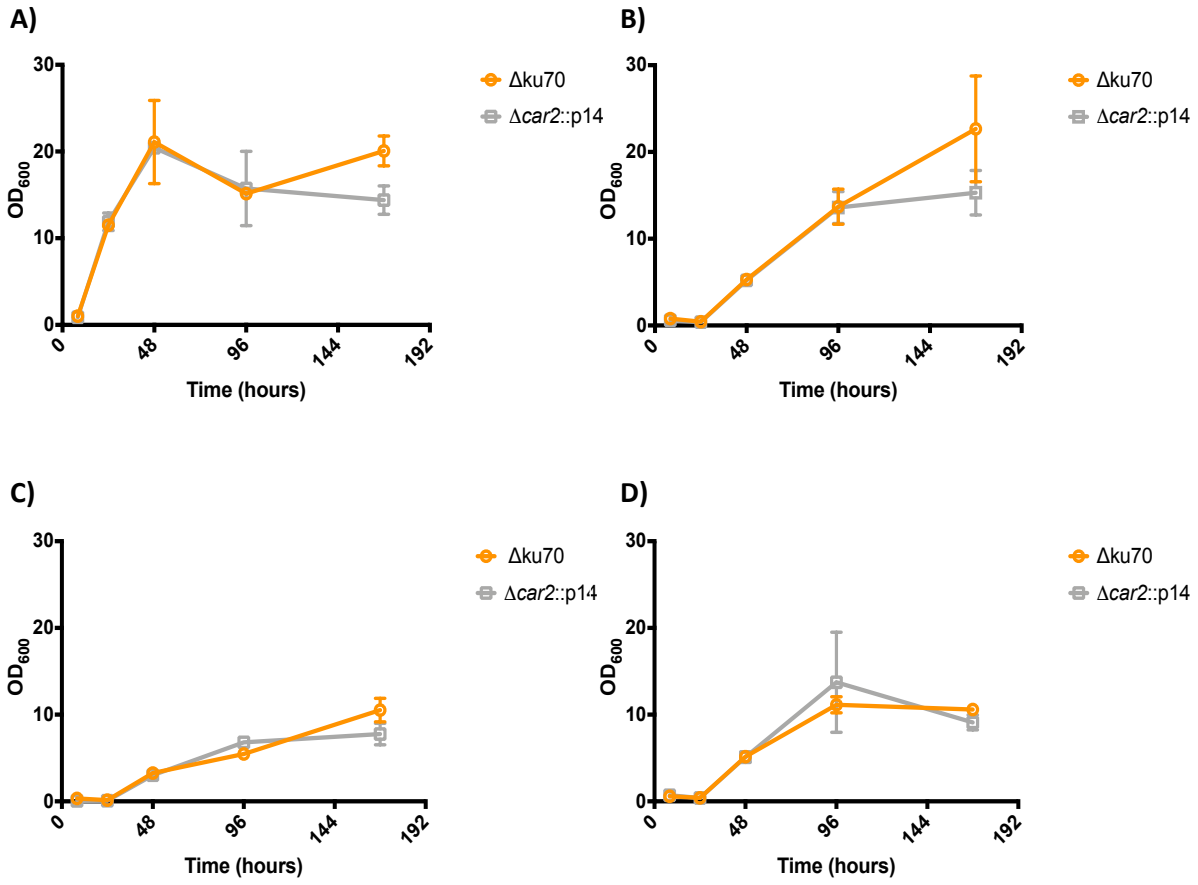

Figure 9. Growth curves for *R. toruloides*. Representation of growth curves for both parental strain *R. toruloides* Δ*ku70* and mutant strain with promoter 14 integrated in the genome. OD600 was measured in spectophotometer in triplicates for all the strains at 8, 24, 48, 96 and 168 hours. **(A)** Growth in YPD. **(B)** Growth in SD glucose 1% and xylose 1%. **(C)** Growth in Glucose 1%. **(D)** Growth in Xylose 1%.

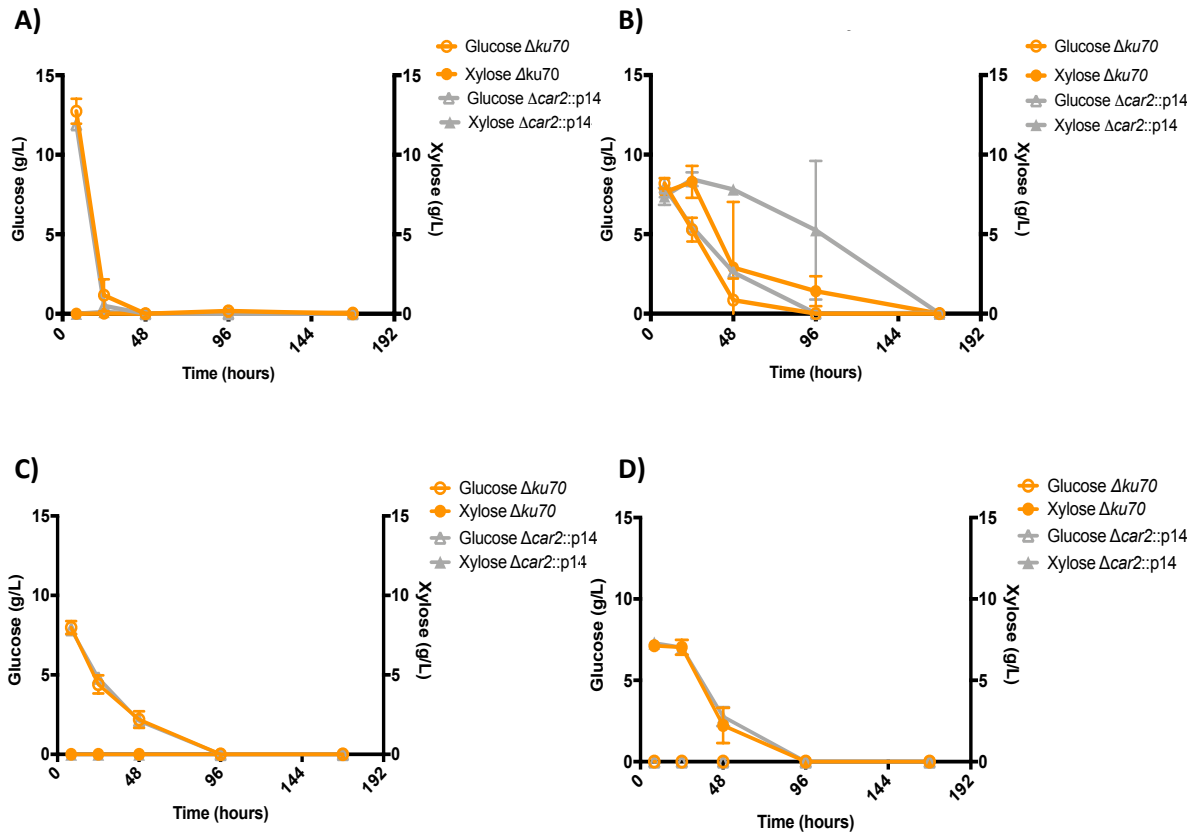

Figure 10. Sugar consumption. Comparison of glucose and xylose consumption from parental strain *R. toruloides* Δ*ku70* and mutant strain with promoter 14 integrated in the genome. Prevalence of sugars in each media was measured using HPL in triplicates for both strains and samples were collected at 8, 24, 48, 96 and 168 hours. Glucose concentration can be seen in left axis and xylose concentration is explicited in the right axis. **(A)** Sugar consumption in YPD. **(B)** Sugar consumption in SD glucose 1% and xylose 1%. **(C)** Sugar consumption in Glucose 1%. **(D)** Growth in Xylose 1%.

3.2. Expression of promoters facing GFP

In **Figure 11**, a heat-map with all the forward constructs is shown. Clusterization through similarity show that the three strongest promoters (promoters 9, 12 and 19) are bidirectional and express GFP and mRuby at similar rates. Promoter 9 was shown to be the strongest promoter of all set. We can also observe that promoters 21, 22, 28 and 13 were clustered as medium bidirectional promoters. Promoters 10, 18, 27, 4, 15, 17 and 14 are strong monodirectional promoters, with promoter 14 being the strongest one of this specific set. Promoters 7, 5, 1, 3 and 6 were clustered as medium monodirectional promoters. It is interesting to note that promoter number 7 is the common benchmark promoter for the GAPDH gene and

through this clusterization it is classified as medium expression. Promoters 2, 25, 23, 16, 24, 20, 26 and 11 have not shown significant expression and thus were considered weak promoters in our data set. Promoter 8 is stronger at the opposite side in which is the promoter construct is faced (which means that, although it is facing GFP in this data set, it has shown a higher expression of the mRuby fluorescent protein). Likewise, promoter 29 is also a medium bidirectional promoter such as promoters 21, 22, 28 and 13, but it was clustered separately because it shows higher expression of mRuby in relation to GFP.

3.3. Expression of promoters facing mRuby

It is possible to obseve in **Figure 12** the expression levels of the same promoters mentioned above but that were built facing the mRuby protein, what we call promoters-R or reverse promoters. The clusterization of the promoters facing the opposite direction showed that promoters behave somehow differently when they are built in the reverse direction. Still, once again, promoters 9, 19 and 12 were considered the strongest bidirectional ones. Promoter 16, that had not shown expression in the forward direction, behaved as a medium bidirectional one when in reverse position. Promoter 22 was classified as medium bidirectional promoter in both analysis and when constructed in both directions. Promoters 1, 11, 29, 25, 24, 2, 26 and 20 have not presented significative expression and were considered weak promoters. Promoter 8 was stronger for mRuby expression in the forward construct but did not show expression for neither protein when facing mRuby. Promoters 23, 6, 3 and 21 were classified as medium monodirectional promoters. Furthermore, promoters 28, 4, 13, 7, 18, 5, 27, 10, 17, 15 and 14 were clustered as strong monodirectional promoters 21, 3, 6, 10, 27, 5, 18, 7, 4, 28, 17, 15 and 14 were considered strong monodirectional promoters.

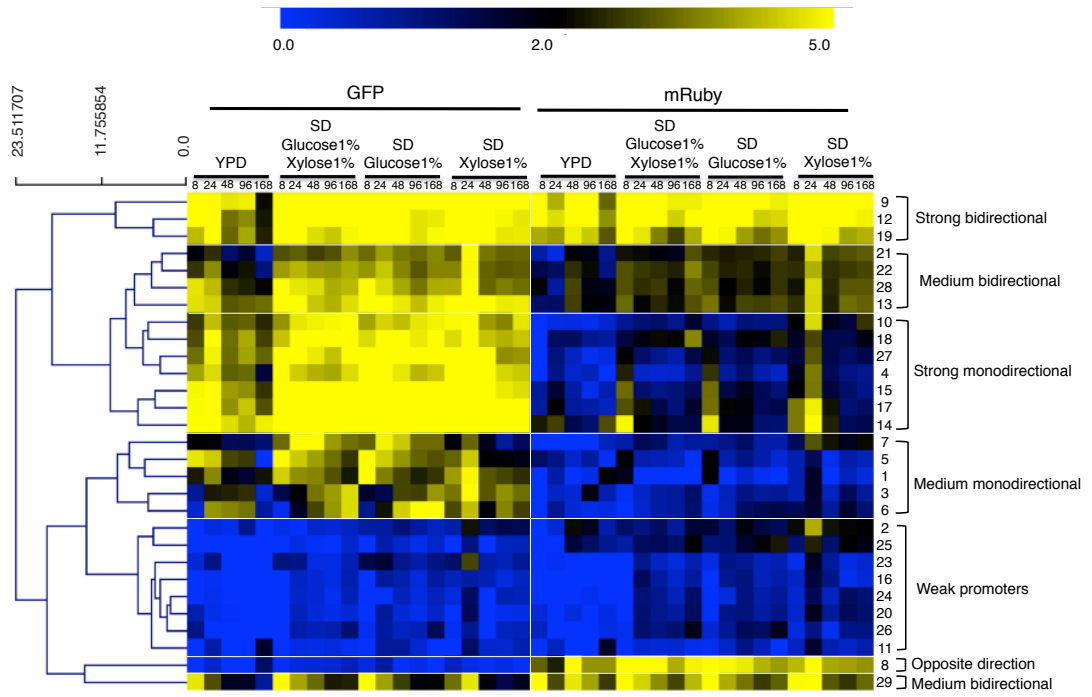

Figure 11. Heatmap for promoters facing GFP. The heatmap was made using MeV and promoters were clustered by hierarchical clustering using Euclidean distance. Expression values are on a Log2 scale, and were calculated as described in Methods. All time points for all media are shown for both GFP and mRuby. A rough guide to the characteristics of each group of promoters is shown on the far right side.

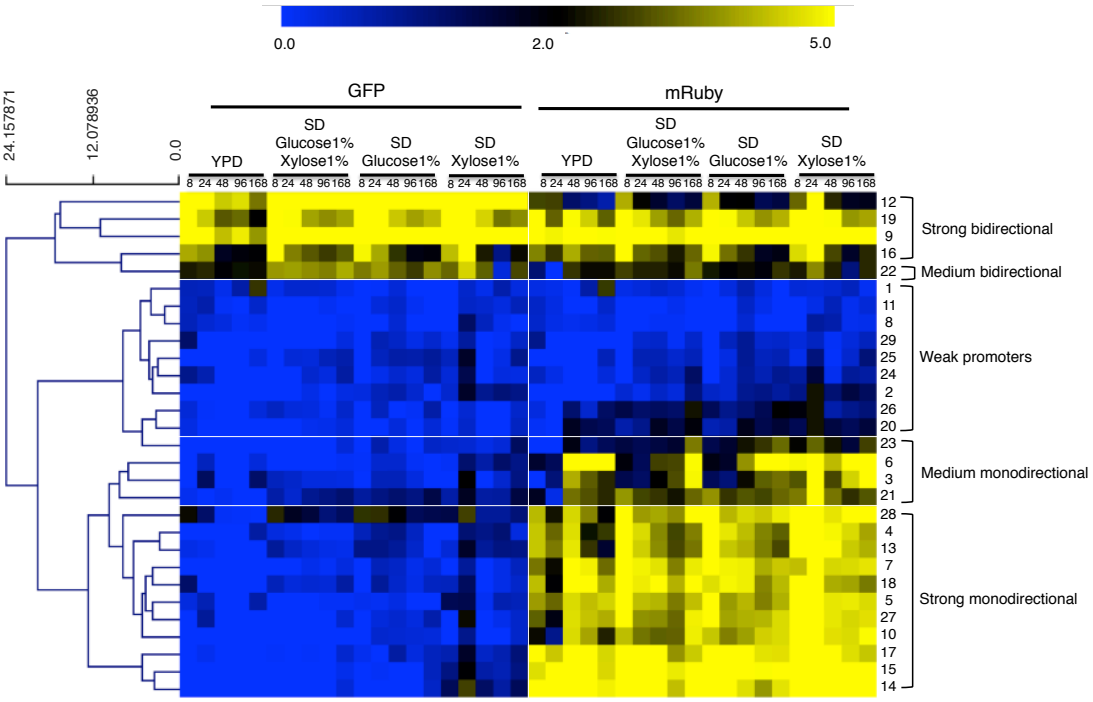

Figure 12. Heatmap for promoters facing mRuby. The heatmap was made using MeV and promoters were clusterd by hierarchical clustering using Euclidean distance. Expression values are in Log2 and were calculated as described in Methods. All time points for all media are shown for both GFP and mRuby. Characteristics of each group of promoters is shown on the far right side.

3.4. Correlation of mRuby and GFP expression and predictability of promoters

In order to check if the dual-reporter method is efficient to report expression of bidirectional promoters we performed correlation analysis between GFP and mRuby expression. In **Figure 13A**, correlation between the strongest bidirectional promoters of all set is shown: 9, 12, 19 and 29 grown for 24 hours in SD media with both glucose and xylose as carbon sources. The analysis resulted in a R^2 of 0.8756 for a P value of 0.0643. In **Figure 13B** the correlation of GFP and mRuby expression values in Log2 is shown for bidirectional promoter 9 for the all the conditions tested in all 5 time points. The linear regression analysis resulted in a R^2 of 0.6942 for a P value<0.0001.

Further, we verified if the amount of transcripts detected in the RNA sequencing experiments actually correlates with the expression values for its respective promoters (**Figure 14**). The linear regression analysis contains only promoters that were predicted to strong by RNA sequencing analysis and were also strong dataset. This analysis resulted in a R^2 of 0.4080 for a P value of 0.0058.

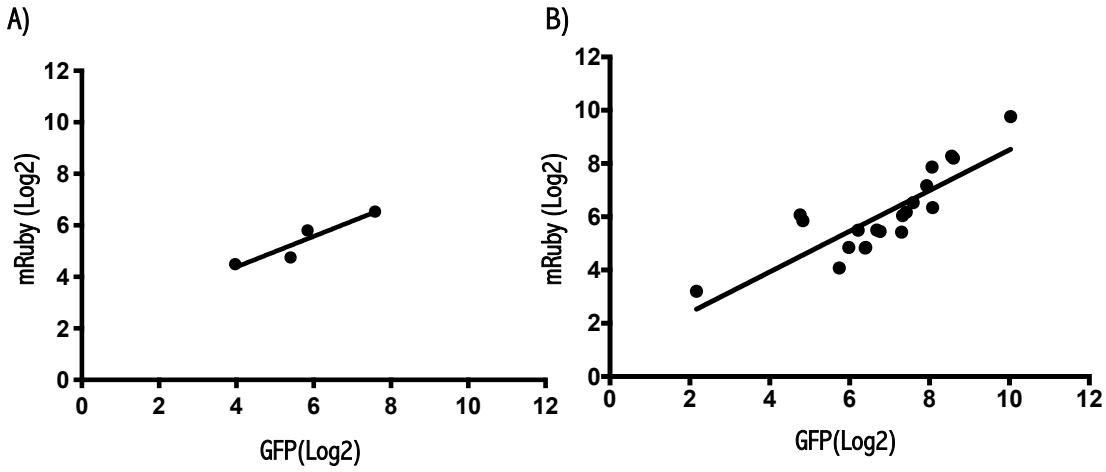

Figure 13. Correlation between GFP and mRuby expression suggest that the dual-reporter method is efficient to demonstrate bidirectionality of promoters in the same experiment. (A) Correlation between GFP and mRuby expression values in Log2 for the bidirectional promoters 9, 12, 19 and 29 at 24 hours in the constructs that were facing GFP grown in SD Glucose 1% Xylose 1%. R^2 = 0.8756 for a P value of 0.0643. **(B)** Correlation of GFP and mRuby expression values in Log2 for bidirectional promoter 9 for the 4 media conditions in all 5 time points. R^2 = 0.6942 for a P value<0.0001.

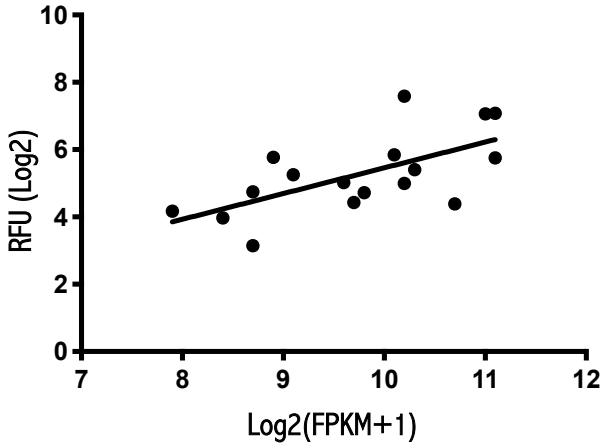

Figure 14. Predictability of promoter expression. Correlation between the average expression values of the RNA sequencing datasets and the average RFU values of the strongest promoters selected from our set. The promoters chosen were: 4, 5, 7, 9, 10, 12, 13, 14, 15, 7, 18, 19, 21, 22, 27, 28 and 29. Promoters that were predicted to be strong but were actually weak in our dataset were excluded from this analysis. R^2=0.4080 for a P value of 0.0058.

4. Discussion

R. toruloides is emerging as a promising host for the production of bioproducts through metabolic engineering. A key requisite to enable pathway engineering strategies is the development of genetic tools and resources for this organism. A set of characterized promoters is an important basic tool to facilitate synthetic biology in a new host, and a step towards the development of standard and orthogonal parts. From our transcriptomic data, a set of 29 different *R. toruloides* promoters was selected for characterization. Since most of them were predicted to be bidirectional, we used a dual-reporter system in order to test for directionality. The promoter-reporter constructs were designed for site-specific integration and transformed into *R. toruloides* using an efficient lithium acetate method protocol that was established in this study. Flow cytometry was used to measure promoter strength over time and in response to relevant carbon sources. The resulting data describes a set of native *R. toruloides* promoters, characterized over time in several useful conditions, which are expected to be of value to future metabolic engineering projects in *R. toruloides*. Robustness of the promoters is considered, both from the perspective of native

transcript abundance versus reporter fluorescence, and comparison of fluorescence from each of the two reporters.

In our dataset, all promoters show higher expression in the media containing glucose and xylose as carbon source, even though they grow much faster in YPD (**Figure 9**). The growth curves are important to show to different growth rates that the organism achieve depending on the carbon source. Comparing the growth curves with expression profiles in the heatmap it becomes clear that promoter strength is related to growth rates**.** Thus, it is not possible to compare the time points for different media, since the organism grow at different speed depending which carbon source it is consuming. But we still can observe kinetics of expression over time for each different media – and choose which promoter would be more suitable depending on the medium of interest for any given study.

For the strong monodirectional promoters, for example, the majority of them have a peak of high expression in the beginning of the cultures and then start to decrease throughout the culture – or at least maintain a lower but stable expression over time (**Figures 10 and 11**). They are interesting for those whose bioproduct needs to be expressed right in the start of the fermentation process. Interestingly enough, promoter 3 and 6 both showed an increase over time in all the cultures. Those two are of much interest for who is looking for bioproducts that need to be processed in the end of the cultivation time (e.g. products that are toxic for growth or that the organism itself can consume decreasing the productivity). These two promoters will be useful for production of bisabolene – the aforementioned precursor of biodiesel.

Besides the monodirectional promoters, we also found natural bidirectional promoters present in *R. toruloides* genome. Bidirectional promoters are important since they offer the ability to improve the design of pathways and gene co-expression (Vogl *et al.*, 2018). The need and applicability of bidirectional promoters is increasing for several organisms, including in some emergent bioconversion hosts (Neil *et al.*, 2009; Xu *et al.*, 2009; Vogl, *et al.*, 2018). From the 14 promoters that were predicted to be bidirectional in the RNA-sequencing data (1, 2, 8, 9, 11, 12, 13, 19, 21, 22, 24, 25, 28 and 29) 7 had proven to have bidirectional expression properties, and promoter 16 showed a bidirectional tendency, even though it was not predicted to. These type of promoters are valuable tools, especially because they were selected from naturally-occurring genes from the fungus genome.

Promoters 9, 12 and 19 are not only bidirectional, but also are the strongest promoters of the set – and in both directions. They could be useful for well-tuned co-expression of functional components for efficient synthetic biology endeavors. Promoter 9, predicted to modulate Histone 3 gene in one direction and Histone 4 in the opposite direction (**Table 3**), was shown to be the strongest promoter of all set. Interestingly enough, most bidirectional promoters described in literature came from histone genes (Vogl *et al.*, 2018). Promoter 12 was predicted to modulate small subunit ribosomal protein S28e in the forward direction and small subunit ribosomal protein S5e in the reverse direction. Yet, promoter 19 was predicted to modulate the expression of small subunit ribosomal protein S15e in the forward direction and large subunit ribosomal protein LP2 in the opposite direction.

A fair amount of promoters behave differently when they were constructed facing mRuby when compared to the constructs facing GFP. As an example is promoter 8, which presented a high expression of mRuby in the strain containing the promoter in the forward direction – facing GFP – but did not present any expression when built in the opposite direction. This could suggest a composability effect, in which the promoter behaves differently depending on the position of regulatory elements and where sequence composition can determine the levels of gene expression (Kosuri *et al.*, 2013; Mutalik *et al.*, 2013). Another suggestion is that mRNA secondary structure surrounding the ATG is unfavorable for one of the reporter proteins (Goodman *et al.*, 2013).

Promoters 10, 18, 27, 4, 15, 17 and 14 are the ones that were classified as strong promoters in both analysis, which means they have the same properties independently of the direction they are facing, and are reliable promoters to be used in metabolic engineering studies.

Since the absolute expression of GFP and mRuby cannot be compared, we performed a correlation analysis using relative values of expression for both proteins (**Figure 13**). All analysis presented signifcant deviations from zero, which means that the methods we applied in our work are valid: from using dual-reporter systems to identify bidirectional promoters to using RNA sequencing to identify candidate promoters to build a toolkit. The correlation show that these bidirectional promoters can be used in metabolic engineering studies for optimization of multi-gene co-expression.

One of the greatest findings of our promoter library were the 13 promoters that are stronger than the GAPDH promoter (GAPDH is described as promoter 7 in the heatmaps, **Figures 11** and **12**). P*gapdh* is one of the most used promoters for a variety of yeasts, including *R. toruloides*, as can be seen in the work by Yaegashi *et al.* (2017). However, as shown in the heatmaps, GAPDH is far from being one of the strongest promoters of the set. Thus, our promoter library can now be used for developing more accurate metabolic engineering studies and achieving higher yields and titers for the bioproducts of interest.

5. Conclusion

The collection of promoters presented in this work is the largest set published for *R. toruloides*. We presented the first bidirectional promoters specific from the *R. toruloides* genome. Additionally, from the seven bidirectional promoters that modulated expression of the fluorescent proteins, three were the strongest of all set, which will be a great benefit for future pathway engineering studies. We also found 13 promoters that are stronger than the benchmark promoter GAPDH, the most commonly used promoter for yeasts. That means that is possible our promoter toolbox can generate more accurate and reliable metabolic engineering studies for this emergent bioconversion host. Thus, we efficiently used synthetic biology methods to create a library of promoters with distinct characteristics to be applied in *R. toruloides*. Even more, we established a lithium acetate transformation method to facilitate genetic manipulation of this yeast and we demonstrated that RNA sequencing is a useful method to select genetic parts from the organism own genome.

IV. GENERAL CONCLUSIONS

The results of this dissertation have given rise to the following conclusions:

1 – Synthetic Biology was successfully applied to build collection of minimalist, synthetic and modular binary vectors for transformation through *Agrobacterium tumefaciens*-mediated transformation, which could be applicable to several distinct fungi species. This includes pathogenic species such as *Paracoccidioides lutzii*, for which common tools were usually larger than 10 kb.

2 – Synthetic Biology was also effectively applied to build a promoter library to the emergent bioconversion host *Rhodosporidium toruloides*. This library is an unprecedented collection of twenty-nine promoters – the largest ever published for this yeast – also consisting of some bidirectional promoters that will facilitate molecular cloning for the purpose of metabolic engineering in this fungal strain.

Altogether, the current work has provided resourceful information regarding the use of Synthetic Biology in such important Eukaryotic organisms. We believe these tools are of timely value not only for understanding the underlying molecular mechanisms of infection from pathogenic organisms but also to understand general principles of gene expression and extrapolate that for metabolic engineering of organisms for biotechnological applications.

V. REFERENCES

Abe, A., Elegado, E.B., and Sone, T. (2006) Construction of pDESTR, a Gateway vector for gene disruption in filamentous fungi. *Curr. Microbiol.* **52**: 210–215.

Abuodeh, R.O., Orbach, M.J., Mandel, M.A., Das, A., and Galgiani, J.N. Genetic Transformation of Coccidioides immitis Facilitated by *Agrobacterium tumefaciens.*

Alberti, S., Gitler, A.D., and Lindquist, S. (2007) A suite of Gateway cloning vectors for high-throughput genetic analysis in *Saccharomyces cerevisiae*. *Yeast* **24**: 913–919.

Almeida, A.J., Carmona, J.A., Cunha, C., Carvalho, A., Rappleye, C.A., Goldman, W.E., et al. (2007) Towards a molecular genetic system for the pathogenic fungus *Paracoccidioides brasiliensis*. *Fungal Genet. Biol.* **44**: 1387–1398.

Almeida, F.B. dos R., Oliveira, L.L. De, Sousa, M.V. De, Barreira, M.C.R., and Hanna, E.S. (2010) Paracoccin from *Paracoccidioides brasiliensis* ; purification through affinity with chitin and identification of N -acetyl- β - D - glucosaminidase activity. *Yeast* **27**: 67–76.

Amores, G., Guazzaroni, M.-E., Arruda, L., and Silva- Rocha, R. (2016) Recent Progress on Systems and Synthetic Biology Approaches to Engineer Fungi As Microbial Cell Factories. *Curr. Genomics* **17**: 85–98.

Andrianantoandro, E., Subhayu, B., Karing, D.K., Weiss, R., Basu, S., Karig, D.K., and Weiss, R. (2006) Synthetic biology: new engineering rules for an emerging discipline. *Mol. Syst. Biol.* **2**: 2006.0028.

Arentshorst, M., Lagendijk, E.L., and Ram, A.F. (2015) A new vector for efficient gene targeting to the pyrG locus in *Aspergillus niger*. *Fungal Biol. Biotechnol.* **2**: 2.

Avalos, J.L., Fink, G.R., and Stephanopoulos, G. (2013) Compartmentalization of metabolic pathways in yeast mitochondria improves the production of branched-chain alcohols. *Nat. Biotechnol.* **31**: 335–341.

Baker, S.E. (2006) *Aspergillus niger* genomics: Past, present and into the future. *Med. Mycol.* **44**: 17–21.

Benedetti, I.M., de Lorenzo, V., and Silva-Rocha, R. (2012) Quantitative, Non-Disruptive Monitoring of Transcription in Single Cells with a Broad-Host Range GFP-luxCDABE Dual Reporter System. *PLoS One* **7**:.

Bhat, M.K. (2000) Cellulases and related enzymes in biotechnology. *Biotechnol. Adv.* **18**: 355–383.

Bischof, R.H., Ramoni, J., and Seiboth, B. (2016) Cellulases and beyond: The first 70 years of the enzyme producer *Trichoderma reesei*. *Microb. Cell Fact.* **15**: 1–13.

Bitinaite, J., Rubino, M., Varma, K.H., Schildkraut, I., Vaisvila, R., and Vaiskunaite, R. (2007) USER friendly DNA engineering and cloning method by uracil excision. *Nucleic Acids Res.* **35**: 1992–2002.

Blazeck, J., Garg, R., Reed, B., and Alper, H.S. (2012) Controlling promoter strength and regulation in *Saccharomyces cerevisiae* using synthetic hybrid promoters. *Biotechnol. Bioeng.* **109**: 2884–2895.

Blazeck, J., Liu, L., Redden, H., and Alper, H. (2011) Tuning gene expression in *Yarrowia lipolytica* by a hybrid promoter approach. *Appl. Environ. Microbiol.* **77**: 7905–7914.

Borodina, I. and Nielsen, J. (2014) Advances in metabolic engineering of yeast *Saccharomyces cerevisiae* for production of chemicals. *Biotechnol. J.* **9**: 609–620.

Bredeweg, E.L., Pomraning, K.R., Dai, Z., Nielsen, J., Kerkhoven, E.J., and Baker, S.E. (2017) A molecular genetic toolbox for *Yarrowia lipolytica*. *Biotechnol. Biofuels* **10**: 1–22.

Cameron, J.R., Philippsen, P., and Davis, R.W. (1983) Analysis of chromosomal integration and deletions of yeast plasmids. *Nucleic Acids Res.* **1**: 1429–1448.

Campbell, K., Xia, J., and Nielsen, J. (2017a) The Impact of Systems Biology on Bioprocessing. 1156–1168.

Campbell, K., Xia, J., and Nielsen, J. (2017b) The Impact of Systems Biology on Bioprocessing. *Trends Biotechnol.* **35**: 1156–1168.

Carrero, L.L., Nino-Vega, G., Teixeira, M.M., Carvalho, M.J.A., Soares, C.M.A., Pereira, M., et al. (2008) New *Paracoccidioides brasiliensis* isolate reveals unexpected genomic variability in this human pathogen. *Fungal Genet. Biol.* **45**: 605–612.

Chee, M.K. and Haase, S.B. (2012) New and Redesigned pRS Plasmid Shuttle Vectors for Genetic Manipulation of Saccharomyces cerevisiae. *Genes|Genomes|Genetics* **2**: 515–526.

Christianson, T.W., Sikorski, R.S., Dante, M., Shero, J.H., and Hieter, P. (1992) Multifunctional yeast high-copy-number shuttle vectors. *Gene* **110**: 119–122.

Clarke, L. and Carbon, J. (1980) Isolation of a yeast centromere and construction of functional small circular chromosomes. *Nature* **287**:.

Colussi, P.A. and Taron, C.H. (2005) *Kluyveromyces lactis* LAC4 promoter variants that lack function in bacteria but retain full function in K. lactis. *Appl. Environ. Microbiol.* **71**: 7092–7098.

Courbet, A., Endy, D., Renard, E., Molina, F., and Bonnet, J. (2015) Detection of pathological biomarkers in human clinical samples via amplifying genetic switches and logic gates. *Sci. Transl. Med.* **7**:.

Darvishi, F., Ariana, M., Marella, E.R., and Borodina, I. (2018) Advances in synthetic biology of oleaginous yeast *Yarrowia lipolytica* for producing non-native chemicals. *Appl. Microbiol. Biotechnol.* **102**: 5925–5938.

Din, M.O., Danino, T., Prindle, A., Skalak, M., Selimkhanov, J., Allen, K., et al. (2016) Synchronized cycles of bacterial lysis for in vivo delivery. *Nature* **536**: 81–5.

Ellis, T., Adie, T., Baldwin, G.S., Andrianantoandro, E., Basu, S., Karig, D.K., et al. (2011) DNA assembly for synthetic biology: from parts to pathways and beyond. *Integr. Biol.* **3**: 109.

Fang, F., Salmon, K., Shen, M.W.Y., Aeling, A.K., Ito, E., Irwin, B., et al. (2011) A vector set for systematic metabolic engineering in *Saccharomyces cerevisae*. *Yeast* **28**: 123–136.

Ferguson, J., Groppe, J.C., and Reed, S.I. (1981) Construction and characterization of three yeast-*Escherichia coli* shuttle vectors designed for rapid subcloning of yeast genes on small DNA fragments. *Gene* **16**: 191–197.

Fernandez-Abalos, J.M., Fox, H., Pitt, C., Wells, B., and Doonan, J.H. (1998) Plant-adapted green fluorescent protein is a versatile vital reporter for gene expression, protein localization and mitosis in the filamentous fungus, *Aspergillus nidulans*. *Mol. Microbiol.* **27**: 121–130.

Florencio, C., Cunha, F.M., Badino, A.C., Farinas, C.S., Ximenes, E., and Ladisch, M.R. (2016) Secretome analysis of *Trichoderma reesei* and *Aspergillus niger* cultivated by submerged and sequential fermentation processes: Enzyme production for sugarcane bagasse hydrolysis. *Enzyme Microb. Technol.* **90**: 53–60.

Gavrilescu, M., Demnerová, K., Aamand, J., Agathos, S., and Fava, F. (2015) Emerging pollutants in the environment: Present and future challenges in biomonitoring, ecological risks and bioremediation. *N. Biotechnol.* **32**: 147–156.

Gegembauer, G., Araujo, L.M., Pereira, E.F., Rodrigues, A.M., Mello, A., Paniago,

M., et al. (2014) Serology of Paracoccidioidomycosis Due to *Paracoccidioides lutzii*. *PLoS Negl. Trop. Dis.* **8**:.

Gietz, R.D. (2014) Yeast transformation by the LiAc/SS carrier DNA/PEG method. *Methods Mol. Biol.* **1163**: 33–44.

Gietz, R.D. and Sugino, A. (1988) New yeast-*Escherichia coli* shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites. *Gene* **74**: 527–534.

Glass, N.L., Schmoll, M., Cate, J.H.D., and Coradetti, S. (2013) Plant cell wall deconstruction by ascomycete fungi. *Annu. Rev. Microbiol.* **67**: 477–98.

Gnügge, R., Liphardt, T., and Rudolf, F. (2016) A shuttle vector series for precise genetic engineering of *Saccharomyces cerevisae*. *Yeast* **33**: 83–98.

Gnügge, R. and Rudolf, F. (2017) *Saccharomyces cerevisiae* Shuttle vectors. *Yeast* **26**: 545–551.

Goodman, D.B., Church, G.M., and Kosuri, S. (2013) Causes and Effects of N-Terminal Codon Bias in Bacterial Genes. *Science (80-.).* **342**: 475–479.

Gressler, M., Hortschansky, P., Geib, E., and Brock, M. (2015) A new high-performance heterologous fungal expression system based on regulatory elements from the *Aspergillus terreus* terrein gene cluster. *Front. Microbiol.* **6**:.

Guerineau, M., Grandchamp, C., Paoletti, C., and Slonimski, P. (1971) Characterization of a new class of circular DNA molecules in yeast. *Biochem. Biophys. Res. Commun.* **42**: 550–557.

Guerineau, M., Slonimski, P.P., and Avner, P.R. (1974) Yeast Episome: Oligomycin resistance associated with a small covalently closed non-mitochondrial circular DNA. *Biochem. Biophys. Res. Commun.* **61**:.

Hahn, R.C., Rodrigues, A.M., Jesus, C., Fontes, F., Nery, A.F., and Tadano, T. (2014) Case Report : Fatal Fungemia due to *Paracoccidioides lutzii*. **91**: 394–398.

Hamer, J.E. and Timberlake, W.E. (1987) Functional organization of the Aspergillus nidulans trpC promoter. *Mol. Cell. Biol.* **7**: 2352–2359.

Hartner, F.S., Ruth, C., Langenegger, D., Johnson, S.N., Hyka, P., Lin-Cereghino, G.P., et al. (2008) Promoter library designed for fine-tuned gene expression in *Pichia pastoris*. *Nucleic Acids Res.* **36**: 1–15.

He, R., Zhang, C., Guo, W., Wang, L., Zhang, D., and Chen, S. (2013) Construction of a plasmid for heterologous protein expression with a constitutive promoter in *Trichoderma reesei*. *Plasmid* **70**: 425–429.

Hinnen, A., Hicks, J.B., and Fink, G.R. (1978) Transformation of yeast. *Proc. Natl. Acad. Sci.* **75**: 1929–1933.

Holkenbrink, C., Dam, M.I., Kildegaard, K.R., Beder, J., Dahlin, J., Doménech Belda, D., and Borodina, I. (2018) EasyCloneYALI: CRISPR/Cas9-Based Synthetic Toolbox for Engineering of the Yeast *Yarrowia lipolytica*. *Biotechnol. J.* **1700543**: 1–8.

Hooykaas, P.J.J., Heusden, G.P.H. van, Niu, X., Roushan, M.R., Soltani, J., Zhang, X., and Zaal, B.J. van der (2018) *Agrobacterium*-Mediated Transformation of Yeast and Fungi. In, *Current Topics in Microbiology and Immunology.* Springer, Berlin, Heidelberg, pp. 1–26.

Jensen, N.B., Strucko, T., Kildegaard, K.R., David, F., Maury, J., Mortensen, U.H., et al. (2014) EasyClone: Method for iterative chromosomal integration of multiple genes in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **14**: 238–248.

Johns, A.M.B., Love, J., and Aves, S.J. (2016) Four inducible promoters for controlled gene expression in the oleaginous yeast *Rhodotorula toruloides*. *Front. Microbiol.* **7**: 1–12.

Jørgensen, M.S., Skovlund, D.A., Johannesen, P.F., and Mortensen, U.H. (2014) A novel platform for heterologous gene expression in *Trichoderma reesei* (Teleomorph *Hypocrea jecorina*). *Microb. Cell Fact.* **13**: 1–9.

Jullesson, D., David, F., Pfleger, B., and Nielsen, J. (2015) Impact of synthetic biology and metabolic engineering on industrial production of fine chemicals. *Biotechnol. Adv.* **33**: 1395–1402.

Kawai, S., Hashimoto, W., and Murata, K. (2010) Transformation of *Saccharomyces cerevisiae* and other fungi: methods and possible underlying mechanism. *Bioeng. Bugs* **1**: 395–403.

Kim, J., Salvador, M., Saunders, E., González, J., Avignone-Rossa, C., and Jiménez, J.I. (2016) Properties of alternative microbial hosts used in synthetic biology: towards the design of a modular chassis. *Essays Biochem.* **60**: 303–313.

Kircher, M. (2015) Sustainability of biofuels and renewable chemicals production from biomass. *Curr. Opin. Chem. Biol.* **29**: 26–31.

Kontoyiannis, D.P. and Lewis, R.E. (2002) Antifungal drug resistance of pathogenic fungi. **359**: 1135–1144.

Kosuri, S., Goodman, D.B., Cambray, G., Mutalik, V.K., and Gao, Y. (2013) Composability of regulatory sequences controlling transcription and translation

in *Escherichia coli.* **110**:.

Küberl, A., Schneider, J., Thallinger, G.G., Anderl, I., Wibberg, D., Hajek, T., et al. (2011) High-quality genome sequence of *Pichia pastoris* CBS7435. *J. Biotechnol.* **154**: 312–320.

Kubicek, C.P., Mikus, M., Schuster, A., Schmoll, M., and Seiboth, B. (2009) Metabolic engineering strategies for the improvement of cellulase production by *Hypocrea jecorina*. *Biotechnol. Biofuels* **2**: 19.

Lambertz, C., Eckert, C., Fischer, R., and Commandeur, U. (2016) *Kluyveromyces lactis* as an Expression Host for Enzymes that Degrade Lignocellulosic Biomass. *Cellul. Chem. Technol* **50**: 3–4.

Leal, C. V, Montes, B.A., Mesa, A.N.A.C., Rua, A.L., Corredor, M., Restrepo, A., and Mcewen, J.G. (2004) *Agrobacterium tumefaciens*-mediated transformation of *Paracoccidioides brasiliensis*. 391–395.

Lee, M.E., Aswani, A., Han, A.S., Tomlin, C.J., and Dueber, J.E. (2013) Expression-level optimization of a multi-enzyme pathway in the absence of a high-throughput assay. *Nucleic Acids Res.* **41**: 10668–10678.

Lee, M.E., DeLoache, W.C., Cervantes, B., and Dueber, J.E. (2015) A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth. Biol.* **4**: 975–986.

Leigh, J., Seif, E., Rodriguez, N., Jacob, Y., and Lang, B.F. (2003) Fungal Evolution meets Fungal Genomics. *Handb. Fungal Biotechonology* 145–161.

Leplat, C., Nicaud, J.M., and Rossignol, T. (2015) High-throughput transformation method for *Yarrowia lipolytica* mutant library screening. *FEMS Yeast Res.* **15**: 1–9.

Lin, X., Wang, Y., Zhang, S., Zhu, Z., Zhou, Y.J., Yang, F., et al. (2014) Functional integration of multiple genes into the genome of the oleaginous yeast *Rhodosporidium toruloides*. *FEMS Yeast Res.* **14**: 547–555.

Liu, G., Lanham, C., Buchan, J.R., and Kaplan, M.E. (2017) High-throughput transformation of *Saccharomyces cerevisiae* using liquid handling robots. *PLoS One* **12**: 1–15.

Liu, H., Jiao, X., Wang, Y., Yang, X., Sun, W., Wang, J., et al. (2017) Fast and efficient genetic transformation of oleaginous yeast *Rhodosporidium toruloides* by using electroporation. *FEMS Yeast Res.* **17**: 1–11.

Liu, Y., Koh, C.M.J., Sun, L., Hlaing, M.M., Du, M., Peng, N., and Ji, L. (2013)

Characterization of glyceraldehyde-3-phosphate dehydrogenase gene RtGPD1 and development of genetic transformation method by dominant selection in oleaginous yeast *Rhodosporidium toruloides*. *Appl. Microbiol. Biotechnol.* **97**: 719–729.

Liu, Y., Yap, S.A., Koh, C.M.J., and Ji, L. (2016) Developing a set of strong intronic promoters for robust metabolic engineering in oleaginous *Rhodotorula* (*Rhodosporidium*) yeast species. *Microb. Cell Fact.* **15**: 1–9.

Lv, D., Wang, W., and Wei, D. (2012a) Construction of two vectors for gene expression in *Trichoderma reesei*. *Plasmid* **67**: 67–71.

Lv, D., Wang, W., and Wei, D. (2012b) Construction of two vectors for gene expression in *Trichoderma reesei*. *Plasmid* **67**: 67–71.

Ma, H., Kunes, S., Schatz, P.J., and Botstein, D. (1987) Plasmid construction by linker-assisted homologous recombination in yeast. *Gene* **58**: 201–216.

Magee, P.T., Gale, C., Berman, J., and Davis, D. (2003) MINIREVIEW Molecular Genetic and Genomic Approaches to the Study of Medically Important Fungi. **71**: 2299–2309.

Margulis, L. and Chapman, M.J. (1998) Endosymbioses: Cyclical and permanent in evolution. *Trends Microbiol.* **6**: 342–345.

Markham, K.A., Vazquez, S., and Alper, H.S. (2018) High efficiency transformation of *Yarrowia lipolytica* using electroporation.

Marques-da-silva, S.H., Rodrigues, A.M., Hoog, G.S. De, and Camargo, Z.P. De (2012) Case Report : Occurrence of *Paracoccidioides lutzii* in the Amazon Region : Description of Two Cases. **87**: 710–714.

Martinez, D., Berka, R.M., Henrissat, B., Saloheimo, M., Arvas, M., Baker, S.E., et al. (2008) Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nat. Biotechnol.* **26**: 553–560.

Martins-Santana, L., Nora, L.C., Sanches-Medeiros, A., Lovate, G.L., Cassiano, M.H.A., and Silva-Rocha, R. (2018) Systems and Synthetic Biology Approaches to Engineer Fungi for Fine Chemical Production. *Front. Bioeng. Biotechnol.* **6**: 1–18.

Matute, D.R., Mcewen, J.G., Puccia, R., Montes, B.A., San-blas, G., Bagagli, E., et al. (2004) Cryptic Speciation and Recombination in the Fungus Paracoccidioides brasiliensis as Revealed by Gene Genealogies.

Maury, J., Germann, S.M., Baallal Jacobsen, S.A., Jensen, N.B., Kildegaard, K.R.,

Herrgàrd, M.J., et al. (2016) EasyCloneMulti: A set of vectors for simultaneous and multiple genomic integrations in *Saccharomyces cerevisiae*. *PLoS One* **11**: 1–22.

Menino, J.F., Almeida, A.J., and Rodrigues, F. (2012) Gene Knockdown in *Paracoccidioides brasiliensis* Using Antisense RNA. In, Brand,A.C. and MacCallum,D.M. (eds), *Host-Fungus Interactions: Methods and Protocols*. Humana Press, Totowa, NJ, pp. 187–198.

Michielse, C.B., J Hooykaas, P.J., J J van den Hondel, C.A.M., and J Ram, A.F. (2008) *Agrobacterium*-mediated transformation of the filamentous fungus *Aspergillus awamori*. *Nat. Protoc.* **3**: 1671–1678.

Mikkelsen, M.D., Buron, L.D., Salomonsen, B., Olsen, C.E., Hansen, B.G., Mortensen, U.H., and Halkier, B.A. (2012) Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. *Metab. Eng.* **14**: 104–111.

Mutalik, V.K., Guimaraes, J.C., Cambray, G., Lam, C., Christoffersen, M.J., Mai, Q., et al. (2013) Precise and reliable gene expression via standard transcription and translation initiation elements. **10**:.

Näätsaari, L., Mistlberger, B., Ruth, C., Hajek, T., Hartner, F.S., and Glieder, A. (2012) Deletion of the pichia pastoris ku70 homologue facilitates platform strain generation for gene expression and synthetic biology. *PLoS One* **7**:.

Nambu-Nishida, Y., Nishida, K., Hasunuma, T., and Kondo, A. (2017) Development of a comprehensive set of tools for genome engineering in a cold- A nd thermo-tolerant *Kluyveromyces marxianus* yeast strain. *Sci. Rep.* **7**: 1–7.

Neil, H., Malabat, C., Aubenton-carafa, Y., Xu, Z., Steinmetz, L.M., and Jacquier, A. (2009) Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature* **457**: 1038–1042.

Nguyen, K.T., Ho, Q.N., Pham, T.H., Phan, T.-N., and Tran, V.-T. (2016) The construction and use of versatile binary vectors carrying pyrG auxotrophic marker and fluorescent reporter genes for *Agrobacterium*-mediated transformation of *Aspergillus oryzae*. *World J. Microbiol. Biotechnol.* **32**: 204.

Nora, L.C., Westmann, C.A., Martins-Santana, L., Alves, L. de F., Monteiro, L.M.O., Guazzaroni, M.E., and Silva-Rocha, R. (2018) The art of vector engineering: towards the construction of next-generation genetic tools. *Microb. Biotechnol.*

Obst, U., Lu, T.K., and Sieber, V. (2017) A Modular Toolkit for Generating Pichia

pastoris Secretion Libraries. *ACS Synth. Biol.* **6**: 1016–1025.

Van Ooyen, A.J.J., Dekker, P., Huang, M., Olsthoorn, M.M.A., Jacobs, D.I., Colussi, P.A., and Taron, C.H. (2006) Heterologous protein production in the yeast Kluyveromyces lactis. *FEMS Yeast Res.* **6**: 381–392.

Pan, R., Zhang, J., Shen, W.L., Tao, Z.Q., Li, S.P., and Yan, X. (2011) Sequential deletion of *Pichia pastoris* genes by a self-excisable cassette. *FEMS Yeast Res.* **11**: 292–298.

Park, Y.K., Nicaud, J.M., and Ledesma-Amaro, R. (2018) The Engineering Potential of *Rhodosporidium toruloides* as a Workhorse for Biotechnological Applications. *Trends Biotechnol.* **36**: 304–317.

Partow, S., Siewers, V., Bjørn, S., Nielsen, J., and Maury, J. (2010) Characterization of different promoters for designing a new expression vector in *Sacharomyces cerevisiae. Yeast* **27**: 955–964.

Pasin, F., Bedoya, L.C., Bernab?-Orts, J.M., Gallo, A., Simón-Mateo, C., Orzaez, D., and García, J.A. (2017) Multiple T-DNA Delivery to Plants Using Novel Mini Binary Vectors with Compatible Replication Origins. *ACS Synth. Biol.* acssynbio.6b00354.

Perea, S. and Patterson, T.F. (2002) Antifungal Resistance in Pathogenic Fungi. **35**: 1073–1080.

de Pourcq, K., Tiels, P., van Hecke, A., Geysens, S., Vervecken, W., and Callewaert, N. (2012) Engineering *Yarrowia lipolytica* to produce glycoproteins homogeneously modified with the universal man3GlcNAc2 N-glycan core. *PLoS One* **7**:.

Purnick, P.E.M. and Weiss, R. (2009) The second wave of synthetic biology: From modules to systems. *Nat. Rev. Mol. Cell Biol.* **10**: 410–422.

Ribeiro, O., Gombert, A.K., Teixeira, J.A., and Domingues, L. (2007) Application of the Cre-loxP system for multiple gene disruption in the yeast *Kluyveromyces marxianus*. *J. Biotechnol.* **131**: 20–26.

Ruiz-Díez, B. (2002) A Review: Strategies for the transformation of filamentous fungi. *J. Appl. Microbiol.* **92**: 189–195.

De Schutter, K., Lin, Y.C., Tiels, P., Van Hecke, A., Glinka, S., Weber-Lehmann, J., et al. (2009) Genome sequence of the recombinant protein production host *Pichia pastoris. Nat. Biotechnol.* **27**: 561–566.

Scorzoni, L., Carolina, A., Paula, A. De, Lacorte, J. De, Leite, F.S., Oliveira, H.C. De,

et al. (2015) Comparison of virulence between *Paracoccidioides brasiliensis* and *Paracoccidioides lutzii* using *Galleria mellonella* as a host model. **5594**.

Sheldon, R.A. (2014) Green and sustainable manufacture of chemicals from biomass: state of the art. *Green Chem.* **16**: 950–963.

Sikorski, R.S. and Hieter, P. (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**: 19–27.

Silva-Rocha, R., Martínez-García, E., Calles, B., Chavarría, M., Arce-Rodríguez, A., De Las Heras, A., et al. (2013) The Standard European Vector Architecture (SEVA): A coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res.* **41**: 666–675.

Soccol, C.R., Dalmas Neto, C.J., Soccol, V.T., Sydney, E.B., da Costa, E.S.F., Medeiros, A.B.P., and Vandenberghe, L.P. de S. (2017) Pilot scale biodiesel production from microbial oil of *Rhodosporidium toruloides* DEBB 5533 using sugarcane juice: Performance in diesel engine and preliminary economic study. *Bioresour. Technol.* **223**: 259–268.

Spohner, S.C., Schaum, V., Quitmann, H., and Czermak, P. (2016) *Kluyveromyces lactis:* An emerging tool in biotechnology. *J. Biotechnol.* **222**: 104–116.

Steiger, M.G., Vitikainen, M., Uskonen, P., Brunner, K., Adam, G., Pakula, T., et al. (2011) Transformation system for *Hypocrea jecorina* (*Trichoderma reesei)* that favors homologous integration and employs reusable bidirectionally selectable markers. *Appl. Environ. Microbiol.* **77**: 114–121.

Sthol, L.L. and Lambowitz, A.M. (1983) Construction of a Shuttle Vector for the Filamentous Fungus *Neurospora crassa*. Proceedings of the National Academy of Sciences of the United States of America, **80**: 1058–1062.

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

Stovicek, V., Borja, G.M., Forster, J., and Borodina, I. (2015) EasyClone 2.0: expanded toolkit of integrative vectors for stable gene expression in industrial *Saccharomyces cerevisiae* strains. *J. Ind. Microbiol. Biotechnol.* **42**: 1519–1531.

Struhl, K., Stinchcomb, D.T., Scherer, S., and Davis, R.W. (1979) High-frequency transformation of yeast: autonomous replication of hybrid DNA molecules. *Proc. Natl. Acad. Sci.* **76**: 1035–1039.

Teixeira, M.D.M., Theodoro, R.C., Oliveira, F.F.M. De, Machado, G.C., Hahn, R.C., Bagagli, E., et al. (2013) *Paracoccidioides lutzii* sp. nov.: biological and clinical implications. *Med. Mycol.* 1–10.

Teixeira, M.M., Theodoro, R.C., Carvalho, M.J.A. De, Fernandes, L., Paes, H.C., Hahn, R.C., et al. (2009) Molecular Phylogenetics and Evolution Phylogenetic analysis reveals a high level of speciation in the *Paracoccidioides genus*. *Mol. Phylogenet. Evol.* **52**: 273–283.

Theodoro, R.C., Bagagli, E., and Oliveira, C. (2008) Phylogenetic analysis of PRP8 intein in *Paracoccidioides brasiliensis* species complex. **45**: 1284–1291.

Tsai, Y.Y., Ohashi, T., Kanazawa, T., Polburee, P., Misaki, R., Limtong, S., and Fujiyama, K. (2017a) Development of a sufficient and effective procedure for transformation of an oleaginous yeast, *Rhodosporidium toruloides* DMKU3-TK16. *Curr. Genet.* **63**: 373.

Tsai, Y.Y., Ohashi, T., Kanazawa, T., Polburee, P., Misaki, R., Limtong, S., and Fujiyama, K. (2017b) Erratum to: Development of a sufficient and effective procedure for transformation of an oleaginous yeast, *Rhodosporidium toruloides* DMKU3-TK16. *Curr. Genet.* **63**: 373.

Vogl, T., Gebbie, L., Palfreyman, R.W., and Speight, R. (2018) Effect of plasmid design and type of integration event on recombinant protein expression in *Pichia pastoris*. *Appl. Environ. Microbiol.* **84**:.

Vogl, T. and Glieder, A. (2013) Regulation of *Pichia pastoris* promoters and its consequences for protein production. *N. Biotechnol.* **30**: 385–404.

Vogl, T., Hartner, F.S., and Glieder, A. (2013) New opportunities by synthetic biology for biopharmaceutical production in *Pichia pastoris*. *Curr. Opin. Biotechnol.* **24**: 1094–1101.

Vogl, T., Kickenweiz, T., Pitzer, J., Sturmberger, L., Weninger, A., Biggs, B.W., et al. (2018) Engineered bidirectional promoters enable rapid multi-gene co-expression optimization. *Nat. Commun.* **9**: 3589.

Wagner, J.M. and Alper, H.S. (2016) Synthetic biology and molecular genetics in non-conventional yeasts: Current tools and future advances. *Fungal Genet. Biol.* **89**: 126–136.

Wang, Y., Lin, X., Zhang, S., Sun, W., Ma, S., and Zhao, Z.K. (2016) Cloning and evaluation of different constitutive promoters in the oleaginous yeast *Rhodosporidium toruloides*. *Yeast* **33**: 99–106.

Wang, Y., Zhang, S., Zhu, Z., Shen, H., Lin, X., Jin, X., et al. (2018) Biotechnology for Biofuels Systems analysis of phosphate - limitation - induced lipid accumulation by the oleaginous yeast *Rhodosporidium toruloides*. *Biotechnol. Biofuels* 1–15.

Ward, O.P. (2012) Production of recombinant proteins by filamentous fungi. *Biotechnol. Adv.* **30**: 1119–1139.

Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011) A modular cloning system for standardized assembly of multigene constructs. *PLoS One* **6**: e16765.

Westermann, S., Drubin, D.G., and Barnes, G. (2007) Structures and Functions of Yeast Kinetochore Complexes. *Annu. Rev. Biochem.* **76**: 563–591.

Westmann, C.A., Guazzaroni, M.-E., and Silva-Rocha, R. (2018) Engineering Complexity in Bacterial Regulatory Circuits for Biotechnological Applications. *mSystems* **3**: e00151-17.

Xu, Z., Camblong, J., Wei, W., Gagneur, J., Perocchi, F., Clauder-mu, S., et al. (2009) Bidirectional promoters generate pervasive transcription in yeast. *Nature* **457**:.

Xuan, Y., Zhou, X., Zhang, W., Zhang, X., Song, Z., and Zhang, Y. (2009) An upstream activation sequence controls the expression of AOX1 gene in *Pichia pastoris*. *FEMS Yeast Res.* **9**: 1271–1282.

Yaegashi, J., Kirby, J., Ito, M., Sun, J., Dutta, T., Mirsiaghi, M., et al. (2017) *Rhodosporidium toruloides*: A new platform organism for conversion of lignocellulose into terpene biofuels and bioproducts. *Biotechnol. Biofuels* **10**: 1–13.

Yamanishi, M., Ito, Y., Kintaka, R., Imamura, C., Katahira, S., Ikeuchi, A., et al. (2013) A genome-wide activity assessment of terminator regions in *Saccharomyces cerevisiae* provides a “terminatome” toolbox. *ACS Synth. Biol.* **2**: 337–347.

Yang, C., Hu, S., Zhu, S., Wang, D., Gao, X., and Hong, J. (2015) Characterizing yeast promoters used in *Kluyveromyces marxianus*. *World J. Microbiol. Biotechnol.* **31**: 1641–1646.

Yang, F., Zhang, S., Tang, W., and Zhao, Z.K. (2008) Identification of the orotidine-5′-monophosphatedecarboxylase gene of the oleaginous yeast *Rhodosporidium toruloides. Yeast* **25**: 623–630.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Improved Ml3 phage cloning

vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103–119.

Yin, X., Shin, H.D., Li, J., Du, G., Liu, L., and Chen, J. (2017) Pgas, a low-pH-induced promoter, as a tool for dynamic control of gene expression for metabolic engineering of *Aspergillus niger*. *Appl. Environ. Microbiol.* **83**:.

Yuan, J. and Ching, C.B. (2016) Mitochondrial acetyl-CoA utilization pathway for terpenoid productions. *Metab. Eng.* **38**: 303–309.

Zhang, L., Zhang, S., Jiang, X., Wei, W., Wang, W., and Wei, D. (2016) A novel host-vector system for heterologous protein co-expression and purification in the *Trichoderma reesei* industrial strain RUT-C30. *Biotechnol. Lett.* **38**: 89–96.

Zhang, S., Ito, M., Skerker, J.M., Arkin, A.P., and Rao, C. V. (2016) Metabolic engineering of the oleaginous yeast *Rhodosporidium toruloides* IFO0880 for lipid overproduction during high-density fermentation. *Appl. Microbiol. Biotechnol.* **100**: 9393–9405.

Zhong, Y.H., Wang, X.L., Wang, T.H., and Jiang, Q. (2007) *Agrobacterium*-mediated transformation (AMT) of *Trichoderma reesei* as an efficient tool for random insertional mutagenesis. *Appl. Microbiol. Biotechnol.* **73**: 1348–1354.

Zhu, T., Wang, W., Yang, X., Wang, K., and Cui, Z. (2009) Construction of two Gateway vectors for gene expression in fungi. *Plasmid* **62**: 128–133.